



irc 2022
XVI. international research conference
proceedings

open science index 16 2022

june 02-03, 2022 san francisco united states
international scholarly and scientific research & innovation



Open Science

Open Science Philosophy

Open science encompasses unrestricted access to scientific research articles, access to data from public research, and collaborative research enabled by information and communication technology tools, models, and incentives. Broadening access to scientific research publications and data is at the heart of open science. The objective of open science is to make research outputs and its potential benefits available to the entire world and in the hands of as many as possible:

- Open science promotes a more accurate verification of scientific research results. Scientific inquiry and discovery can be sped up by combining the tools of science and information technologies. Open science will benefit society and researchers by providing faster, easier, and more efficient availability of research outputs.
- Open science reduces duplication in collecting, creating, transferring, and re-using scientific material.
- Open science increases productivity in an era of tight budgets.
- Open science results in great innovation potential and increased consumer choice from public research.
- Open science promotes public trust in science. Greater citizen engagement leads to active participation in scientific experiments and data collection.

Open Science Index

The Open Science Index (OSI) currently provides access to over thirty thousand full-text journal articles and is working with member and non-member organizations to review policies to promote and assess open science. As part of the open science philosophy, and by making open science a reality; OSI is conducting an assessment of the impact of open science principles and restructuring the guidelines for access to scientific research. As digitalization continues to accelerate science, Open science and big data hold enormous promise and present new challenges for policymakers, scientific institutions, and individual researchers.

OSI is helping the global scientific research community discover, evaluate, and access high-quality research output. Renowned for its editorially curated and refereed collection of the highest-quality publications, OSI has always been and will remain free-of-charge.

OSI provides an efficient and thorough discovery process to the open science research database and provides links and free access to full-text articles. There are 50 open access journal categories that are curated and refereed by international scientific committees, the in-house editorial team, and trusted partners. Since its inception in 2007, OSI has made more than thirty thousand peer-reviewed open access full-text journal articles (PDF versions) freely available online without cost, barriers, or restrictions.

Open Science Access

With the Open Science Index, researchers can discover and access trusted peer-reviewed open access full-text scientific research articles with confidence. OSI helps researchers find appropriate non-profit open access journals to publish their work.

OSI gives one-click access to online full-text PDFs and expands the reach to global society by giving users free access from anywhere around the globe. Through cutting-edge open science collaboration, in an innovative public partnership, the non-profit OSI is devoted to making science open and reusable.

To learn more, visit online at waset.org

Open Science

Open Society

An open society allows individuals to change their roles and to benefit from corresponding changes in status. Open science depends to a greater or lesser extent on digital technologies and innovations in structural processes by an open society. When realized, open science research and innovation can create investment opportunities for new and better products and services and therefore increase competitiveness and employment. Open science research and innovation is a key component of thematic open science priorities. Central to the open science digital infrastructure is enabling industry to benefit from digital technology and to underpin scientific advances through the development of an open society. Open science research and innovation can also contribute to society as a global actor because scientific relations can flourish even where global relations are strained. Open science has a critical role across many areas of decision making in providing evidence that helps understand the risks and benefits of different open science choices. Digital technology is making the conduct of open science and innovation more collaborative, more global, and more open to global citizens. Open society must embrace these changes and reinforce its position as the leading power for science, for new ideas, and for investing sustainably in the future.

It is apparent in open society that the way science works is fundamentally changing, and an equally significant transformation is taking place in how organizations and societies innovate. The advent of digital technology is making research and innovation more open, collaborative, and global. These exchanges are leading open society to develop open science and to set goals for research and innovation priority. Open science goals are materializing in the development of scientific research and innovation platforms and greater acceptance of scientific data generated by open science research. Open science research and innovation do not need help from open society to come up with great ideas, but the level of success ideas ultimately reach is undoubtedly influenced by regulation, financing, public support, and market access. Open society is playing a crucial role in improving all these success factors.

Open Science

Open science represents a new approach to the scientific process based on cooperative work and new ways of diffusing knowledge by using digital technologies and collaborative tools. These innovations capture a systemic change to the way science and research have been carried out for the last fifty years. Science is shifting from the standard practice of publishing research results in scientific publications after the research and reviews are completed. The shift is towards sharing and using all available knowledge at an earlier stage in the research process. Open science is to science what digital technology is to social and economic transactions: allowing end users to be producers of ideas, relations, and services and in doing so, enabling new working models, new social relationships and leading to a new modus operandi for science. Open science is as important and disruptive as e-commerce has been for the retail industry. Just like e-commerce, the open science research paradigm shift affects the whole business cycle of doing science and research. From the selection of research subjects to the carrying out of research, to its use and re-use, to the role of universities, and that of publishers are all dramatically changed. Just as the internet and globalization have profoundly changed the way we do business, interact socially, consume culture, and buy goods, these changes are now profoundly impacting how one does research and science.

The discussion on broadening the footprint of science and on novel ways to produce and spread knowledge gradually evolved from two global trends: Open Access and Open Source. The former refers to online, peer-reviewed scholarly outputs, which are free to read, with limited or no copyright and licensing restrictions, while open source refers to software created without any proprietary restriction and which can be accessed and freely used. Although open access became primarily associated with a particular publishing

Open Science

or scientific dissemination practice, open access already sought to induce a broader practice that includes the general re-use of all kinds of research products, not just publications or data. It is only more recently that open science has coalesced into the concept of a transformed scientific practice, shifting the focus of researchers' activity from publishing as fast as possible to sharing knowledge as early as possible. Open science is defined as the idea that scientific knowledge of all kinds should be openly shared as early as is practical in the discovery process. As a result, the way science is done in the future will look significantly different from the way it is done now. Open science is the ongoing evolution in the modus operandi of doing research and organizing science. This evolution is enabled by digital technology and is driven by both the globalization of the scientific community and increasing public demand to address the societal challenges of our times. Open science entails the ongoing transitions in the way research is performed, researchers collaborate, knowledge is shared, and science is organized.

Open science impacts the entire research cycle, from the inception of research to its publication, and on how this cycle is organized. The outer circle reflects the new interconnected nature of open science, while the inner circle shows the entire scientific process, from the conceptualization of research ideas to publishing. Each step in the scientific process is linked to ongoing changes brought about by open science, including the emergence of alternative systems to establish a scientific reputation; changes in the way quality and impact of research are evaluated; the growing use of scientific blogs; open annotation; and open access to data and publications. All institutions involved in science are affected, including research organizations, research councils, and funding bodies. The trends are irreversible, and they have already grown well beyond individual projects. These changes predominantly result from a bottom-up process driven by a growing number of researchers who increasingly employ social media in their research and initiate globally coordinated research projects while sharing results at an early stage in the research process.

Open science is encompassed in five schools of thought:

- the infrastructure school, concerned with technological architecture
- the public school, concerned with the accessibility of knowledge creation
- the measurement school, concerned with alternative impact assessment
- the democratic school, concerned with access to knowledge
- the pragmatic school, concerned with collaborative research

According to the measurement school, the reputation and evaluation of individual researchers are still mainly based on citation-based metrics. The h-index is an author-level metric that attempts to measure both the productivity and citation impact of the publications of a scientist or scholar. The impact factor is a measure reflecting the average number of citations to articles published in an academic journal and is used as a proxy for the relative importance of a journal.

Numerous criticisms have been made of citation-based metrics, primarily when used, and often misused, to assess the performance of individual researchers. These metrics:

- are often not applicable at the individual level
- do not take into account the broader social and economic function of scientific research
- are not adapted to the increased scale of research
- cannot recognize new types of work that researchers are performing

Web-based metrics for measuring research output, popularized as altmetrics, have recently received much attention: some measure the impact at the article level, others make it possible to assess the many outcomes of research in addition to the number of scientific articles and references. The current reputation and evaluation system has to adapt to the new dynamics of open science and acknowledge and incentivize

Open Science

engagement in open science. Researchers engaging in open science have growing expectations that their work, including intermediate products such as research data, will be better rewarded or taken into account in their career development. Vice-versa, the use, and reuse of open data will require appropriate codes of conduct requiring, for example, the proper acknowledgment of the original creator of the data.

These ongoing changes are progressively transforming scientific practices with innovative tools to facilitate communication, collaboration, and data analysis. Researchers that increasingly work together to create knowledge can employ online tools and create a shared space where creative conversation and collaboration can occur. As a result, the problem-solving process can be faster, and the range of problems that can be solved can be expanded. The ecosystem underpinning open science is evolving very rapidly. Social network platforms for researchers already attract millions of users and are being used to begin and validate more research projects.

Furthermore, the trends towards open access are redefining the framework conditions for science and thus have an impact on how open innovation is produced by encouraging a more dynamic circulation of knowledge. It can enable more science-based startups to emerge thanks to the exploitation of openly accessible research results. Open science, however, does not mean free science. It is essential to ensure that intellectual property is protected before making knowledge publicly available in order to subsequently attract investments that can help translate research results into innovation. If this is taken into account, fuller and broader access to scientific publications and research data can help to accelerate innovation. Investments that boost research and innovation in open science would benefit society with fewer barriers to knowledge transfer, open access to scientific research, and greater mobility of researchers. In this context, open access can help overcome the barriers that innovative organizations face in accessing the results of research funded by the public.

Open innovation

An open society is the largest producer of knowledge, but the phenomenon of open science is changing every aspect of the scientific method by becoming more open, inclusive, and interdisciplinary. Ensuring open society is at the forefront of open science means promoting open access to scientific data and publications alongside the highest standards of research integrity. There are few forces in this globe as engaging and unifying as science. The universal language of science maintains open channels of communication globally. Open society can maximize its gains through maintaining its presence at the highest level of scientific endeavor, and by promoting a competitive edge in the knowledge society of the information age. The ideas and initiatives described in this publication can stimulate anyone interested in open science research and innovation. It is designed to encourage debate and lead to new ideas on what and open society should do, should not do, or do differently.

An open society can lead to a research powerhouse; however, open society rarely succeeds in turning research into innovation and in getting research results to the global market. Open society must improve at making the most of its innovation talent, and that is where open innovation comes into play. The basic premise of open innovation is to open up the innovation process to all active players so that knowledge can circulate more freely and be transformed into products and services that create new markets while fostering a stronger culture of entrepreneurship. Open innovation is defined as the use of purposive inflows and outflows of knowledge to accelerate internal innovation. This original notion of open innovation was primarily based on transferring knowledge, expertise, and even resources from one company or research institution to another. This notion assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as they seek to improve their performance. The concept of open innovation is continually evolving and is moving from linear, bilateral transactions and collaborations

Open Science

towards dynamic, networked, multi-collaborative innovation ecosystems. This means that a specific innovation can no longer be seen as the result of predefined and isolated innovation activities but rather as the outcome of a complex co-creation process involving knowledge flows across the entire economic and social environment. This co-creation takes place in different parts of the innovation ecosystem and requires knowledge exchange and absorptive capacities from all the actors involved, whether businesses, academia, financial institutions, public authorities, or citizens.

Open innovation is a broad term, which encompasses several different nuances and approaches. Two main elements underpin the most recent conceptions of open innovation: the users are in the spotlight and invention becomes an innovation only if users become a part of the value creation process. Notions such as user innovation emphasize the role of citizens and users in the innovation processes as distributed sources of knowledge. This kind of public engagement is one of the aims of open science research and innovation. The term 'open' in these contexts has also been used as a synonym for 'user-centric'; creating a well-functioning ecosystem that allows co-creation and becomes essential for open innovation. In this ecosystem, relevant stakeholders are collaborating along and across industry and sector-specific value chains to co-create solutions for socio-economic and business challenges. One important element to keep in mind when discussing open innovation is that it cannot be defined in absolutely precise terms. It may be better to think of it as a point on a continuum where there is a range of context-dependent innovation activities at different stages, from research to development through to commercialization, and where some activities are more open than others. Open innovation is gaining momentum thanks to new large-scale trends such as digitalization and the mass participation and collaboration in innovation that it enables. The speed and scale of digitalization are accelerating and transforming the way one designs, develops, and manufactures products, the way one delivers services, and the products and services themselves. It is enabling innovative processes and new ways of doing business, introducing new cross-sector value chains and infrastructures.

Open society must ensure that it capitalizes on the benefits that these developments promise for citizens in terms of tackling societal challenges and boosting business and industry. Drawing on these trends, and with the aim of helping build an open innovation ecosystem in open society, the open society's concept of open innovation is characterized by:

- combining the power of ideas and knowledge from different actors to co-create new products and find solutions to societal needs
- creating shared economic and social value, including a citizen and user-centric approach
- capitalizing on the implications of trends such as digitalization, mass participation, and collaboration

In order to encourage the transition from linear knowledge transfer towards more dynamic knowledge circulation, experts agree that it is essential to create and support an open innovation ecosystem that facilitates the translation of knowledge into socio-economic value. In addition to the formal supply-side elements such as research skills, excellent science, funding and intellectual property management, there is also a need to concentrate on the demand side aspects of knowledge circulation, making sure that scientific work corresponds to the needs of the users and that knowledge is findable, accessible, interpretable and reusable. Open access to research results aims to make science more reliable, efficient, and responsive and is the springboard for increased innovation opportunities, e.g. by enabling more science-based startups to emerge. Prioritizing open science does not, however, automatically ensure that research results and scientific knowledge are commercialized or transformed into socio-economic value. In order for this to happen, open innovation must help to connect and exploit the results of open science and facilitate the faster translation of discoveries into societal use and economic value.

Open Science

Collaborations with global partners represent important sources of knowledge circulation. The globalization of research and innovation is not a new phenomenon, but it has intensified in the last decade, particularly in terms of collaborative research, international technology production, and worldwide mobility of researchers and innovative entrepreneurs. Global collaboration plays a significant role both in improving the competitiveness of open innovation ecosystems and in fostering new knowledge production worldwide. It ensures access to a broader set of competencies, resources, and skills wherever they are located, and it yields positive impacts in terms of scientific quality and research results. Collaboration enables global standard-setting, allows global challenges to be tackled more effectively, and facilitates participation in global value chains and new and emerging markets.

To learn more, visit online at waset.org

Open Science

Scholarly Research Review

The scholarly research review is a multidimensional evaluation procedure in which standard peer review models can be adapted in line with the ethos of scientific research, including accessible identities between reviewer and author, publishing review reports and enabling greater participation in the peer review process. Scholarly research review methods are employed to maintain standards of quality, improve performance, provide credibility, and determine suitability for publication. *Responsible Peer Review Procedure:* Responsible peer review ensures that scholarly research meets accepted disciplinary standards and ensures the dissemination of only relevant findings, free from bias, unwarranted claims, and unacceptable interpretations. Principles of responsible peer review:

- Honesty in all aspects of research
- Accountability in the conduct of research
- Professional courtesy and fairness in working with others
- Good stewardship of research on behalf of others

The responsibilities of peer review apply to scholarly researchers at all stages of peer review: Fairness, Transparency, Independence, Appropriateness and Balance, Participation, Confidentiality, Impartiality, Timeliness, Quality and Excellence, Professionalism, and Duty to Report.

Scholarly Research Review Traits:

- Scholarly Research Review Identities: Authors and reviewers are aware of each other's identity
- Scholarly Research Review Reports: Review reports are published alongside the relevant article
- Scholarly Research Review Participation: The wider academic community is able to contribute to the review process
- Scholarly Research Review Interaction: Direct reciprocal discussion between author(s) and reviewers, and/or between reviewers, is allowed and encouraged
- Scholarly Research Pre-review Manuscripts: Manuscripts are made immediately available in advance of any formal peer review procedures
- Scholarly Research Review Final-version Reviewing: Editorial revision of the language and format is conducted on the final version of the manuscript for publication
- Scholarly Research Review Platforms: The scholarly research review process is independent of the final publication of the manuscript and it is facilitated by a different organizational entity than the venue of publication

All submitted manuscripts are subject to the scholarly research review process, in which there are three stages of evaluation for consideration: pre-review manuscripts, chair-review presentation, and final-review manuscripts. All submitted full text papers, that may still be withstand the editorial review process, are presented in the conference proceedings. Manuscripts are tracked and all actions are logged by internal and external reviewers according to publication policy. External reviewers' editorial analysis consists of the evaluation reports of the conference session chairs and participants in addition to online internal and external reviewers' reports. Based on completion of the scholarly research review process, those manuscripts meeting the publication standards are published 10 days after the event date.

To learn more, visit online at waset.org

TABLE OF CONTENTS

1	Strengthening Evaluation of Steel Girder Bridge under Load Rating Analysis: Case Study <i>Qudama Albu-Jasim, Majdi Kanaan</i>	1
2	Geometric Continuity in the Form of Iranian Domes, Study of Prominent Safavid and Sasanian Domes <i>Nima Valibeig, Haniyeh Mohammadi, Neda Sadat Abdelahi</i>	2
3	Upconversion Nanomaterials for Applications in Life Sciences and Medicine <i>Yong Zhang</i>	3
4	A Post-Occupancy Evaluation of LEED-Certified Residential Communities Using Structural Equation Modeling <i>Mohsen Goodarzi, George Berghorn</i>	4
5	A Virtual Reality Cybersecurity Training Knowledge-Based Ontology <i>Shaila Rana, Wasim Alhamdani</i>	5
6	Hybrid Multipath Congestion Control <i>Akshit Singhal, Xuan Wang, Zhijun Wang, Hao Che, Hong Jiang</i>	11
7	SPA-VNDN: Enhanced Smart Parking Application by Vehicular Named Data Networking <i>Bassma Aldahlan, Zongming Fei</i>	25
8	A Dynamic Model for Circularity Assessment of Nutrient Recovery from Domestic Sewage <i>Anurag Bhambhani, Jan Peter Van Der Hoek, Zoran Kapelan</i>	37
9	Floating Populations, Rooted Networks Tracing the Evolution of Russeifa City in Relation to Marka Refugee Camp <i>Dina Dahood Dabash</i>	38
10	To Improve or Not to Improve Reflections from Jerash Urban Improvement Project, Jordan <i>Dina Dahood Dabash</i>	39
11	Legal Issues Relating to Sustainable Buildings <i>Richard J. Sobelsohn</i>	40
12	Does Index-Based Livestock Insurance Change the Cash Saving Behaviour of Pastoralists in Southern Ethiopia? <i>Tnsue Gebrekidan Bezabh</i>	41
13	Analyze and Visualize Eye-Tracking Data <i>Aymen Sekhri, Emmanuel Kwabena Frimpong, Bolaji Mubarak Ayeyemi, Aleksii Hirvonen, Matias Hirvonen, Tedros Tesfay Andemichael</i>	42
14	Future Focused Strategy: How To Avoid Getting Fracked! <i>Dev Patnaik</i>	48
15	Multi-Robotic Partial Disassembly Line Balancing with Robotic Efficiency Difference via HNSGA-II <i>Tao Yin, Zeqiang Zhang, Wei Liang, Yanqing Zeng, Yu Zhang</i>	50
16	Resilient Manufacturing in Times of Mass Customisation: Using Augmented Reality to Improve Training and Operating Practices of EV's Battery Assembly <i>Lorena Caires Moreira, Marcos Kauffman</i>	56
17	Overview of Adaptive Spline interpolation <i>Rongli Gai, Zhiyuan Chang</i>	61
18	LACGC: Business Sustainability Research Model for Generations Consumption, Creation, and Implementation of Knowledge: Academic and Non-Academic <i>Satpreet Singh</i>	69

19	Diagnosis and Resolution of Intermittent High Vibration Spikes at Exhaust Bearing of Mitsubishi H-25 Gas Turbine using Shaft Vibration Analysis and Detailed Root Cause Analysis <i>Fahad Qureshi</i>	70
20	Transmission Characteristics and Anisotropy of Epsilon Near Zero Behaviour in Photonic Hypercrystal <i>Munazza Zulfiqar Ali</i>	79
21	Fibrations Property of Embedding Maps of Orbifold Charts <i>Hakeem A. Othman</i>	80
22	Generalized Bivariate Blending Type Bernstein Operators by a New Class of Coefficients <i>Seyed Yashar Zaheerani</i>	81
23	Contribution of the SidePlate Beam-Column Connections to the Seismic Responses of Special Moment Frames <i>Gökhan Yüksel, Serdar Akça, İlker Kalkan</i>	97
24	Verifying Environmental Performance through Inventory and Assessment: Case Study of the Los Alamos National Laboratory Waste Compliance and Tracking System <i>Oral S. Saulters, Shanon D. Goldberg, Wendy A. Staples, Ellena I. Martinez, Lorie M. Sanchez, Diego E. Archuleta, Deborah L. Williams, Scot D. Johnson</i>	105
25	Programmed Speech to Text Summarization Using Graph-Based Algorithm <i>Hamsini Pulugurtha, P. V. S. L. Jagadamba</i>	112
26	Review on PETG Material Parts Made Using Fused Deposition Modeling <i>Dhval Chauhan, Mahesh Chudasama</i>	115
27	Knowledge of Artificial Insemination and Agribusiness Management for Social Innovation in Rural Populations <i>Yasser Y. Lenis, Daniela Garcia Gonzalez, Cristian Solarte Bacca, Diego F. Carrillo González, Amy Jo Montgomery, Dursun Barrios</i>	116
28	Evaluation of Interspecific Pollination of <i>Elaeis guineensis</i> and <i>Elaeis oleifera</i> Carried Out in the Ucayali Region-Peru <i>Victor Sotero, Cindy Castro, Ena Velazco, Ursula Monteiro, Dora Garcia</i>	117
29	Identification of Gm15441, a Txnip Antisense lncRNA, as a Critical Regulator in Liver Metabolic Homeostasis <i>Mingyang Xin, Qian Guo, Qingchun Lu, Juan Lu, Po-Shun Wang, Yun Dong, Tao Li, Ye Chen, Glenn S. Gerhard, Xiao-Feng Yang, Michael Autieri, Ling Yang</i>	118
30	Structural Insights into the Bypass of the Major Deaminated Purines by Translesion Synthesis DNA Polymerase <i>Hunmin Jung, Michael Hawkins, Seongmin Lee</i>	134
31	Goal-Oriented Hope and Sustainable Technology Adoption by Smallholder Farmers: Evidence from Senegal <i>Shira Peles (Bukchin), Dorit Kerret</i>	154
32	Mistreatment of Women in Mental Health Systems <i>T. A. Rossman</i>	155
33	Formation of an Empire in the 21st Century: Theoretical Approach in International Relations and Worldview of the New World Order <i>Rami Georg Johann</i>	156
34	Trust in the Police: How Political Identity Shapes Violence Victims' Engagement with the State <i>Tasnia Symoom</i>	157
35	Memories of Lost Fathers: The Unfinished Transmission of Generational Values in Hungarian Cinema by Peter Falanga <i>Peter Falanga</i>	158
36	Representing the Un-representable in Zhang Yimou's <i>Coming Home</i> <i>Min YAng</i>	159

37	The Impact of Technology on Media Content Regulation <i>Eugene Mashapa</i>	160
38	Life Orientation Teachers' Perceptions of HIV Education for Diverse Sexual Identities in South African Schools <i>Jesse Ashley, Makaringe Gezani Given</i>	163
39	Buthainaal-Issa's Ques, Lilao Thaleb (Beauty and The Beast): A Feminist Cultural Study <i>Anwa Albader</i>	164
40	AI in Hungerstation <i>Latifa AlZamil, Fatima Mohammed Baba, Fatimah AlGhalib, Joanne Jifri</i>	165
41	SweatCoin <i>Talah Reda, Orjwan Jaad, Felwa Alotaibi, Wafa Alsaqqaf</i>	170
42	Heritage Spanish Speaker's Bilingual Practices and Linguistic Varieties: Challenges and Opportunities <i>Ana C. Sanchez</i>	174
43	Social Internet of Things: Future Security Development <i>Muhammad Ali Fahroni, Tutun Juhana S.T. M. T.</i>	178
44	TESTING REPLIKA (CHATBOT): AN AI APP <i>SALWA QURESHI, HYA ALLUGMANI, ASSIL BAJNAID, TIYANA KHALDIEH</i>	184
45	Learning Communities and Collaborative Reflection for Teaching Improvement <i>Mariana Paz Sajon, Paula Cecilia Primogerio, Mariana Albarracin</i>	189
46	Effect of Training on Digitalization in Manufacturing Companies - A Quantitative Study <i>Maria Hulla, Patrick Herstätter, Daniela Moser, Johanna Pirker, Sandra Brettschuh, Michael Holly, Harald Burgsteiner, Christian Ramsauer</i>	190
47	A Note on the Effect of Covid-19 and Vaccine Rollout on School Enrollment in the United States <i>Kolawole Ogundari</i>	191
48	Ethical and Personality Factors and Accounting Professional Judgement <i>Shannon Hashemi, Alireza Daneshfar</i>	203
49	Good Communication Methods and Stakeholder Engagement Resulting in Effective Project Management <i>Thakaramahlaha Lehohla</i>	220
50	Developing a Document of Good Practice to Support SEND Professionals in Their Role of Eliciting Pupil Voice and Achieving Meaningful Participation <i>Pooja Sharma</i>	229
51	Coping for Academic Women Departmental Heads during COVID-19: A Capabilities Approach Perspective <i>Juliet Ramohai</i>	230
52	The Development of Educational Video Games Aimed at Enhancing Academic Motivation and Learning Among African American Males <i>Kenneth Philip Jones</i>	231
53	Identifying Family Needs, Support, and Barriers for More Effective Involvement in Early Intervention Services <i>Sadeem A. Alolayan</i>	241
54	Principal Creative Leadership for Teacher Learning and School Culture <i>Yashi Ye</i>	260
55	Using Game-Based Learning Platforms in the Student Experience: Inclusion for All <i>Ellen Spender</i>	266
56	Cultural and Social Conflicts in Former British Settler Colonies (the US, Canada, Australia and New Zealand): The Fate of Indigenous Communities and Rights <i>Sophie Croisy</i>	267
57	Sheriffs and Localized Racial Injustice in the American Jim Crow South, 1880-1965 <i>Megan Qingxi Li</i>	268

58	Diversity and Intersectionality Through a Disability Lens <i>Leigha Shoup</i>	269
59	Arousal, Encoding, And Intrusive Memories <i>Hannah Gutmann, Rick Richardson, Richard Bryant</i>	270
60	Indigenous Healers and Indigenous Trauma: Healing at the Intersections of Colonial, Intergenerational, and Individual Trauma for Indigenous Peoples in Canada <i>Suzanne L. Stewart, Mikaela D. Gabriel</i>	271
61	Preventative Programs for At-Risk Families of Child Maltreatment: Using Home Visiting and Intergenerational Relationships <i>Kristina Gordon</i>	272
62	Sceletium Tortuosum: A review on its Phytochemistry, Pharmacokinetics, Biological and Clinical Activities <i>Tomi Lois Olatunji, Frances Siebert, Ademola Emmanuel Adetunji, Brian Harvey, Johane Gericke, Josias Hamman, Frank Van Der Kooy</i>	273
63	Micromechanical Compatibility Between Cells and Scaffold Mediates the Efficacy of Regenerative Medicine <i>Li Yang, Yang Song, Martin Y. M. Chiang</i>	275
64	Design of Experiment for Optimizing Immunoassay Microarray Printing <i>Alex J. Summers, Jasmine P. Devadhasan, Douglas Montgomery, Brittany Fischer, Jian Gu, Frederic Zenhausem</i>	277
65	Application of Nanoparticles in Biomedical and MRI <i>Raziyeh Mohammadi</i>	278
66	The Effect of Pregabalin on the Prevention of Succinylcholine-Induced Fasciculation and Myalgia <i>Shahryar Sane, Hedyeh Rezaei, Shima Khanahmadi</i>	279
67	Evaluating the Effectiveness of Combined Psychiatric and Psychotherapeutic Care versus Psychotherapy Alone in the Treatment of Depression and Anxiety in Cancer Patients <i>Nathen A. Spitz, Dennis Martin Kivlighan III, Arwa Aburizik</i>	280
68	Bias Prevention in Automated Diagnosis of Melanoma: Augmentation of a Convolutional Neural Network Classifier <i>Kemka Ithemelandu, Chukwuemeka Ithemelandu</i>	281
69	Childhood Epilepsy: Solutions and Study Methods That Have Been Attracting Researchers <i>Jhonas Geraldo Peixoto Flauzino, Pedro Pompeo Boechat Araujo, Alexia Allis Rocha Lima, Giovanna Biângulo Lacerda Chaves, Victor Ryan Ferrão Chaves</i>	282
70	Expert Opinions about Barriers to Physical Activity among Ghanaian Adults with Type 2 Diabetes Mellitus: A Qualitative Descriptive Study <i>Mohammed Amin, Debra Kerr, Yacoba Atiase, Andrea Driscoll</i>	283
71	A Comparison of qCON/qNOX to the Bispectral Index as Indices of Antinociception in Surgical Patients Undergoing General Anesthesia with Laryngeal Mask Airway <i>Roya Yumul, Ofelia Loani Elvir-Lazo, Sevan Komshian, Ruby Wang, Jun Tang</i>	284
72	The Impact of COVID-19 on Childhood Academic Functioning and Anxiety: A Literature Review <i>Lindsey Giunta</i>	285
73	Preoperative Anxiety Evaluation: Comparing the Visual Facial Anxiety Scale/Yumul Faces Anxiety Scale, Numerical Verbal Rating Scale, Categorization Scale, and the State-Trait Anxiety Inventory <i>Roya Yumul, Chse, Ofelia Loani Elvir Lazo, David Chernobylsky, Omar Durra</i>	286
74	Efficacy of Gamma Knife Radiosurgery in Refractory Obsessive-Compulsive Disorder: An Indian Experience <i>Milind Sankhe, Sanjeev Pattankar</i>	287
75	Molecular Profiles of Microbial Etiologic Agents Forming Biofilm in Urinary Tract Infections of Pregnant Women by RTPCR Assay <i>B. Nageshwar Rao</i>	288

76	Data Augmentation for Early-Stage Lung Nodules Using Deep Image Prior and Pix2pix <i>Qasim Munye, Juned Islam, Haseeb Qureshi, Syed Jung</i>	289
77	Ultrasound-Guided Intraarticular Electrical Stimulation, a New Intervention for Improving Physical Activity and Quality of Life in Patients with Severe Knee Osteoarthritis: Cases Report <i>Azar Behnamfar</i>	290
78	Implementing a Multimodality, Multidisciplinary Approach to Novice Nurse Success in the Emergency Department <i>Rocio Garcia, Guillermo Herrera, Rhiannon Rasmussen, Nicole Fraire</i>	296
79	In Vitro Dissolution Profile of Different Brands of Low Solubility Drugs Available in the Nigerian and Brazilian Pharmaceutical Market <i>Ezeagu Henry Ifeanyi, Issa Michele Georges, Ofoefule Sabinus Ifeanyi, Ferraz Humberto Gomes</i>	297
80	Computer-Aided Drug Repurposing for Mycobacterium Tuberculosis by Targeting Tryptophanyl-tRNA Synthetase <i>Neslihan Demirci, Serdar Durdađı</i>	305
81	Bluetooth Proximity Estimation with Machine Learning Algorithms for Private Automated Contact Tracing to limit the spread of COVID-19 <i>Collin Li</i>	306
82	Individualized Emotion Recognition Through Dual-Representations and Ground-Established Ground Truth <i>Valentina Zhang</i>	325
83	Utilizing Artificial Intelligence to Predict Post Operative Atrial Fibrillation in Non-Cardiac Transplant <i>Alexander Heckman, Rohan Goswami, Zachi Attia, Paul Friedman, Peter Noseworthy, Demilade Adedinsewo, Pablo Moreno-Franco, Rickey Carter, Tathagat Narula</i>	336
84	Preparation of Papers - Developing a Leukemia Diagnostic System Based on Hybrid Deep Learning Architectures in Actual Clinical Environments <i>Skyler Kim</i>	337
85	Integrated Machine Learning Framework for At-Home Patients Personalized Risk Prediction Using Activities, Biometric, and Demographic Features <i>Claire Xu, Welton Wang, Manasvi Pinnaka, Anqi Pan, Michael Han</i>	338
86	Establishing a Data Processing Pipeline Using Artificial Intelligence to Capture Human Kinematic Data: DeepLabCut as a Markerless Motion Capture System <i>Fatemeh Khounsarian, Sowmya Gopalakrishnan, Devyani McLaren, Sara Kraeutner, Lara Boyd</i>	339
87	Introduction of Integrated Image Deep Learning Solution and How It Brought Laboratorial Level Heart Rate and Blood Oxygen Results to Everyone <i>Zhuang Hou, Xiaolei Cao</i>	342
88	Protocol Selection of Advanced Imaging Exams Using Multi-Steps Deep Learning Models <i>C. Cronister, T. B. Issa, D. Uminsky, R. W. Filice</i>	346
89	The Importance of Fruit Trees for Prescribed Burning in a South American Savanna <i>Rodrigo M. Falleiro, Joaquim P. L. Parime, Luciano C. Santos, Rodrigo D. Silva</i>	354
90	Agroforestry Systems and Practices and Its Adoption in Kilombero Cluster of Sagcot, Tanzania <i>Lazaro E. Nnko, Japhet J. Kashaigili, Gerald C. Monela, Pantaleo K. T. Munishi</i>	355

Strengthening Evaluation of Steel Girder Bridge under Load Rating Analysis: Case Study

Qudama Albu-Jasim, Majdi Kanaan

Abstract— A case study about the load rating and strengthening evaluation of the six-span of steel girders bridge in Colton city of State of California is investigated. To simulate the load rating strengthening assessment for the Colton Overhead bridge, a three-dimensional finite element model built in the CSiBridge program is simulated. Three-dimensional finite-element models of the bridge are established considering the nonlinear behavior of critical bridge components to determine the feasibility and strengthening capacity under load rating analysis. The bridge was evaluated according to Caltrans Bridge Load Rating Manual 1st edition for rating the superstructure using the Load and Resistance Factor Rating (LRFR) method. The analysis for the bridge was based on load rating to determine the largest loads that can be safely placed on existing I-girder steel members and permitted to pass over the bridge. Through extensive numerical simulations, the bridge is identified to be deficient in flexural and shear capacities, and therefore strengthening for reducing the risk is needed. An in-depth parametric study is considered to evaluate the sensitivity of the bridge's load rating response to variations in its structural parameters. The parametric analysis has exhibited that uncertainties associated with the steel's yield strength, the superstructure's weight, and the diaphragm configurations should be considered during the fragility analysis of the bridge system.

Keywords— load rating, CSiBridge, strengthening, uncertainties, case study.

Geometric Continuity in the Form of Iranian Domes, Study of Prominent Safavid and Sasanian Domes

Nima Valibeig, Haniyeh Mohammadi, Neda Sadat Abdelahi

Nima Valibeig is with the Art University of Isfahan, Italy (e-mail: n.valibeig@au.ac.ir).

Abstract— Persian domes follow different forms depending on the materials used to construct and other factors. One of the factors that shape the form of a dome is the geometric proportion used in the drawing and construction of the dome. Some commonly used proportions are revealed by analysing the shapes and geometric ratio of the monuments' domes. The proportions are achieved by the proficiency of the skilled architects of the buildings. These proportions can be used to reconstruct damaged parts of the historical monuments. Most of the research on domes is about the historical or stability features of domes, and less attention is made to the geometric system in domes. Therefore, in this study, we study the explicit and implicit geometric proportions in Iranian dome structures for the first time. The study is done based on literature review and field survey. This research reveals that the permanent geometric rules are perfectly used in the design and construction of the prominent domes.

Keywords— Geometry in architecture, architectural proportions, prominent domes, Iranian Golden Ratio, geometric proportion

Objectives:

The goal of this paper is to reveal the magnificent geometric proportions in the design and construction of Persian domes. These proportions are discovered by traditional architects and guarantee the beauty and stability of domes at the same time.

Methodology:

The study is based on literature review, descriptive and visual document analysis, and field surveys.

Contribution:

The study reveals the geometric proportion that guarantees the beauty and stability of domes. The revealed proportion can be used for the conservation and restructuring of old buildings.

Upconversion Nanomaterials for Applications in Life Sciences and Medicine

Yong Zhang

Abstract— Light has proven to be useful in a wide range of biomedical applications such as fluorescence imaging, photoacoustic imaging, optogenetics, photodynamic therapy, photothermal therapy, and light controlled drug/gene delivery. Taking photodynamic therapy (PDT) as an example, PDT has been proven clinically effective in early lung cancer, bladder cancer, head, and neck cancer and is the primary treatment for skin cancer as well. However, clinical use of PDT is severely constrained by the low penetration depth of visible light through thick tissue, limiting its use to target regions only a few millimeters deep. One way to enhance the range is to use invisible near-infrared (NIR) light within the optical window (700–1100nm) for biological tissues, extending the depth up to 1cm with no observable damage to the intervening tissue. We have demonstrated use of NIR-to-visible upconversion fluorescent nanoparticles (UCNPs), emitting visible fluorescence when excited by a NIR light at 980nm, as a nanotransducer for PDT to convert deep tissue-penetrating NIR light to visible light suitable for activating photosensitizers. The unique optical properties of UCNPs enable the upconversion wavelength to be tuned and matched to the activation absorption wavelength of the photosensitizer. At depths beyond 1cm, however, tissue remains inaccessible to light even within the NIR window, and this critical depth limitation renders existing phototherapy ineffective against most deep-seated cancers. We have demonstrated some new treatment modalities for deep-seated cancers based on UCNP hydrogel implants and miniaturized, wirelessly powered optoelectronic devices for light delivery to deep tissues.

Keywords— upconversion, fluorescent, nanoparticle, bioimaging, photodynamic therapy.

A Post-Occupancy Evaluation of LEED-Certified Residential Communities Using Structural Equation Modeling

Mohsen Goodarzi, George Berghorn

Abstract— Despite the rapid growth in the number of green building and community development projects, the long-term performance of these projects has not yet been sufficiently evaluated from the users' points of view. This is partially due to the lack of post-occupancy evaluation tools available for this type of project. In this study, a post-construction evaluation model is developed to evaluate the relationship between the perceived performance and satisfaction of residents in LEED-certified residential buildings and communities. To develop this evaluation model, a primary five-factor model was developed based on the existing models and residential satisfaction theories. Each factor of the model included several measures that were adopted from LEED certification systems such as LEED-BD+C New Construction, LEED-BD+C Multifamily Midrise, LEED-ND, as well as the UC Berkeley's Center for the Built Environment survey tool. The model included four predictor variables (factors), including perceived building performance (8 measures), perceived infrastructure performance (9 measures), perceived neighborhood design (6 measures), and perceived economic performance (4 measures), and one dependent variable (factor), which was residential satisfaction (6 measures). An online survey was then conducted to collect the data from the residents of LEED-certified residential communities (n=192) and the validity of the model was tested through Confirmatory Factor Analysis (CFA). After modifying the CFA model, 26 measures, out of the initial 33 measures, were retained to enter into a Structural Equation Model (SEM) and to find the relationships between the perceived buildings performance, infrastructure performance, neighborhood design, economic performance and residential Satisfaction. The results of the SEM showed that the perceived building performance was the most influential factor in determining residential satisfaction in LEED-certified communities, followed by the perceived neighborhood design. On the other hand, perceived infrastructure performance and perceived economic performance did not show any significant relationship with residential satisfaction in these communities. This study can benefit green building researchers by providing a model for the evaluation of the long-term performance of these projects. It can also provide opportunities for green building practitioners to determine priorities for future residential development projects.

Keywords— green building, residential satisfaction, perceived performance, confirmatory factor analysis, structural equation modeling.

A VR Cybersecurity Training Knowledge-Based Ontology

Shaila Rana, Wasim Alhamdani

Abstract—Effective cybersecurity learning relies on an engaging, interactive, and entertaining activity that fosters positive learning outcomes. VR cybersecurity training may promote these aforementioned variables. However, a methodological approach and framework have not yet been created to allow trainers and educators to employ VR cybersecurity training methods to promote positive learning outcomes, to the author’s best knowledge. Thus, this paper aims to create an approach that cybersecurity trainers can follow to create a VR cybersecurity training module. This methodology utilizes concepts from other cybersecurity training frameworks, such as NICE and CyTrONE. Other cybersecurity training frameworks do not incorporate the use of VR. VR training proposes unique challenges that cannot be addressed in current cybersecurity training frameworks. Subsequently, this ontology utilizes concepts unique to developing VR training to create a relevant methodology for creating VR cybersecurity training modules. The outcome of this research is to create a methodology that is relevant and useful for designing VR cybersecurity training modules.

Keywords—Virtual reality cybersecurity training, VR cybersecurity training, traditional cybersecurity training, ontology.

I. INTRODUCTION

EFFECTIVELY designed cybersecurity training programs are of the utmost importance. Poorly designed cybersecurity training programs result in insignificant improvements within a population [1]. Thus, specific and targeted training programs and frameworks need to be developed for an appropriately designed training module. Cybersecurity training frameworks, such as CyTrONE and NICE, exist; however, nuances within VR cybersecurity training must be addressed. Consequently, this study proposes to fill the gap in the cybersecurity training field by creating a VR cybersecurity training framework that can assist organizations and individuals alike. VR cybersecurity training may be a mitigating factor to the problems found in traditional training methods. Furthermore, VR cybersecurity can extend game-based training to account for an immersive and highly interactive experience to support positive learning outcomes. VR cybersecurity training is an immature field with few studies focusing on its potential significance. Researchers studied the use of virtual reality to enforce cybersecurity principles, and it was found that VR and augmented reality tools can teach cybersecurity fundamentals effectively and support active learning [2]. Researchers have demonstrated traditional cybersecurity training methods as ineffective in changing user behavior and defending against cyber threats

[3]. Furthermore, researchers have found that traditional training methods have been cited as “boring and tedious” and lack success in programs [4]. VR cybersecurity training was demonstrated to be a more engaging learning platform for cybersecurity education than traditional training methods [5]. VR systems allow students to learn cybersecurity principles in an interactive way [2]. VR training is utilized in other industries, such as in the healthcare industry, for medical training and is demonstrated as applicable [6]. Thus, the need for exploring VR cybersecurity training modules needs to be addressed. Consequently, a proposed ontology for the development of VR cybersecurity training may encourage additional research and usage of VR training simulations and games. Moreover, creating ontologies is significant for both the cybersecurity field and cybersecurity training field as it can contribute to assist decision-makers in need of effective cybersecurity training modules [7].

This paper aims to fill the gap in current cybersecurity training literature by providing a methodology for creating VR cybersecurity training. VR scenarios require additional planning and designing; thus, this methodology aims to incorporate the unique challenges of developing VR simulations. Some of these challenges include the stylistic components that are unique to VR simulations. The creation of VR simulations is a unique undertaking compared to other cybersecurity training platforms. Consequently, this ontology addresses the challenges mentioned above in a seven-step process.

A. Research Contribution

The contribution of this study is to propose a framework to design a VR cybersecurity training program that is customized to the user. This framework proposes to create an engaging, interactive, and entertaining platform to encourage positive learning outcomes in a user. Furthermore, this study aims to contribute to the cybersecurity training discipline by creating a framework that considers the user’s needs, level of technical prowess, and the industry in which he or she works. This study aims to create an ontological model to create an effective cybersecurity training platform that equips users to work against cyberattacks’ growing number and complexity. This paper has not yet been implemented and is currently a work in progress. The proposed methodology includes the idiosyncrasies involved in the production and creation of VR simulations and games. The methodology addressed in this paper is deduced from both the NICE and CyTrONE training frameworks.

Shaila Rana is with University of the Cumberland, United States (e-mail: shailashifarana@gmail.com).

II. OTHER FRAMEWORKS COMPARED TO THE PROPOSED ONTOLOGY

A. CyTrONE Cybersecurity Training Framework Compared to the Proposed Ontology

Cybersecurity training and education are required to prepare security professionals for the ubiquity and complexity of cybersecurity threats [8]. Thus, the potential workforce needs to be adequately trained to tackle these inevitable scenarios. Researchers note that a methodology is necessary to disseminate the required knowledge in cybersecurity education programs [8]. Furthermore, a methodology can assist educational and training institutions in providing an appropriate cybersecurity training program to equip professionals adequately [8]. Frameworks can allow for a customized approach that incorporates both the students' knowledge and the cybersecurity industry's needs [8]. This framework differs in that it is specifically aimed at developing VR simulations. VR simulations require a different development level and can be either just a simulation or game-based training. Thus, some nuances create another layer beyond only cyber ranges. Furthermore, this framework includes a more rigorous testing process due to potential adverse reactions in users. Simulation sickness occurs if adequate development and testing have not occurred, creating an unpleasant experience for users. Spatial awareness and inclusion of head, eye, and hand gestures also need to be included in developing a VR simulation or game. Consequently, the proposed framework in this study accounts for the differences mentioned above.

CyTrONE is a framework that includes a training database to create training content and the necessary environment within a cyber range for training [9]. These cyber ranges are administered on virtual machines for users. VR simulations differ from virtual machines in that users wear a VR headset and are immersed within the scene or game. On the other hand, virtual machines are administered on a computer or device. This study proposes a framework that is specifically designed for VR simulations and games. VR simulations include another layer of human interactivity and human input, including head and eye movements. For instance, the proposed ontology includes an additional testing phase before the VR simulation or game is administered. Thus, the proposed framework for CyTrONE, while highly flexible and configurable, is not complete for VR cybersecurity training modules. Instead, additional steps need to be included in the framework to account for human input and gestures, the subject in question, and extensive testing.

B. NICE Framework Compared to the Proposed Ontology

This ontology differs from the NICE framework in that this framework is not only focused on cybersecurity job roles and the development of cybersecurity professionals. Work roles are an inherent part of the NICE framework. On the other hand, VR cybersecurity training can apply to many audiences, including non-cyber security users. Thus, this study's proposed ontology aims to be a more inclusive framework that targets a broad audience and can be used outside government, educational institutions, and private organizations. Furthermore, this ontology focuses on VR development,

which includes an immersive training platform requiring more testing and development than traditional training methods. Contrastingly, the NICE framework is primarily focused on traditional cybersecurity training methods. While this ontology builds upon the components proposed in the NICE framework, it extends it by including specificities found in VR simulations and games.

III. PROPOSED ONTOLOGY

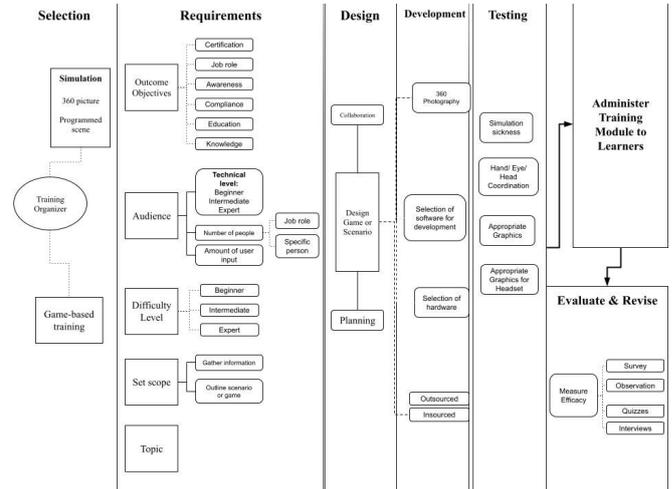


Fig 1: Proposed Ontology with 7 Distinct Phases

A. Stage I: Selection

The first stage of the ontology is the selection done by the training organizer. The training organizer can be one entity or a committee that can communicate, administer, and support the cybersecurity training program. The training organizer is the point of contact for setting the requirements for the VR cybersecurity training program. In the selection stage, the training organizer must decide the direction of the VR training content. VR training content can be created as either simulation or game-based training. Simulation training will allow a user to be immersed in a scenario or scene, walk around, explore, and interact with the scene. On the other hand, game-based training requires more user input. Depending upon user actions and selections, the user will experience different scenarios and challenges.

Scenario-based VR training includes immersing a user within a scene or scenario. Users have the option to interact with the environment or merely observe. Scenario-based VR training can be developed through software, or it can be done through 360-degree photographs. The 360-degree photographs allow a user to wear a headset and be immersed within a scene; however, the user cannot interact with the surroundings (i.e., pick up objects, move objects around). Instead, the user is in a photograph and can observe. This is a less expensive way to create VR-based training while allowing the user to be immersed within a scene. The 360-degree photographs may be appropriate for allowing users to spot security infractions or relying on observation-based training. This is a less interactive form of training and may be appropriate for a wider audience, depending on technical skill level and the amount of interaction and time required for training. Alternatively, a scenario-based simulation may allow the user to interact with

the surroundings, like holding objects and moving objects around. This will be most costly than 360-degree photographs but less expensive than developing game-based training.

Game-based training can have different game types such as role-play, adventure, action, strategy, puzzles, board games, and network simulations [10]. Consequently, the VR simulation game-based training can include the aforementioned or a combination of these game types. It is important to note that most game-based target audiences include students ranging from children to teenagers [10]. Essentially, game-based training design should heavily depend on the target demographic of cybersecurity learners. Demographics include but are not limited to the skill level, age, and gender of users. Game-based VR training may not be appropriate for a wide range of audiences; however, the nature of high user engagement and interactivity can produce more positive learning outcomes. Furthermore, game-based training allows users to view challenges in different ways and scenarios [11]. In general, game-based training is an "excellent platform" to train users [12]. Developing and implementing VR game-based training will be more costly in terms of time, resources, and finances. Furthermore, there is a higher likelihood that game-based VR training development will be outsourced to third-party vendors. Thus, the training organizer must decide whether the VR development will be simulation or scenario-based training or game-based training.

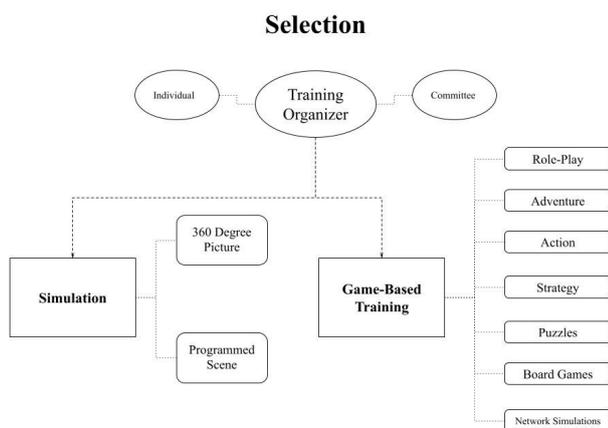


Fig 2. Selection Stage

B. Stage II: Requirements

The second stage consists of building and setting requirements for the VR simulation or game-based training. This stage is relatively laborious compared to the other six stages; however, several factors need to be considered when designing and creating VR training. First, the training organizer(s) must decide the topic of cybersecurity training to be conducted. If the training is meant to be administered to office members, physical cybersecurity policies may be the topic in focus. If the training is meant to be delivered to university students exploring network security, the topic will differ. Some examples of regularly discussed cybersecurity training topics include physical cybersecurity concepts, incident response,

social engineering, malware, patch management, password management, compliance, critical infrastructure, and many more. Selecting a topic will depend upon the audience, outcome objective, difficulty level, and overall scope. Additionally, there may be multiple topics covered in one VR simulation or game. Consequently, the training organizer(s) must select this information to determine what information will be developed, depending upon the outcome objective, audience, scope, and difficulty level. Essentially, this will vary significantly between training organizers and organizations.

The outcome objective needs to be defined by the training organizer(s) to select the topic. There are different outcome objectives for cybersecurity training. Outcome objectives can range from training for a learner to certification, a job or work role, cybersecurity awareness, compliance, education, or merely obtaining knowledge. The goal of VR training will depend upon what the desired outcome objectives are. Additionally, there may be a combination of outcome objectives; however, it will depend upon the training organizer(s) goals. Depending on those objectives, VR cybersecurity training requirements will be drafted to develop, develop, test, and administer.

The audience is another consideration to take into account when outlining the requirements for VR cybersecurity training. The target audience's technical prowess level should be defined to avoid administering too complex training for their knowledge or too simple to extract useful learning from the training module. Thus, the training organizer needs to define what the technical capacity level is for the target audience. Furthermore, the number of people who will receive this training needs to be considered and whether the training is meant for a specific person and job role. These factors will contribute to outlining applicable and appropriate requirements for the VR cybersecurity training module. As aforementioned, VR cybersecurity training creates an immersive experience for a learner, given that a headset will be placed on them only to view the scene created by the training organizer. However, the amount of user input can vary greatly. As mentioned in stage 1, the training organizer must select whether the VR training module will be a simulation or game-based training. Subsequently, there will be different levels of human input and interaction. The training organizer(s) must also consider how much user input is required in the requirements stage. If the training module is simply a simulation or scenario, it should be defined as picking up objects and interact with the scenario. On the other hand, if the training module is defined as game-based training, the organizer(s) must determine how much user input and interaction will be required to progress throughout the game. It is essential to define the amount of user input before designing and developing the game.

The scope of the VR simulation or game should also be considered in the requirements stage. Depending upon the topic(s) selected by the training organizer, the scope will differ from training module to training module. Additionally, the overall objectives of the training module will influence the scope of the VR training module. To set the scope, the training organizer must gather information regarding the overall outcome objectives. If the outcome objective is to spread awareness simply, the scope may be set within a scene in an

office focusing on commonplace office policies. On the other hand, if the scope is to create incident response training, the scope will encumber more scenes, principles to be taught, and human interaction. In general, the scope will vary depending on the training organizer, the outcome objectives, audience, topic, and difficulty level. Thus, given the nature of the high variability of VR training modules, this ontology determines setting the scope should be noted and done within the requirements stage.

It is suggested to create a rough outline of the scenario or game to determine the VR training module’s scope. Finally, another consideration is to determine the difficulty level of the game or simulation. The audience’s technical ability will heavily influence the difficulty level set in the requirements stage. However, it is crucial to differentiate the training module’s overall difficulty level from the audience’s technical capability level. This is because the VR simulation or game may or may not be technical by nature. Instead, games and simulations may be categorized as intermediate to progress throughout the training, while the audience’s technical capability may be a beginner. For example, university students may be categorized as a beginner technical level but find a challenging game for learning cybersecurity fundamentals. Consequently, the difficulty level and audience technical capability level are differentiated in this stage. The next stage will be accumulating the requirements and designing a VR simulation or game.

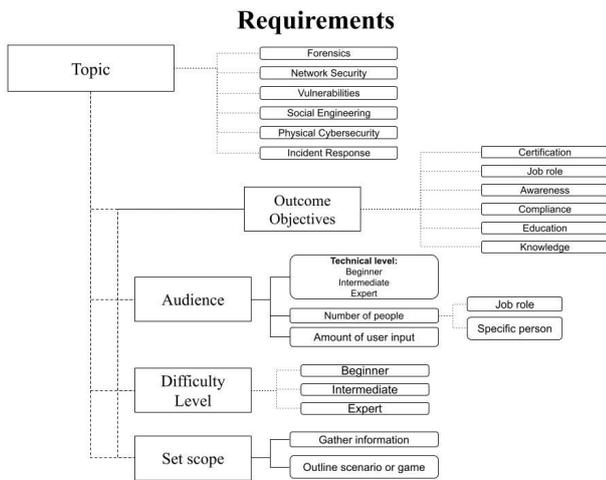


Fig 3: Requirements stage

C. Stage III: Design

Designing the VR training module will take stages one and two and translate it into an overall simulation design. The outputs of stage one include whether the VR training module will be a simulation or game. The outputs of stage two are the overall requirements for the training module. These outputs include the outcome objectives, target audience, scope, difficulty level, and cybersecurity topic. The design stage is a natural next step from the requirements stage. Designing a VR training module will require planning and possibly collaboration. For gamified VR cybersecurity training, designers with past game experience may be necessary.

However, if the VR game is for a beginner audience and beginner’s difficulty level, designers may not require a strong game background. The training organizer(s) must evaluate whether or not a dedicated designer is required. Collaboration can also include gathering input from other stakeholders. Stakeholders can range from the learner to members within an organization or even possibly third parties. This will heavily depend on the training organizer’s goals for the cybersecurity training module, scope, target audience, outcome objectives, and difficulty level.

In the design stage, the elements of the game will be outlined. Elements may include the scene in which the game or simulation will occur, how the user will interact with the VR simulation, and how topic information will be disseminated to the user. If text is utilized for training the user, the way that text will be displayed must also be addressed in this stage. On the other hand, if sound, music, or voiceovers are utilized in the VR game or simulation, that must be appropriately addressed in this stage. The design stage will outline the key elements that will be passed on to developers. Design and development may be outsourced to third parties to ensure the VR training module can support the requirements and selections made in the previous two stages. Depending upon the outputs of stages one and two, outsourcing to third-party vendors may be deemed appropriate. Ultimately, this will depend upon the training organizer whether or not a collaboration with external vendors is required. Overall, focused planning should occur in this stage to create a blueprint for developing a VR simulation or game. The design phase’s output will allow developers to create a simulation or game that supports the requirements delineated in stage two.

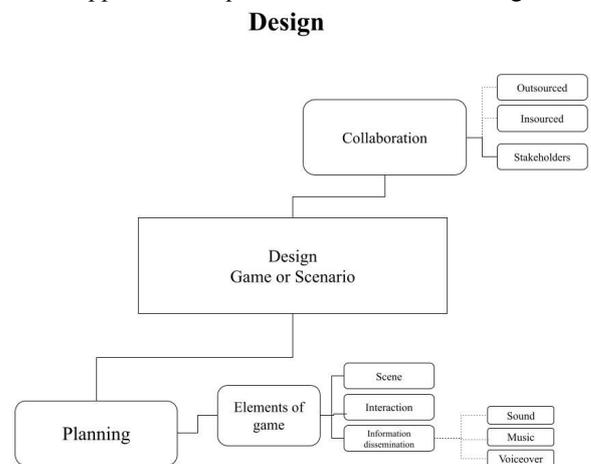


Fig 4: Design stage

D. Stage IV: Development

Stage four focuses on the development of the training module. This includes selecting the software for software development and developing the game or scenario. If 360-degree photographs are decided upon by the training organizer, this is the stage in which the photography would occur. The selection of hardware must happen in this stage or before developing an appropriate training module.

Development of the training module that embodies the requirements and design by the training organizer(s) may be insourced or outsourced depending on the resources, budget, and preferences of the training organizer(s). How development is done is dependent upon the earlier stages and overall design of the game.



Fig 5: VR simulation example of a common office environment

E. Stage V: Testing

Testing the VR training module is essential and sets this ontology apart from other cybersecurity training frameworks. VR training modules can have adverse effects if testing is not done adequately. Simulation sickness is a consequence of development errors and insufficient testing. Furthermore, if the VR module is required to be more interactive, hand, eye, and head coordination must be tested beforehand. Additionally, depending upon the VR simulation or game's hardware, graphics needed to be tested. Graphics should be appropriate not only for the equipment hosting the VR training but also for the learner. Consequently, the testing stage needs to be thorough and rigorous compared to other cybersecurity training development platforms. Poor graphics and simulation sickness will cause a negative experience for the learner and may not yield effective learner outcomes. Thus, the testing stage is vital and should be a specific step or stage in creating VR training modules.

F. Stage VI: Administer Testing

The sixth stage is to administer the training module to learners. After sufficient testing, this is conducted to ensure the VR training module fulfills the requirements and design of the training organizer(s) and avoids adverse outcomes (i.e., simulation sickness).

G. Stage VII: Evaluate and Revise

Evaluations and revisions are encouraged to ensure the VR training platform fulfills the training organizer's requirements and design components. Surveys, observations, quizzes, and interviews may be conducted to measure the VR training module's efficacy. Learner feedback will help revise the training module that supports cybersecurity training requirements and encourages the outcome objectives set by the training organizer(s). On the sample of feedback received from learners, the sample can be divided into groups to measure the effectiveness between learner groups. Thus, this can give a unique perspective on the efficacy of the VR training module.

IV. CONCLUSION

The proposed ontology aims to fill a gap in existing literature and cybersecurity training frameworks by accounting for VR

simulations and games' unique implementation and idiosyncrasies. The ontology is divided into seven distinct phases, building on the four phases included in the CyTrONE framework. The way in which this ontology is distinct from other cybersecurity training frameworks, such as NICE and CyTrONE, is due to the highly interactive and immersive nature of VR simulations and games. Additionally, the way this ontology differs from other training frameworks is discussed. This ontology aims to answer the question of how this methodology differs from existing methodologies. Furthermore, this paper aims to fill a gap in the literature on creating and implementing VR cybersecurity training programs. In general, VR simulations have been demonstrated to provide an interactive and engaging platform that supports active learning [2]. Compared to other training platforms, VR simulations have been demonstrated to be more effective [14]. Therefore, this ontology aims to support the creation of VR cybersecurity training programs by addressing the unique nature of VR simulations. The proposed methodology builds upon the author's previous research determining the need to study the efficacy of VR cybersecurity training compared to traditional methods [13]. This ontology is a work in progress and has not yet been quantitatively studied. Future work aims to address how this framework quantitatively differs from other cybersecurity training frameworks.

V. REFERENCES

- [1] Beuran, R., Inoue, T., Tan, Y., & Shinoda, Y. (2019, June). Realistic cybersecurity training via scenario progression management. In *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)* (pp. 67-76). IEEE.
- [2] Seo, J. H., Bruner, M., Payne, A., Gober, N., & McMullen, D. (2019). Using virtual reality to enforce principles of cybersecurity. *The Journal of Computational Science Education*, 10(1).
- [3] Sabillon, R., Serra-Ruiz, J., & Cavaller, V. (2019). An effective cybersecurity training model to support an organizational awareness program: The Cybersecurity Awareness TRaining Model (CATRAM). A Case Study in Canada. *Journal of Cases on Information Technology (JCIT)*, 21(3), 26-39.
- [4] Aldawood, H., & Skinner, G. (2019). Reviewing cyber security social engineering training and awareness programs—Pitfalls and ongoing issues. *Future Internet*, 11(3), 73.
- [5] Veneruso, S. V., Ferro, L. S., Marrella, A., Mecella, M., & Catarci, T. (2020, September). CyberVR: An Interactive Learning Experience in Virtual Reality for Cybersecurity Related Issues. In *Proceedings of the International Conference on Advanced Visual Interfaces* (pp. 1-8).
- [6] Kasurinen, J. (2017). Usability issues of virtual reality learning simulator in healthcare and cybersecurity. *Procedia computer science*, 119, 341-349.
- [7] Oltramari, A., Henshel, D. S., Cains, M., & Hoffman, B. (2015). Towards a Human Factors Ontology for Cyber Security. *Stids, 2015*, 26-33.
- [8] Kim, E., & Beuran, R. (2018). On designing a cybersecurity educational program for higher education. *Proceedings of the 10th International Conference on Education Technology and Computers*, 195-200. <https://doi.org/10.1145/3290511.3290524>
- [9] Beuran, R., Pham, C., Tang, D., Chinen, K. I., Tan, Y., & Shinoda, Y. (2017). Cytrone: An integrated cybersecurity training framework. In *Proceedings of the 3rd International Conference on Information Systems Security and Privacy (ICISSP 2017)* (pp. 157-166). SCITEPRESS – Science and

Technology Publications

- [10] Hendrix, M., Al-Sherbaz, A., & Victoria, B. (2016). Game based cyber security training: are serious games suitable for cyber security training?. *International Journal of Serious Games*, 3(1), 53-61.
- [11] Boopathi, K., Sreeiith, S., & Bithin, A. (2015). Learning cyber security through gamification. *Indian Journal of Science and Technology*, 8(7), 642-649.
- [12] Jin, G., Tu, M., Kim, T. H., Heffron, J., & White, J. (2018, February). Game based cybersecurity training for high school students. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education* (pp. 68-73).
- [13] Rana, S., & Alhamdani, W. Exploring the Need to Study the Efficacy of VR Training Compared to Traditional Cybersecurity Training. *International Journal of Computer and Information Engineering*, 15(1), 10-17.
- [14] Seymour, N. E., Gallagher, A. G., Roman, S. A., O'brien, M. K., Bansal, V. K., Andersen, D. K., & Satava, R. M. (2002). Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of Surgery*, 236(4), 458.

Hybrid Multipath Congestion Control

Akshit Singhal, Xuan Wang, Zhijun Wang, Hao Che, Hong Jiang

Akshit Singhal is with the University of Texas at Arlington,
United States (e-mail: akshit.singhal2@mavs.uta.edu).

Abstract

Multiple Path Transmission Control Protocols (MPTCPs) allow flows to explore path diversity to improve the throughput, reliability and network resource utilization. However, the existing solutions may discourage users to adopt the solutions in the face of multipath scenario where different paths are charged based on different pricing structures, e.g., WiFi vs. cellular connections, widely available for mobile phones. In this paper, we propose a Hybrid MPTCP (H-MPTCP) with a built-in mechanism to incentivize users to use multiple paths with different pricing structures. In the meantime, H-MPTCP preserves the nice properties enjoyed by the state-of-the-art MPTCP solutions. Extensive real Linux implementation results verify that H-MPTCP can indeed achieve the design objectives.

1 Introduction

Multiple Path Transmission Control Protocols (MPTCPs) that split a flow into multiple subflows to be sent via different paths towards the destination explore path diversity to improve flow throughput, network reliability and utilization. They play an important role in today's Internet, especially for end-to-end MPTCPs that do not explicitly involve the Internet routers for the control, as evidenced by the standardization of such protocols by IETF, e.g., [2]. Like end-to-end TCP, end-to-end MPTCPs can be deployed quickly at scale. Hence, the work in this paper aims at developing an end-to-end MPTCP.

A widely accepted design objective for MPTCP is three-pronged [1]: (i) the overall flow rate for an MPTCP flow should be at least as high as the highest flow rate a single-path TCP can achieve using any of the sub-flow paths of the MPTCP flow, and the single-path TCP flow that achieves this flow rate is called the best single-path TCP flow; (ii) to be fair to a single-path TCP, the total flow rate for any number of subflows of an MPTCP flow sharing a bottleneck link with a single-path TCP flow should not exceed the flow rate of the single-path TCP flow; and (iii) an MPTCP flow must be able to balance the load among subflow paths [1]. Notable MPTCPs that meet the three-pronged objective include Semi-coupled, LIA, OLIA and Balia [1], [3-5]. In this paper, we generally call these MPTCPs Equal bandwidth shared MPTCP (EMPTCP), because they tend to distribute the flow rates evenly among all flows, regardless how many subflows there are in a flow.

In this paper, we argue that the above three-pronged design objective should be augmented with yet another aspect, making it a four-pronged design objective, i.e. to provide a mechanism to incentivize users to use the protocol. Although the first of the three prongs of the above objective does provide some incentive for a user to use EMPTCP over single-path TCP, an EMPTCP may end up discouraging users from using it. This is simply because different paths used by an EMPTCP may be owned by different Internet service providers based on different pricing structures. For example, mobile devices, such as cell phones, are equipped with both WiFi connectivity and cellular data service (e.g., 3G/4G/5G) most of the time. Usually, the WiFi connection is without usage fee and the mobile data service may charge a usage fee based on the amount of data sent/received. Considering the scenario of a social gathering, e.g., a family party, where many guests may want to use their mobile phones to browse the Internet, watch online videos, sending/receiving messages and so on, via the host's WiFi network. This may result in low flow rates seen by and hence, poor Internet experiences for individual guests. When this happens, a guest with a cellular data service may be tempted to turn on an EMPTCP that meets the three-pronged design objective, thinking that by doing so, his/her cellular data service may help prop up the bandwidth needed to gain good experience at the cost of paying a small amount of cellular data service fee. In reality, however, he/she may end up with using the cellular data service almost entirely and receiving a hefty bill later. This may well be the case because an EMPTCP may attain the desired flow rate using the cellular connection only, hence giving up much of the free/low-cost bandwidth on the WiFi side. Obviously, this resource allocation is unfair to the guest who uses EMPTCP, and hence, would discourage him/her from using EMPTCP again in the future. This example clearly indicates that to incentivize users to use MPTCP in the face of multiple paths with different pricing structures, the three-pronged design objective should be augmented to a four-pronged one, with the fourth one being subflow path pricing structure aware.

A naive solution is to simply use a weighted single-path TCP for each subflow and assign a heavier weight to a subflow with a lower price. In this paper, we call this kind of MPTCP weighted MPTCP (WMPTCP). Note that an earlier MPTCP protocol, known as equal-weight MPTCP [3] is a special case of WMPTCP, which assigns the same weight to all the subflow paths. Again, using the above scenario as an example, one may assign, e.g., 75% and 25% weights to subflows using WiFi and cellular, respectively. By doing so, the guest who uses WMPTCP can get 75% of the single-path TCP flow rate from the WiFi side for sure. In the meantime, it can still compete for the bandwidth on the cellular side, but much less aggressively. Unfortunately, however, WMPTCP does not meet the first and the third prongs of the three-pronged design objective, meaning that it may lead to the overall flow rate lower than that of the best single-path TCP flow and cannot balance the load.

In this paper, we propose a new MPTCP that meets the four-pronged design objective. The idea is to combine an EMPTCP with a WMPTCP to come up with a hybrid MPTCP (H-MPTCP). H-MPTCP leverages the ability of

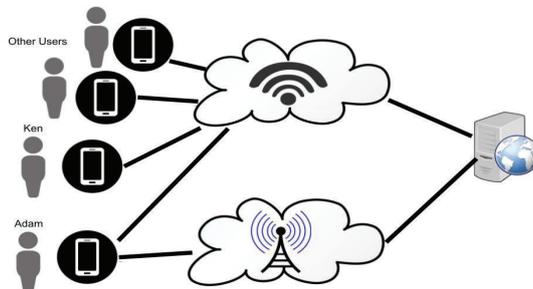


Fig. 1: A example network

WMPTCP to allow price-aware subflow path rate allocation and the ability of EMPTCP to meet the three-pronged objective, hence resulting in its ability to achieve the four-pronged design objective. We implement the proposed protocol in Linux Kernel based on the open source MPTCP source codes [6]. Extensive test results demonstrate that H-MPTCP can indeed achieve the four-pronged design objective. In the meantime, it outperforms EMPTCP, WMPTCP and some well-known MPTCP protocols (i.e., LIA [1] and Balia [4]) in terms of throughput and responsiveness.

2 Background and Motivations

A straightforward but naive approach to end-to-end MPTCP design is to simply run independent end-to-end TCP flows as subflows on different paths. This approach, however, is too aggressive and unfair to single-path TCP flows, and cannot balance the loads among subflow paths. This leads to the design of EWTCP [3], a member in the WMPTCP family. EWTCP attempts to achieve TCP fairness by modifying the previous approach, i.e., reducing the TCP window increase rate by a factor of $w_l = 1/S^2$ for all S TCP-based subflows. However, besides the lack of load balancing capability inherited from the previous approach, EWTCP may lead to inefficient use of networks and flow rate lower than the best single-path TCP. The shortcomings of EWTCP further leads to the design of a coupled congestion control algorithm, also known as Linked Increases Algorithm (LIA) [2], [1]. LIA is purposely designed to meet the three-pronged design objective and standardized by IETF. OLIA [5] improves over LIA in terms of the Pareto optimality. More recently, semi-coupled and Balanced linked adaptation (Balial) algorithm [4] were proposed to strike a good balance between TCP-friendliness, responsiveness and window oscillation, especially to further improve responsiveness when network condition changes. All these solutions (i.e., LIA, OLIA, semi-coupled, and Balial) meet the three-pronged design objective and hence, are members in the EMPTCP family.

Table 1: Flow rate allocations for the example

Scenario	WiFi=Cellular=100			Cellular=5, WiFi=100				
	Adam			Ken	Adam			Ken
Subflows	Cellular	WiFi	Total	WiFi	Cellular	WiFi	Total	WiFi
Best-case TCP	100(\$10)	N.A.	100	100	N.A.	50	50	50
EMPTCP	100(\$10)	0	100	100	5(\$0.5)	47.5	52.5	52.5
WMPTCP	57 (\$5.7)	43	100	57	5(\$0.5)	43	48	57

However, neither the WMPTCP family nor the EMPTCP family is capable of achieving the four-pronged design objective. To demonstrate this, we take a closer look at the family party example. Consider the network topology shown in Fig. 1. Adam is a MPTCP user who can use both WiFi free of charge and cellular with 10 cents per unit bandwidth per hour charge. Ken is a single-path TCP user who only uses WiFi for free. So, we have two users where Adam and Ken can both use WiFi service and Adam can also use cellular service. Assume that the maximum data rates for both Adam’s and Ken’s applications are 100 bandwidth units. The link bandwidths for both WiFi and cellular connections represent the capacity of the bottleneck link bandwidths of the connections. Now, we consider two different scenarios, where in the first scenario, both WiFi and cellular connections can support up to 100 units of bandwidth and in the second scenario where WiFi has 100 units of bandwidth but cellular has only 5 units of bandwidth. For both scenarios, EMPTCP will attempt to equalize the flow rate allocation between both users, whereas WMPTCP will attempt to allocate flow rates in proportion to their weights. For WMPTCP, we further assume that the weights assignments are 0.75 and 0.25 for WiFi and cellular, respectively. This will allow the subflow using the WiFi network to be allocated roughly 75% of the flow rate of a single-path TCP flow on the same network, resulting in a potentially much reduced flow rate needed for the subflow on the cellular network side.

Table I gives the flow rate allocation and costs for both Adam and Ken as well as the best single-path TCP flow when Adam uses either EMPTCP or WMPTCP for both scenarios. For the first scenario on the left, EMPTCP for Adam equalizes the flow rate allocation by fully utilizing the cellular connection, yielding to the single-path TCP flow for Ken completely on the WiFi connection. This results in Adam paying for the full usage fee of the cellular link of \$10 per hour, whereas Ken enjoys the same flow rate performance for free. In contrast, WMPTCP that has a weight of 0.75 for WiFi and 0.25 for cellular is able to rip 43% of the WiFi link bandwidth and 57% of cellular link bandwidth, resulting in the same total flow rate of 100 units, as the cost of \$5.7 per hour, more than 40% lower than the case of EMPTCP. In other words, in this scenario, WMPTCP offers better incentive to users to use MPTCP than EMPTCP by reducing the usage costs. These weights may be set as a function of the pricing models of individual subflow paths to further incentivize users to use it.

Table 2: MPTCP Design objectives and performance parameters

Solutions	Design Objectives				Performance Evaluation
	Best-case TCP	TCP Friendliness	Load Balancing	Adoption Incentive	Responsiveness
EWTCP [3]	No	Yes	No	Yes	High
LIA [1]	Yes	Yes	Yes	No	Low
Balia [4]	Yes	Yes	Yes	No	Medium
H-MPTCP	Yes	Yes	Yes	Yes	High

For the second scenario on the right in Table I, to equalize the flow rate allocation, EMPTCP has to take up a much bigger chunk of the bandwidth from WiFi connection, i.e. 47.5, to be exact. So, EMPTCP is able to achieve a total bandwidth of 52.5 units. In contrast, with the weight of 0.75, WMPTCP still can only grab 43% of the bandwidth from WiFi connection, leaving it with a total flow rate, 48 units, lower than both EMPTCP and the flow rate of 50 for the best single-path TCP.

The above example clearly demonstrates the need for a new MPTCP, which motivates us to propose H-MPTCP. As we shall demonstrate, in scenario 1, H-MPTCP will automatically select WMPTCP over EMPTCP, whereas in scenario 2, it will automatically select EMPTCP. In either case, H-MPTCP results in better than the best single-path TCP performance at a lower cost than EMPTCP, hence, meeting the four-pronged design objective.

Based on the discussion so far and the performance data to be presented in the later sections, Table 2 provides a summary of the features for some notable end-to-end MPTCP protocols as well as H-MPTCP. As one can see, H-MPTCP possesses the most desirable features among all MPTCPs.

Finally, we note that there are other works that focus on specific aspects of the MPTCP protocol design challenges, e.g., bottleneck detection [12]–[14] and packet scheduling [8]–[11]. However, they are not concerned with new end-to-end MPTCP protocol design, the main focus of this paper.

3 Hybrid Multipath Congestion Control

In this section, we first briefly describe EMPTCP and WMPTCP. Then we introduce H-MPTCP.

3.1 EMPTCP

For EMPTCP, the congestion window size (W_l) for subflow (l) in each RTT in the slow start phase is the same as that in TCP Reno. The congestion window

change for subflow l in each RTT in the congestion avoidance phase is,

$$\Delta W_l = \begin{cases} \frac{W_l}{\tau_l \sum_{j=1}^S W_j / \tau_j} & \text{if } cg = 0 \\ -\frac{W_l}{2} & \text{if } cg = 1. \end{cases} \quad (1)$$

where W_j and τ_j are the congestion window size and RRT for subflow j ($j = 1, \dots, S$). In this paper, we choose semicoupled mentioned in Balia [4] for EMPTCP, but in general, we can use any members in the EMPTCP family.

From Equation (1), we know that the subflow increase rate is proportional to the ratio of the subflow rate with the overall flow rate, i.e., x_l/x while its decrease rate is the same as that in TCP Reno. It means that the subflow rate increase is slower for a multipath flow with higher flow rate, or the network tries to evenly allocate the flow rate to all flows. For single path TCP, $W_l = W$, then the congestion window change is degenerated to single path TCP Reno.

3.2 WMPTCP

For WMPTCP, the change in congestion window size (W_l) for subflow (l) in each RTT in the slow start phase is same as that in TCP Reno in the slow start phase. The congestion window change for subflow l with weight (ω_l) in each RTT in the congestion avoidance phase is,

$$\Delta W_l = \begin{cases} \omega_l & \text{if } cg = 0 \\ -\frac{W_l}{2} & \text{if } cg = 1. \end{cases} \quad (2)$$

From Equation (2), we know that the rate increase for a subflow l is proportional to the weight ω_l . If $\omega_l < 1$, the subflow increase rate is smaller than that in a single path TCP flow. If we set $\omega_l < 1$ for any subflow l , then each subflow obtains no more flow rate than that of a single path TCP in a shared link. WMPTCP can allocate more bandwidth than the best single path TCP in some cases, but it may not guarantee the rate of WMPTCP flow always be no less than the best single TCP flow rate as shown in Table 1. If $\omega_l = 1$ for any subflow l , the congestion window change of each subflow is the same as in TCP Reno. In this case, a multipath flow is just the combination of S individual single TCP flows.

3.3 H-MPTCP

To meet the four-pronged design objective, we now design H-MPTCP. We set the weight $1/S \leq 1$ for each subflow of an MPTCP flow with S subflows. With such weight assignment, each subflow can get no more flow rate than that of a single TCP flow rate for a shared link. But the overall flow rate of an MPTCP may have chance to get more rate than its best single path TCP flow. In slow start phase, H-MPTCP behaves the same as TCP-Reno. In the congestion avoidance phase, H-MPTCP selects the larger subflow increase rate from the

increase rates of EMPTCP and WMPTCP, i.e.,

$$\Delta W_l = \max(\Delta W_l^{Equal}, \Delta W_l^{Weighted}), \quad (3)$$

where ΔW_l^{Equal} and $\Delta W_l^{Weighted}$ are the congestion window increases in each RTT defined in Equation. (1) and Equation. (2) in case of no congestion, respectively.

Under "normal" situation, e.g., scenario one in the example given in Section 2, WMPTCP is likely to be automatically selected because it is more responsive/aggressive than EMPTCP [4]. EMPTCP will take over only under "abnormal" situation, e.g., scenario two in Section 2, when WMPTCP fails to reach flow rate equal to or higher than the best single-path TCP. In this case, EMPTCP will ensure that flow rate will be balanced to further improve the flow rate performance.

In summary, in the slow start phase, we have,

$$\Delta W_l = \begin{cases} W_l & \text{if } cg_l = 0 \\ -\frac{W_l}{2} & \text{if } cg_l = 1. \end{cases} \quad (4)$$

and in the congestion avoidance phase, we have,

$$\Delta W_l = \begin{cases} \max(\omega_l, \frac{W_l}{\tau_l \sum_{j=1}^S W_j / \tau_j}) & \text{if } cg = 0 \\ -\frac{W_l}{2} & \text{if } cg = 1. \end{cases} \quad (5)$$

4 Performance Evaluation

In this section, we evaluate the proposed H-MPTCP compared to the EMPTCP and WMPTCP and to the existing MPTCP solutions LIA [1] and Balia [4] in Linux kernel implementation on a testbed. The hosts (i.e., S_1 , S_2 and D) are Dell Poweredge servers, each equipped with 8-core processors with 10 GB memory and running Linux 16.04. Node b is a Dell N4032F switch with multiple 1 Gbps Ethernet interfaces running Ubuntu 16.04.1 LTS (Linux kernel 4.19.98). The link bandwidth for all the links can be configured at any rate lower than or equal to 1 Gbps through the networking interface traffic control command tc , allowing for the testing of the MPTCP responsiveness to sudden link bandwidth changes. H-MPTCP is implemented by modifying the open source Linux kernel codes of LIA and Balia [6].

We use the network topology as shown in Fig. 2 for both implementation. In the network, source S_1 has a single path TCP flow x_2 running TCP Reno, and source S_2 has an MPTCP flow with two subflows x_{11} and x_{12} . Subflow x_{12} representing Wi-Fi link shares the link from node b to destination D with flow x_2 . We set different bandwidth combination of C_1 and C_2 to test the flow rate allocation in different MPTCP solutions. This network topology is also adopted in Balia [4] for the testing of their solution.

We test the performance of the proposed protocol in three different cases, each has a different network setup (i.e., different bandwidth C_1 and C_2). For

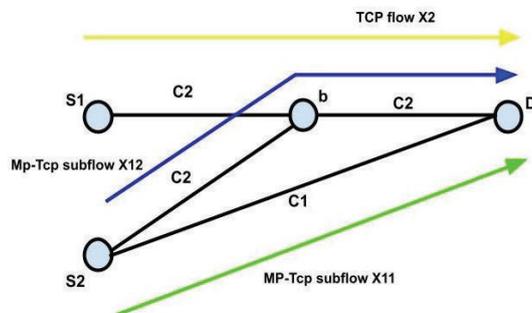


Fig. 2: Network Topology for Performance Evaluation

the test, each source has a huge data to send to the destination D , and each flow can be viewed as an infinite data flow during our test (i.e., the test is finished before the data is fully sent out). In this case, MPTCP user is trying to get the maximum bandwidth available to improve the flow completion time. Each flow reaches its stable rate quickly and we present the first 10-second results in the first two cases and the first 20-second for the third case. The weight values in WMPTCP is set as $1/2$ for both subflows x_{11} and x_{12} . Flow x_2 and the two subflows of x_1 (i.e., x_{11} and x_{12}) start at the same time in each test.

Case 1: First we test the performance of MPTCP solutions in a symmetric network by set $C_1 = C_2 = 1024$ Mbps (i.e. 1 Gbps link). With this setup, the rate allocation for EMPTCP are $x_{11} = 1024$ Mbps, $x_{12} = 0$ Mbps (i.e., the overall flow rate $x_1 = 1024$ Mbps) and $x_2 = 1024$ Mbps and for WMPTCP are $x_{11} = 1024$ Mbps, $x_{12} \approx 341$ Mbps (i.e., the overall flow rate $x_1 \approx 1365$ Mbps) and $x_2 \approx 683$ Mbps, respectively.

Fig. 3 shows the results of rate allocation (the average rate in a second) in the proposed protocol compared with EMPTCP, WMPTCP, LIA and Balia. The rate of each flow/subflow is also presented using dashed line with the same color as the real measured rate in an MPTCP solution. As LIA and Balia have equally sharing of the bandwidth, they achieve similar rate allocations as EMPTCP, and hence the theoretical rates in EMPTCP is also listed in LIA and Balia results for comparison. From the results, we know that the rate allocations in EMPTCP and WMPTCP are closely matches to their corresponding theoretical rates with little lower rates. The results indicate that EMPTCP and WMPTCP can really achieve the rate allocation. The lower rate is due to the discrete time control and network condition feedback delay, these make the protocols unable to achieve full bandwidth usage. LIA and Balia also achieve their design objective, i.e., sharing network bandwidth to all flows as even as possible. In this case, WMPTCP allows the MPTCP flow x_1 to obtain higher flow rate than the best single path TCP flow. As WMPTCP has higher flow rate than EMPTCP, H-MPTCP selects WMPTCP and hence achieves the same flow rate allocation

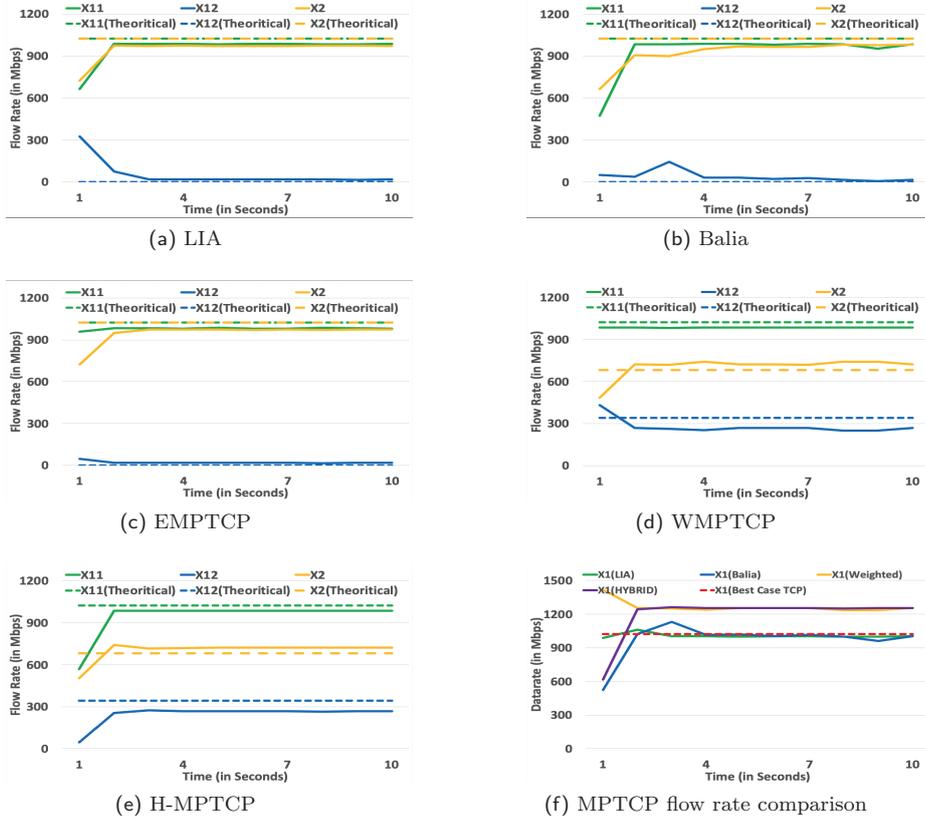


Fig. 3: Performance for MPTCP in Case 1

as that in WMPTCP (see Fig. 3 (e)).

The overall flow rate of the MPTCP flow (i.e., x_1) and its theoretical best single path TCP flow rate (denoted as dashed red line) are shown in Fig. 3 (f). From the results, we can see that all EMPTCP, LIA, and Balia are just achieve close to the theoretical best single path TCP flow rate. Although they make flow x_2 to get higher flow rate, they may have no incentives by usage of MPTCP. WMPTCP/H-MPTCP achieves higher MPTCP flow rate than the best single flow rate. It also benefits flow x_2 in case that x_1 was using a single path TCP as congestion control and chooses the path x_{12} . Hence WMPTCP/H-MPTCP gives more incentives to users to apply MPTCP.

Case 2: Now we test the performance of MPTCP solutions in an asymmetric network by set $C_1 = 8$ Mbps and $C_2 = 1024$ Mbps. With this setup, the rate allocation for EMPTCP are $x_{11} = 8$ Mbps, $x_{12} = 508$ Mbps (i.e., the overall flow rate $x_1 = 516$ Mbps) and $x_2 = 516$ Mbps and for WMPTCP are $x_{11} = 8$ Mbps, $x_{12} \approx 341$ Mbps (i.e., the overall flow rate $x_1 \approx 349$ Mbps) and $x_2 \approx 683$

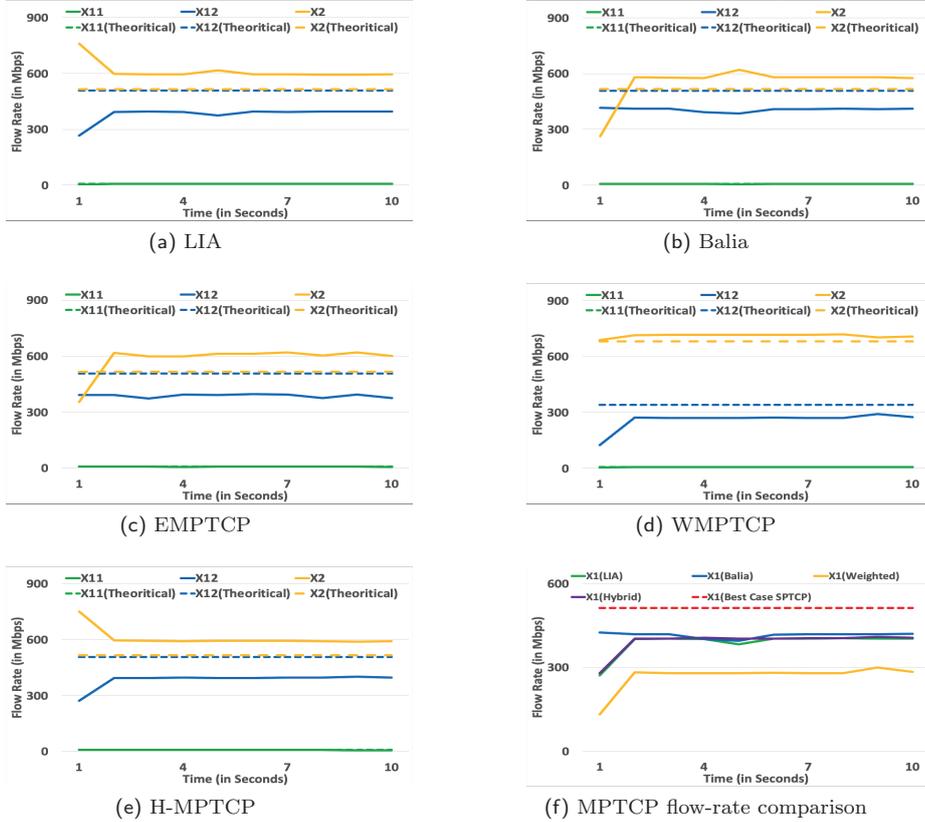


Fig. 4: Performance for MPTCP in Case 2

Mbps, respectively. In this case, x_1 in WMPTCP has lower flow rate than that of the best single path TCP flow. As EMPTCP can achieve higher flow rate, H-MPTCP selects EMPTCP increase rate and hence achieves the same flow rate allocation as that in EMPTCP, as shown in Fig. 4 (c) and Fig. 4 (e). Fig. 4 shows the results of rate allocation in the MPTCP solutions. Again, we see that the rate allocations of EMPTCP/H-MPTCP and WMPTCP closely match to their corresponding theoretical rates. From Fig. 4 (f), we can see that all the protocols can still get more rates (516 vs. 512 Mbps theoretically) than that of the best single path TCP. Although the path S_2 to D has very limited bandwidth but EMPTCP, H-MPTCP, LIA and Balia can still balance the rates and benefit the single path flow x_2 . This indicates that MPTCP can be useful to balance the traffic and increase the network utilization. From both case 1 and case 2, we know that H-MPTCP can really guarantee an MPTCP flow to achieve at least the best single path TCP flow rate as that in EMPTCP while it can try to obtain higher flow rate as many as possible and hence can encourage

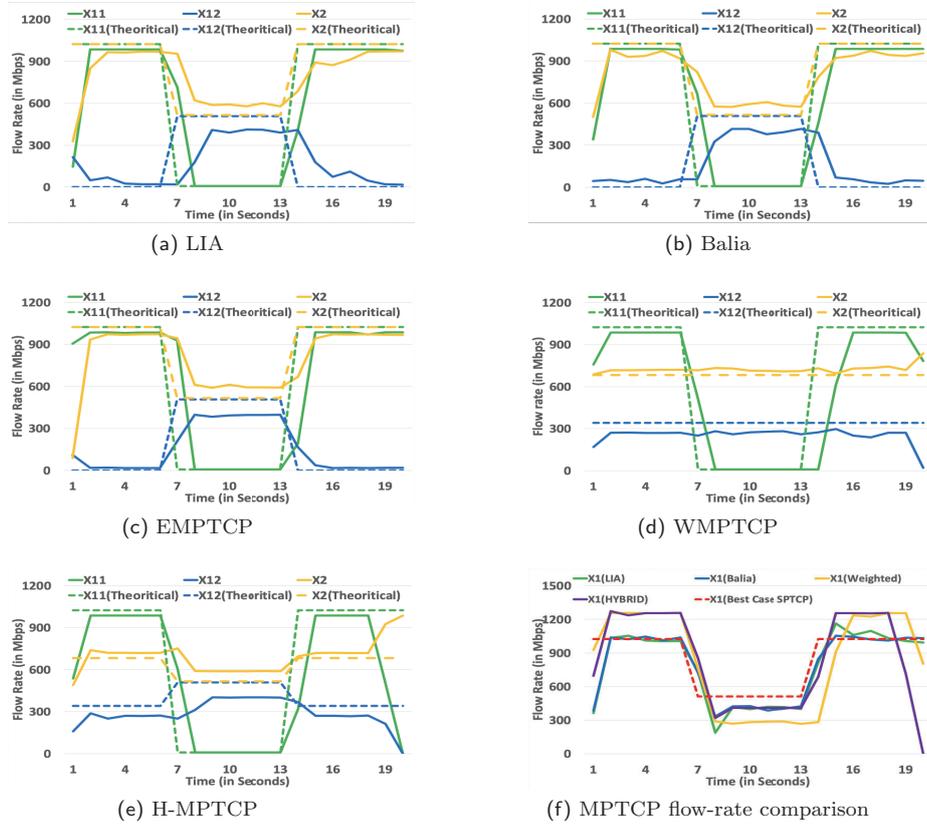


Fig. 5: Performance for MPTCP in Case 3

users to use it.

Case 3: Finally, we test the performance of the proposed protocol in a dynamic network environments to see the responsiveness to network changes. In this case, we set the link bandwidth $C_1 = C_2 = 1024$ Mbps in the first 6 seconds and change the bandwidth $C_2 = 8$ Mbps in the next 7 seconds (i.e., from second 7 to second 13) and then C_2 is switched back to 1024 Mbps. As this setup, the rate allocation in the proposed protocol during the first 6 seconds and after the second 13 are the same as that in the case 1, and the rate allocation during the time period between second 6 to second 13 is the same as that in the case 2. For H-MPTCP, it always selects the higher flow from EMPTCP and WMPTCP, and hence it selects WMPTCP congestion control in the first 6 seconds and after second 13 and chooses EMPTCP during the time period from second 6 to second 13.

Fig. 5 shows the results of rate allocation in the five MPTCP solutions. From the results, we see that EMPTCP, WMPTCP and H-MPTCP can closely

match to their rates in this dynamical network environment. Now let us look at the flow rate change during the network bandwidth changes.

First, we note that the subflow rate x_{12} in WMPTCP has no change, because the weight for the subflow is only depend on the number of subflows, and hence the flow rate allocation in the shared link does not change.

Second, in EMPTCP, LIA and Balia, we can see that the subflow rate of x_{11} (denoted as green line) drops to 8 Mbps almost the same time in all the three protocols, this is because subflow x_{11} does not compete the bandwidth with any other flow/subflow. As the subflow rate x_{11} is reduced, the rate of subflow x_{12} is then increased in all the three protocols, and hence the flow rate x_2 should be reduced because of the shared bandwidth with x_{12} . From Fig. 5 (a), Fig. (b) 5 and Fig. 5 (c), we can see that the transition time for flow x_2 (i.e., yellow line) dropped to its new balanced rate has almost the same time in the three protocols.

Third, from Fig. 5 (a), Fig. (b) 5 and Fig. 5 (c), we know that the transition time for subflow x_{12} (i.e., blue line) reaching its new balanced rate is different in the three protocols. In LIA and Balia, the transition time takes about 3 seconds from second 6 to about second 9 while in EMPTCP, the transition time is about 2 seconds from second 6 to about second 8. This indicates that EMPTCP can be more quickly to catch up the network condition change to reach the new balanced rate than LIA and Balia. This is because Semicoupled has higher responsiveness as compared to LIA and Balia as mentioned in the Balia paper [4].

Fourth, H-MPTCP switches from WMPTCP at second 6 and switches back to Weighted-MPCTP at second 13. The transition time in H-MPTCP is similar as that in EMPTCP as shown in Fig. 5 (d).

From the results, we can also see that H-MPTCP always choose the higher flow rate from EMPTCP and WMPTCP, and hence benefits the user to applying MPTCP. The results show that EMPTCP/H-MPTCP can quickly response to the network condition change and reaches to its new balanced state.

Through the three-case studies, we conclude that the proposed H-MPTCP solution is ready to be applied in today's Internet and can meet the MPTCP design goals and give more incentives for users to apply the MPTCP solutions.

5 Conclusion

In this paper, we propose a Hybrid-MPTCP(H-MPTCP)that always achieves higher flow rate from EMPTCP and WMPTCP and provides a built-in mechanism that can encourage users to apply it over other MPTCPs as well as single-path protocols with different pricing structures. Extensive real Linux implementation test results verify that the proposed H-MPTCP can indeed achieve the design objectives.

References

- [1] Wischik, Damon and Raiciu, Costin and Greenhalgh, Adam and Handley, Mark, *Design, Implementation and Evaluation of Congestion Control for Multipath TCP*, Proceedings of NSDI, 2011.
- [2] Raiciu, Costin and Handley, Mark and Wischik, Damond, *Coupled Congestion Control for Multipath Transport Protocols*, RFC 6356, 2011.
- [3] Honda, Michio and Nishida, Yoshifumi and Eggert, Lars and Sarolahti, Pasi and Tokuda, Hideyuki, *Multipath congestion control for shared bottleneck*, Proceedings of PFLDNeT workshop, 2009.
- [4] Peng, Qiuyu and Walid, Anwar and Hwang, Jaehyun and Low, Steven H, *IEEE/ACM Transactions on Networking (TON)*, RFC 6356, 2016.
- [5] Khalili, Ramin and Gast, Nicolas and Popovic, Miroslav and Le Boudec, Jean-Yves, *MPTCP is not Pareto-optimal: Performance issues and a possible solution*, IEEE/ACM Transactions on Networking, 2013.
- [6] C. Paasch, S. Barre, *Multipath TCP in the Linux Kernel*, <http://www.multipath-tcp.org/>, 2019.
- [7] Kheirkhah, Morteza and Wakeman, Ian and Parisi, Goerge, *Multipath-TCP in ns3*, Proceedings of ACM Workshop on ns-3, 2014.
- [8] Mirani, Farhan H and Boukhatem, Nadia and Tran, Minh Anh, *A data-scheduling mechanism for multi-homed mobile terminals with disparate link latencies*, IEEE 72nd Vehicular Technology Conference-Fall, 2010.
- [9] Wei, Wenjia and Xue, Kaiping and Han, Jianping and Wei, David S. L. and Hong, Peilin, *Shared Bottleneck-Based Congestion Control and Packet Scheduling for Multipath TCP*, IEEE Transaction on Networking, 2020.
- [10] Xu, Changqiao and Li, Zhuofeng and Li, Jinglin and Zhang, Hongke and Muntean, Gabriel-Miro, *Cross-layer fairness-driven concurrent multipath video delivery over heterogeneous wireless networks*, IEEE Transactions on Circuits and Systems for Video Technology, 2014.
- [11] Xue, Kaiping and Han, Jiangping and Ni, Dan and Wei, Wenjia and Cai, Ying and Xu, Qing and Hong, Peilin, *DPSAF: Forward prediction based dynamic packet scheduling and adjusting with feedback for multipath TCP in lossy heterogeneous networks*, IEEE Transactions on Vehicular Technology, 2017.
- [12] Hassayoun, Sofiane and Iyengar, Janardhan and Ros, David, *Dynamic window coupling for multipath congestion control*, 19th IEEE International Conference on Network Protocols, 2011.

- [13] Ferlin, Simone and Alay, Özgü and Dreibholz, Thomas and Hayes, David A and Welzl, Michael, *Revisiting congestion control for multipath TCP with shared bottleneck detection*, IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications, 2016.
- [14] Wei, Wenjia and Wang, Yansen and Xue, Kaiping and Wei, David SL and Han, Jiangping and Hong, Peilin, *Shared bottleneck detection based on congestion interval variance measurement*, IEEE Communications Letters, 2018.

SPA-VNDN: Enhanced Smart Parking Application by Vehicular Named Data Networking

Bassma Aldahlan,¹ Zongming Fei²

Abstract—Today, there is great attention on exchanging information regarding smart parking applications to assist drivers in finding suitable parking spaces and making a reservation. Smart parking applications are enhanced by Vehicular ad-hoc Network, which involves assisting drivers in requesting the availability of parking spaces and making a reservation for a period of time ahead of time.

Named Data Networking (NDN) is a future Internet architecture that benefits Vehicular Ad-hoc Networks because of its clean-slate design and pure communication model. In this paper, we proposed an NDN-based framework for smart parking that involves a fog computing architecture. The proposed application has two main directions : First, we allow drivers to query the number of parking spaces in a particular parking lot. Second, we introduce a technique that enables drivers to make an intelligent reservation before their arrival time. We also introduce a “push-based” model that supports the NDN-based framework for smart parking application.

To evaluate the proposed solution performance, we evaluate the function for finding parking lots with available parking spaces and the function for reservation a parking space. Our system shows high-performance results in terms of response time and push overhead. The proposed reservation application performs better compared with a baseline approach.

Index Terms—Named Data Networking, Vehicular Named Data Networking, Smart Parking Applications, Fog Computing.

I. INTRODUCTION

Smart Parking applications have attracted many researchers in recent years, especially with the increasing growth in the number of vehicles in urban cities. Finding an available parking spot could be stressful in big crowded cities. Drivers may circle around several

rounds to find the parking spaces. It also can be time-consuming when a vehicle tries to find a parking space while wasting gasoline consumption. When more than one vehicle attempt to find parking space in a mostly vacant parking area, the traffic congestion and pollution could be increased. To solve this issue, we need to have a smart parking application to find a suitable parking lot sufficiently.

Some existing projects such as SFpark (San Francisco) and LA Express Park (Los Angeles) [1] take advantage of smart parking applications to find available parking space with the association of smartphones. In those projects, many sensors were deployed to collect the sensed data and send it to the nearby meters. Drivers with smartphone devices can request the availability of parking spaces. The projects are promising solutions; however, those projects still have to be improved from several prospective to satisfy the driver’s need in real-time. First, all requested drivers can get the same information regarding parking space availability. Second, real-time availability data is only available for vehicles near the parking location. Third, installing sensors should be more robust to avoid road obstacles, and transforming and processing those sensed data should be through a wireless network in a way that eliminates sending high traffic to the same spectrum.

Intuitively, the development of a smart parking solution can reduce parking search time and traffic congestion. It can also alleviate environmental pollution and fuel consumption. In addition, smart parking brings economic benefits that increase activities and business opportunities.

In this context, Vehicular Ad-hoc Network (VANET) is a critical technology that can be considered in smart parking applications in order to allow vehicles to communicate with each other or enable infrastructure such as Road Side Unit (RSU) and Base Station (BS) to serve the smart parking systems. In VANETs, the main communication modes are Vehicle-TO-Vehicle (V2V), in which VANET can form a pure communications mode between only On-Board Units (OBU) via a wireless, IEEE 802.11p. Second, Vehicle-TO-Infrastructure(V2I), the communication between vehicles and infrastructure

¹Bassma Aldahlan is with Yanbu University College, Yanbu, Saudi Arabia (correspondence author’s email: aldahlanb@rcyci.edu.sa, baldahlan@uky.edu)

²Zongming Fei is with University of Kentucky, Lexington, KY 40503 USA. He is now with the Department of Computer Science, (email: fei@netlab.uky.edu)

such as RSU and BS through either IEEE 802.11p or cellular communication.

Using V2V communication in smart parking systems could face some challenges. Vehicles in VENETs are lack in sharing their sensed information via V2V communication. Also, sending that huge information through rely nodes can be energy-consuming, while VANET supports real-time information dissemination. Utilizing V2I communication to support the smart parking system performs better. However, deploying RSUs everywhere can be costly in terms of installing, deploying, and maintenance. Also, the capability of RSUs in terms of computation and storage can be limited.

Currently, smart parking applications use an IP-based network to enable communication. However, IP-based network faces many issues. First, it is not easy to assign a unique IP address to all sensors and vehicles because of the limited resources. Second, it is not trivial to achieve a request via those unique addresses in a mobile environment like VANETs. Third, the IP-based method depends on where the data is residing, not on where the content is, which forces us to find the target node that has the information regarding the smart parking system [2].

Vehicular Named Data Networking (V-NDN) has brought more attention in academic and industrial fields. It replaced the traditional IP network with Named Data Networking to overcome the limitations in IP networks. V-NDN is a promising network architecture that supports high mobility and intermittent disconnectivity. The nature of NDN, where the data content can be fetched from the producers or any associated cache, can play a role in VANETs, such as enhancing self-certifying and layer security. NDN supports multicast communication in a mobile environment; however, there are still some challenges that need to be addressed.

This paper provides a design of an NDN-based framework for the smart parking application, which takes the benefit of the fog-computing to answer the query from drivers about whether there is parking space available in a particular geographical area. Further, the application introduces a new technique that allows drivers to make a reservation for some time ahead of time. We assume sensors are installed in the parking lot to collect availability information. Sensors on vehicles can also contribute to the parking application. We also introduced a "push-based model" to integrate sensor nodes into the system for sensor nodes to deliver the information proactively.

II. RELATED WORK

Recently, smart parking applications have been considered as important applications in VANET. Besides that, smart parking applications have attracted many researchers in industrial and research fields. Many survey

papers [1], [3]–[5] are conducted to study and investigate the smart parking applications. Lin et al. [1] propose a survey paper that depends on a smart parking ecosystem and a comprehensive classification of smart parking applications by identifying the functionalities of the existing work and their problematic focuses. They divided the existing work into three macro-themes based on the target of each work: First, the information collection tools that assist in gathering data such as sensors, actuators, and parking meters. Second, the system deployment, in which it handles the software system exploitation that assists with the statistical analysis of the collected data and provides data prediction. Third, the service dissemination that deals with the relationship between the gathered information and some social feature related to the driver's behaviors, including some aspects such as parking competition, information dissemination, and parking behavior. Both works [4] and [3] discussed some technologies about parking availability monitoring, parking reservation and dynamic pricing. In [4], the authors investigate the different technologies regarding parking availability monitoring. These technologies help in the dissemination of the associated information regarding parking space availability and parking reservation. The authors aimed to ensure higher customer satisfaction and increase revenue from parking services. Parking reservation allows the drivers to reserve the parking space in advance before arriving at the parking area. The authors studied the parking infrastructure from different perspectives, including the variant parking lots infrastructure. Faheem et al. [5] present a review paper for showing the different Intelligent Parking Services (IPS), which are used for parking guidance and parking facility management. The authors represent all the techniques that contribute to having an efficient and modern parking system. The authors provide economic analysis for assessing and exploring the project's feasibility.

The following work has been done on parking reservations and competition between vehicles. Parking reservation enables drivers to reserve the parking spaces before their arrival time. Integrating reservation policies with smart parking provides many benefits for both the drivers and managers of parking areas. Delot et al. [6] proposed a reservation protocol that allows vehicles to search for the available parking slots that are related to their requested event. Vehicles with similar interests in a relevant event will be cooperatively gathered. Besides that, the relevant direction of vehicles is taken into account to support the reservation strategy. They aim to avoid competition between the vehicles. Doulamis et al. [7] introduce an intelligent parking reservation management that relies on optimal strategy, aiming to promote service quality of drivers and increase the parking spaces utilization. The proposed approach utilizes the interval

scheduling principles represented as a list of parking requests that are provided as a set of requested time intervals. They presented a scheduler that can determine whether to accept the task and assign it to some resource or to deny it. They also introduced an adaptive pricing policy that is proportional to the rejection probability of a parking request.

Karbab et al. [8] proposed a scalable and low-cost car parking framework (CPF). The authors introduced driver guidance, automatic payment, parking lot retrieval, security, and vandalism detection. They also introduced a standard I2C protocol to cluster a group of sensors into a single mote. They used smartphone applications to reserve parking slots and hybrid wireless communications. The GPS helps in getting the real-time location and guidance toward the destination. However, smartphone applications are only beneficial if the driver is within 2 Km of the location.

These smart parking applications involve a high amount of data traffic caused by a variety of sensors. This high amount of traffic caused by the Internet of Things (IoT) sensed data are considered. Both [9] and [10] have been studied the ability of cloud computing, especially fog computing, on VANET, to support the computational demands and reduce the response time. The fog computing at edges can bring the benefits of computing, processing, and storage to the edges to be more closer to the sensor devices. Houet al. [11] introduced the idea of exploiting fog computing in VANET. Vehicular fog computing (VFC) utilizes a vehicular node to act as a fog node for communication and computation. They used both the mobile and parked vehicles as infrastructure to perform communication and computation. Xiao and Zhu [12] also introduced a similar idea of vehicular fog computing. They considered the connected vehicles as mobile fog nodes. They presented cost-effective and on-demand fog computing for vehicular applications. Tang et al. [13] proposed a parking slot allocation strategy that considered the real-time parking slots information by exploiting the fog computing-based capabilities on a smart parking architecture, aiming to enhance smart parking in real-time. The deployed fog nodes (RSUs) at parking lots can communicate cooperatively with each other wirely, allowing parking request processing. Meanwhile, the centralized cloud can promote smart parking capability by enforcing global optimization on parking requests. They provided an allocation strategy that considered the comprehensive factors which can affect the decision marking, such as walking costs, driving costs, and waiting for costs. Installing, deploying, and maintaining RSU for every parking to achieve one-hop communication could be expensive. Yi et al. [14] proposed a packing reservation auction system that uses cloud fog computing with the association of parked

vehicle cloud. They aim to guide mobile vehicles to the available parking spaces with less effort. They also used the fog capability of parked vehicles to help the delay-sensitive computing services by monetary rewards to compensate for their service cost and allocate the workload to each CPU. This solution results in performance due to the enhancement of the fog node controller, the smart vehicles, and the parking places. However, this work considered only the parked vehicles located in some parking areas and did not consider on-street parking. The authors did not mention any criteria to select those parked vehicles.

Despite all the previous solutions, the smart parking application is still in its early stages and needs improvements. In this paper, we develop a VNDN-based smart parking application using a fog computing architecture that supports IoT data collection and avoids the overcrowdedness of the central cloud approach.

III. SPA-VNDN: DESIGN

Smart parking application is developed to provide two functions. One is to allow a driver to find parking lots that have available parking spaces. The driver can send out the request as an interest packet in the NDN architecture and the replies can come from multiple parking lots. It is up to the driver to pick any specific parking lot. Normally, the driver wants to find available parking space at this moment. We will discuss a case in which the driver may be interested in the availability in the future. This is related to the function we will discuss next.

The other function is to allow a driver to reserve a parking spot in a particular parking lot for a period of time. This can happen after the driver finds a list of parking lots with available spaces in some future time. Then the driver picks one to make the reservation with the arrival time and departure time specified. Similar to the previous function, the driver can also provide geographical-area information in the request packet so that it can be forwarded to the correct location, rather than using the broadcast method.

The smart parking application adopts a fog cloud architecture to provide a flexible structure and improve the response time. NDN naming scheme is designed to identify the request, either for finding an available parking space or making a reservation for the future. We introduce a push-based model into the NDN architecture to streamline the process for allowing sensor nodes installed to get the information into the system.

A. Design Assumption of Smart Parking Application

We consider a topology with roads with traffic in both directions and interactions where decisions are made to change directions. In addition, sensor nodes will be

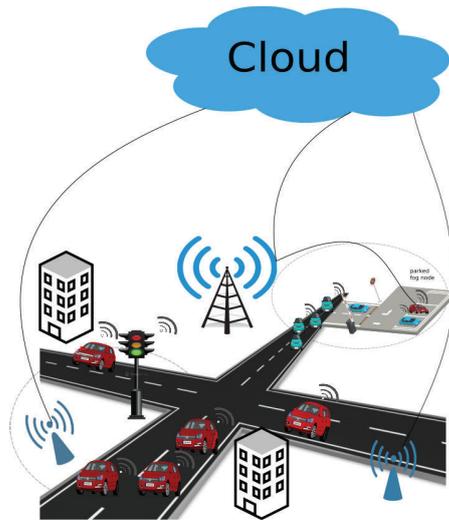


Fig. 1. Network Topology

installed to collect real-time information, and cloud computing elements will be used to meet the communication and computing requirements of the application. Figure 1 illustrates the main components.

- 1) **Vehicles:** Vehicles can communicate with other vehicles and with infrastructure units (fog nodes). They are equipped with a Global Positioning System (GPS), which can tell the location of the vehicle. We can classify vehicles into two types. One is the static vehicles (parked vehicles) that are parked in a parking garage or the curbside of roads. The other is dynamic vehicles (mobile vehicles) that are moving.
- 2) **Sensors:** Sensors are installed to monitor a variety of status. In particular, the parking application may have sensors to monitor whether a parking spot is free or not. In addition, modern vehicles are equipped with a lot of sensors that can also contribute information to the parking application, such as location, speed, and direction.
- 3) **RSUs:** Road Side Units (RSUs) are computing, storage and communication resources install on the road side for supporting modern transportation systems. They are an important component of VANET and will be a good candidate as the fog node in the fog cloud architecture. However, not all roads have RSUs installed. Vehicles and RSUs communicate through WiFi IEEE802.11p/WAVE technology and the quality of the communication can be significantly better than vehicle to vehicle communication.
- 4) **Dedicated Servers:** Dedicated servers may be installed in big parking garages for managing the parking space. They can also be incorporated into

the fog cloud architecture.

- 5) **Central Cloud:** Central cloud provide a central storage and processing service for the fog nodes connected to the architecture. It provide a global information necessary for coordination among the fog nodes.

A1. Fog Cloud Architecture: We considered several architectural candidates for the smart parking application. One possibility is that we can depend on RSUs and dedicated servers on larger parking garage to handle the requests, without any help from the cloud architecture. There are four limitations. First, there are smaller parking lots without dedicated servers or not covered by RSUs. We will not be able to include those parking spaces into the application. Second, if the dedicated servers or the RSUs crashes, the information stored there will be lost. For example, the reservation information will not be able to be recovered and agreements with customers will not be able to be honored. Third, different parking lots will not be able to cooperate with each other, such as recommending alternative parking places if the current lot is full. And finally, deploying RSUs and dedicated servers can be costly.

The second possibility is that we can use a centralized cloud server to handle all the requests from users. There are two potential issues. First, the central cloud becomes the concentration point of all requests. It can take a longer time for a user to get the response than the case in which a request can be processed by a distributed and close-by server. Second, the state information collected by the sensors will be sent over to the cloud. This can be huge, depending on the scope of parking spaces the central cloud is handling.

Instead, we consider an architecture based on the fog cloud computing. We still have a central cloud service that has a large pool of processing and storage resources that can be allocated on demand to meet the ever-changing requirement of the system. In addition, we have fog nodes distributed in different parts of the system, located at the edge of cloud and close to the parking space they serve. They will handle the request for the spaces they are in charge. The central cloud server also acts as the backup service in case of some critical states at distributed fog nodes are lost. It also has the authority to manage the fog infrastructure by authorizing the setup and teardown of fog nodes. The fog nodes can be RSUs, dedicated servers, or even vehicles. Usually status information generated by the sensor nodes is sent to the fog nodes. Only aggregated information or critical states will be forwarded to the central cloud. For example, whether a particular parking spot is occupied is kept at the fog node, while the total number of free spaces may be forwarded from the fog node to the central cloud server.

For large parking garages, dedicated servers will be the first choice as the fog node, because their processing units are more powerful and storage space is larger. Similarly, when there is an RSU next to a parking lot, we will prefer using the RSU as a fog node to take care of processing requirements. These two kinds of fog nodes are easy to set up and we will not focus on them in our following discussion. However, for parking lots that do have dedicated servers available or are not covered by RSUs, or they are out of service due to hardware or software failure, we propose to utilize the capability of Vehicular Fog Computing (VFC) [11], [15], [16] to set up vehicle-based fog nodes to handle the requests for these lots. The VFC is especially useful for handling the cases such as curbside parking and small parking areas.

We have to address several issues. First, we need to select a vehicle as a fog node. Generally we prefer to select a parked vehicle because it will stay in place for a longer period of time. Second, We need to enable vehicles and sensor devices to communicate with the selected fog node. The fog node maintains computing, storage and communication capabilities to handle users requests. Third, we need to connect the fog node with the central cloud so that it can be authorized to make decisions and be trusted by sensor nodes and other vehicles. Also we need to handle the transition to other vehicles once the current fog node needs to leave the location. The rest of this section will describe our approaches in detail.

B. Fog Cloud Architecture

We considered several architectural candidates for the smart parking application. One possibility is that we can depend on RSUs and dedicated servers on larger parking garage to handle the requests, without any help from the cloud architecture. There are four limitations. First, there are smaller parking lots without dedicated servers or not covered by RSUs. We will not be able to include those parking spaces into the application. Second, if the dedicated servers or the RSUs crashes, the information stored there will be lost. For example, the reservation information will not be able to be recovered and agreements with customers will not be able to be honored. Third, different parking lots will not be able to cooperate with each other, such as recommending alternative parking places if the current lot is full. And finally, deploying RSUs and dedicated servers can be costly.

The second possibility is that we can use a centralized cloud server to handle all the requests from users. There are two potential issues. First, the central cloud becomes the concentration point of all requests. It can take a longer time for a user to get the response than the case in which a request can be processed by a distributed and

close-by server. Second, the state information collected by the sensors will be sent over to the cloud. This can be huge, depending on the scope of parking spaces the central cloud is handling.

Instead, we consider an architecture based on the fog cloud computing. We still have a central cloud service that has a large pool of processing and storage resources that can be allocated on demand to meet the ever-changing requirement of the system. In addition, we have fog nodes distributed in different parts of the system, located at the edge of cloud and close to the parking space they serve. They will handle the request for the spaces they are in charge. The central cloud server also acts as the backup service in case of some critical states at distributed fog nodes are lost. It also has the authority to manage the fog infrastructure by authorizing the setup and teardown of fog nodes. The fog nodes can be RSUs, dedicated servers, or even vehicles. Usually status information generated by the sensor nodes is sent to the fog nodes. Only aggregated information or critical states will be forwarded to the central cloud. For example, whether a particular parking spot is occupied is kept at the fog node, while the total number of free spaces may be forwarded from the fog node to the central cloud server.

For large parking garages, dedicated servers will be the first choice as the fog node, because their processing units are more powerful and storage space is larger. Similarly, when there is an RSU next to a parking lot, we will prefer using the RSU as a fog node to take care of processing requirements. These two kinds of fog nodes are easy to set up and we will not focus on them in our following discussion. However, for parking lots that do have dedicated servers available or are not covered by RSUs, or they are out of service due to hardware or software failure, we propose to utilize the capability of Vehicular Fog Computing (VFC) [11], [15], [16] to set up vehicle-based fog nodes to handle the requests for these lots. The VFC is especially useful for handling the cases such as curbside parking and small parking areas.

We have to address several issues. First, we need to select a vehicle as a fog node. Generally we prefer to select a parked vehicle because it will stay in place for a longer period of time. Second, We need to enable vehicles and sensor devices to communicate with the selected fog node. The fog node maintains computing, storage and communication capabilities to handle users requests. Third, we need to connect the fog node with the central cloud so that it can be authorized to make decisions and be trusted by sensor nodes and other vehicles. Also we need to handle the transition to other vehicles once the current fog node needs to leave the location.

B1. Selecting Fog Nodes: Ultimately the central cloud has the authority to select which vehicle can be the fog node for a particular parking space. Information about potential candidates – parked vehicles – should be obtained by the central cloud to make the decision. If there are no existing fog node available, interested parked vehicles can send to the cloud directly. This may be done by vehicle to vehicle communications until the message reaches an RSU or other fog node handling a different parking area that has a channel to the central cloud. If there is an existing fog node in the area, it can collect the information about parked vehicle and send them to the central cloud so that a new fog can be selected once the current fog node needs to leave the area.

The attributes about parked vehicles that will be sent to the central cloud include:

- 1) Vehicle ID: the identification of the vehicle.
- 2) Current location: the x and y coordinates of the vehicle.
- 3) Storage size, computing and communication capacities: the processing power of the vehicle.
- 4) Parking duration: represents the time that the parked vehicle will stay there.
- 5) Willingness: a vehicle is willing to participate in forming a vehicular cloud. A vehicle sending the information usually has a high level of willingness to serve.

The central cloud maintains a data repository to gather vehicles and the fog cloud information. Based on the data collected regarding the vehicle's information, the central cloud uses some criteria to select the fog node as the initial setup for a parking space or as the replacement for the current fog node that is going to leave the area. First, the central cloud have a set of thresholds and only those nodes with metrics equal to or greater than the thresholds can become candidates. We give a list of criteria in the following.

- 1) Computing, Storage and Communication Capacities. Computing, storage and communication capacities vary from vehicle to vehicle. The central cloud specifies a threshold and filter out all those vehicles that do not meet the threshold requirement.
- 2) Duration of Being Active. Another factor to consider is how long the vehicle can provide service. The central cloud will balance the capacities and time being able to serve.
- 3) Behavior Rate. In addition to the vehicle's willingness to server, the central cloud also maintains information about the past behavior of vehicles. Vehicles with high ratings will be preferred. The central cloud performs an adaptive adjustment to rate the vehicle node's behavior based on its previous monitoring of the participating vehicles. Each

time a vehicle participates as a fog node, the cloud monitors the fog nodes and rates them based on how long they serve and performance obtained by other vehicles.

B2. Communications with Fog Nodes: After a static vehicle is selected the fog node, it will broadcast an *advertisement* message to inform other nodes in the same Geo-area, including sensor devices and other vehicles.

The following naming convention is used for the advertisement message: "/advertisement/Fog_nodeID/current position/type/duration". In this format, "advertisement" indicates the type of message, which is a broadcast message; "Fog_nodeID" represents the identification of the selected fog node; "current position" indicates the current position of the fog node at the time; "type" indicates whether the fog node is an RSU, a dedicated server or a parked vehicles; and "duration" shows the duration in which the fog node will be active.

After receiving this message, other nodes will know information about the fog node, including Fog_nodeID, current position, type, and duration. The device (e.g., vehicle or sensor) can include the interest name of the advertisement message in its PIT table. Intermediate nodes in the NDN architecture also record it as a pending interest with the incoming interface.

When the PIT entry is set up in the PIT table of a vehicle or a sensor node, it can push any sensed data messages or send any other requests regarding the smart parking applications to the fog node. The difference is that the device can periodically send updated state information, not just one message to meet the requirement of the interest packet. The device can identify the message as push-based status information as being associated with the fog node using its "Fog_nodeID". To support the push-based model, we need to modify the PIT table entry at intermediate nodes. Instead of removing the entry after one message meeting the interest is forwarded via the incoming interfaces, it will keep the entry for the duration specified by the during in the advertisement message sent by the fog node.

C. Finding Parking Lots with Available Parking Space

The first function we want to implement for the smart parking application is enable drivers to find parking lots with available parking space. We assume that drivers send a query to check parking space availability for a specific area. When these interest packets arrive in the target area, all those parking lots with available parking space will reply.

In order to realize the function, we assume that sensors (IoT devices) are installed to monitor the parking space in each parking lot. We assume that there are n number of

parking spots denoted by $P = \{p_1, p_2, \dots, p_n\}$. The real-time information of a parking spot can be collected by a parking sensor, and it will be pushed to the associated fog node. The fog node is responsible for collecting parking space information regarding occupancy and availability. It maintains a data structure, which is composed of a list of three-tuple: (parking Id, occupancy status, vehicleId).

- 1) parking Id (i): each parking spot is identified by a unique parking Id and is associated with the corresponding position (x_i, y_i) on the Euclidean plane.
- 2) occupancy status (OS): it indicates the status of the monitored parking space. It could be vacant or occupied, which can be represented in 1

$$OS(i) = \begin{cases} 1 & \text{if a vehicle is parked at parking spot } p_i \\ 0 & \text{if the parking spot } p_i \text{ is vacant} \end{cases} \quad (1)$$

- 3) vehicle Id: represents the vehicle's identification that occupies a parking space. This field is only used when the occupancy status value is 1; otherwise, this field will be empty.

To answer the request from the user whether there are available parking spaces, we can simply find the value of $\exists i : OS(i) = 1$. This is only true if we do not accept reservations. In case the fog node allows drivers to reserve parking space ahead of time, we will need a more complicated data structure, which will be described in the next section, to answer the question of whether there is a parking space available.

The naming format of the interest message for requesting whether there are available parking space can be: "application_type/Application_name/data name/timestamp/Geo-area". In this format, "Application_type" should be type A, indicating that it is an application interested in location-associated information. "Application_name" indicates it is a request for parking information, i.e., it is smart parking application; "Data name" represents that the request is specifically requesting for parking lots with available parking spaces. "Timestamp" specifies the time the driver is interested in the availability information. Normally it is the current time for requesting current availability. For the fog nodes that allow reservation, the query can also ask whether the parking lot has space available in some future time. This requires more complicated processing but can be accommodated with the reservation function discussed in the next section.

When a vehicle needs to look for a place to park, it can send a request as an interest packet for finding parking lots with available parking spaces. This request includes the geo-area information indicating the area the drive is interested in. This interest packet can be forwarded by

other intermediate nodes using the algorithm proposed in the previous chapters to the desired target geographical area. The vehicles can use the V2V communication or take advantage of whatever infrastructure is available.

Upon receiving an interest message, an intermediate node will check its CS for matching content. If content with a matching timestamp is found, it will forward the corresponding data back to the vehicle making the request. Otherwise, it will check its PIT table for the same interest name. If the entry with the same interest name exists, the incoming interface of the interest will be added to the corresponding interfaces field in the PIT table. The interest entry in the PIT table will be removed if the entry timer expires.

When a fog node is in charge of a parking space in the target geo-area, it will process the interest packet and send back the data packet with related information. This data packet will be forwarded back to the vehicle making the request.

NDN fits perfectly with this function of the smart parking application because multiple drivers can send interest messages requesting the same content about available parking space. These requests will be aggregated along the way, except the fusion point will have multiple incoming interfaces, which will be used to forward data packets back to these drivers. All drivers can be satisfied with the same corresponding data packet from either the fog node or any caches at the intermediate nodes along the way.

D. Reserving a Parking Space

The other function of the smart parking application provide is to allow a driver to reserve a parking spot before his/her arrival from any location he/she currently is. The reservation will be based on the First Comes First Serve (FCFS) basis.

For the reservation application, we use the fog node to maintain the data related to parking spaces located at the same location and process the corresponding transactions locally. Also the fog node will use the central cloud as the backup service to maintain crucial states including existing user reservations. The central cloud also play a role of managing the setup of a new fog node and transition old states to it.

Multiple drivers may make reservations from the same parking lot with the same fog node. Each requester should be satisfied differently with the unique data content related to the request. NDN can satisfy multiple interest packets that carry the same name by fetching the original data from the producer or any intermediate node's corresponding data content. In this case, the data content will be the same for all the consumers. However, the case for reservation in the smart parking application is different in the sense that each requester should be

satisfied differently, even though their interest packet may have the same name.

To deal with the issue, we need to make some modifications to NDN packets by extending the interest and data packet fields. The following fields are added to the interest packet: [my_id, my_arrival-time, my_departure-time, parkingSpace_location, my_position], where “my_id” indicates the identification of the sender vehicle, “my_arrival-time” is the earliest possible arrival time, “my_departure-time” is the latest possible departure time, “parkingSpace_location” specifies the parking lot the driver wants to make the reservation, and “my_position” indicates the sender vehicle position.

When an intermediate node receives the interest packet regarding the reservation, the interest name will be included in the PIT table if there is no matched data in the CS. The current NDN forwarding daemon has no concept about sending the same interest name and satisfying each consumer with different data packets. We introduce the concept of Identity of Vehicle Interfaces (IVI) by modifying the PIT table to include the IVI, instead of the incoming interface field. The IVI entries include the binding of the incoming interface to the vehicle ID of the requester and a version number. This binding can be presented as <incoming interface, vehicleID, version>, where “version” is used to indicate the version number when a requester vehicle updates its route and sends a new version of the same interest. This binding helps to send the data packet to each requester vehicle separately. Once the intermediate node receives the data packet regarding a reservation request, the PIT will only satisfy the request with the same vehicleID. Only the IVI with the same vehicleID will be deleted. All other IVIs will remain there until the data packets matching their ID arrive. The PIT entry will be deleted if the data packet fails to return by the expiration time, or if the requests from all vehicles are satisfied.

When the reservation request arrives at the fog node responsible for a parking space, it extracts three tuple < vid, r_a, r_d > from the request. vid is the id of the driver, r_a is the arrival time the driver expects to arrive to start parking the vehicle in the parking lot, and r_d is the expected departure time.

We assume that there are n parking spots in the parking space, denoted by $P = \{p_1, p_2, \dots, p_n\}$, as shown in Figure 2. The system will allow drivers to reserve for a certain period of time (e.g., a week, a month, etc.). The y-axis represents the n parking spots in the parking area. The x-axis indicates the time. Any reservation (as marked) is an interval on the x-axis.

The data structure used to represent the reservations is shown in Figure 3. Any parking spot p_i consists of a list of reservations. Each reservation is represented by a node < vid, a, d >, where vid is the identification of

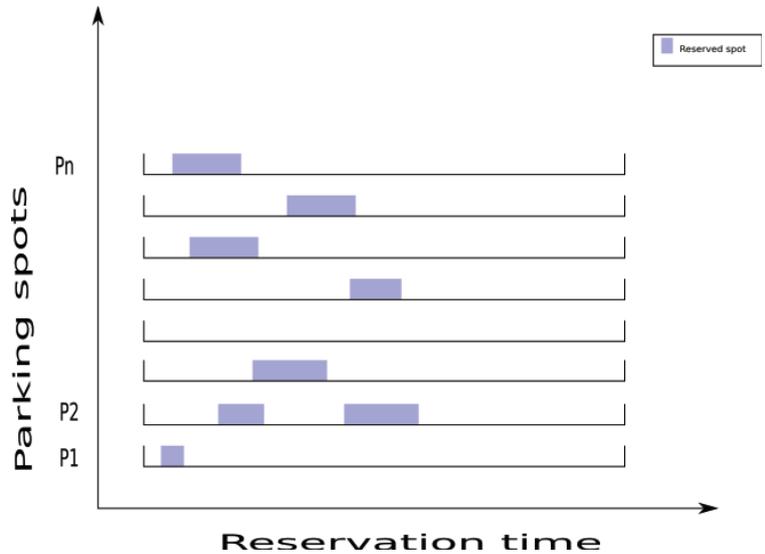


Fig. 2. Schedules for the Reservation System

the vehicle this spot is reserved for, a is the arrival time, and d is the departure time. $\langle a, d \rangle$ represented an interval of the reservations. There is no overlap between all reservations for any given spot. To make the search algorithm for finding a spot that is available for the duration of a new request, we make sure that the list is sorted in increasing order based on the arrival time. If a spot has no reservation, the list is empty, and the pointer from that spot is NULL.

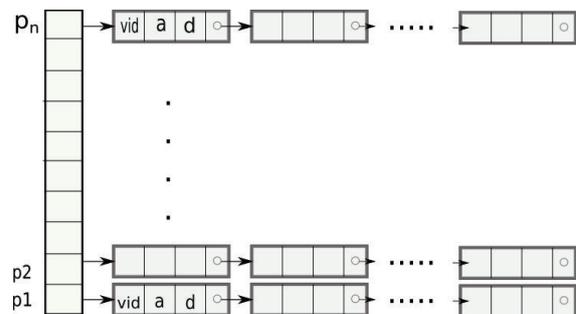


Fig. 3. Data structure for the reservation application

The system searches for a parking spot that is not reserved for the period of time $\langle r_a, r_d \rangle$. If the system finds such an available interval, it reserves the parking spot for the requested duration, stores the reservation information and confirms the reservation. The reservation is in the format of $\langle vid, r_a, r_d, spot \rangle$, where $spot$ is the parking spot reserved for a vehicle identified by vid . If no such spot is found, the system will report failure. Figure 4 shows the detail of the algorithm for reserving a parking spot. It goes through all the parking

spots. If it finds a spot that has not been reserved by any vehicle, it will reserve the time interval (lines 2 to 5). Otherwise, it searches through the ordered list of current reservations. If the departure time is less than the arrival time of the first reservation for a node, the reservation will be added to the beginning of the list (lines 6-12). Otherwise, if it can find in the list two reservations such that the requested arrival time is after the first reservation's departure time and the requested departure time is before the second reservation's arrival time, the requested reservation will be put in between (lines 13 to 25). If we still cannot find a parking spot that is available for the duration of the requested reservation after going through all the parking spots, it returns failure.

Data: the list of Parking Spots $P[1..n]$, the id of the vehicle making the request vid , request arrival time r_a , request departure time r_d

```

1 for  $i = 1$  to  $n$  do
2   if  $P[i] == nil$  then
3      $nptr = new\ node(vid, r_a, r_d)$ ;
4      $P[i] = nptr$ ;
5     return(success,  $i$ );
6   else
7      $cptr = P[i]$ ;
8     if  $r_d < cptr.a$  then
9        $nptr = new\ node(vid, r_a, r_d)$ ;
10       $nptr.next = cptr$ ;
11       $P[i] = nptr$ ;
12      return(success,  $i$ );
13    else
14      while  $((cptr \neq nil) \text{ and } (r_a \geq cptr.d))$  do
15         $qptr = cptr.next$ ;
16        if
17           $((qptr == nil) \text{ or } (r_d \leq qptr.a))$ 
18          then
19             $nptr = new\ node(vid, r_a, r_d)$ ;
20             $nptr.next = qptr$ ;
21             $cptr.next = nptr$ ;
22            return(success,  $i$ );
23          else
24             $cptr = qptr$ ;
25        end
26      end
27    end
28 return(failure);

```

Fig. 4. Algorithm for Making a Reservation

Drivers can also cancel a previous reservation by sending the tuple $\langle vid, spot, r_a, r_d \rangle$ to the fog node

in charge of the parking space. Figure 5 shows the algorithm for canceling a reservation. It first checks whether the parking spot number is valid or not (lines 1 to 3). If not, it will return with a failure. Otherwise, it searches the list of reservations for that spot (lines 4 to 18). If it finds a reservation that matches the user provided, it will delete the reservation from the list and return with success. If it cannot find the matching reservation, it will return with a failure (line 19).

Data: the list of Parking Spots $P[1..n]$, the id of the vehicle making the deletion request vid , the reservation $spot, r_a, r_d$, where $spot$ is the parking spot reserved, r_a is arrival time, r_d is the departure time

```

1 if  $((spot < 1) \text{ or } (spot > n))$  then
2   return("Failure: Invalid spot number");
3 end
4  $cptr = P[spot]$ ;
5  $pptr = nil$ ;
6 while  $(cptr)$  do
7   if  $((cptr.v == vid) \text{ and } (cptr.a == r_a) \text{ and } (cptr.d == r_d))$  then
8     if  $(pptr == nil)$  then
9        $P[spot] = cptr.next$ ;
10    else
11       $pptr.next = cptr.next$ ;
12    end
13    return("Success: Reservation deleted");
14  else
15     $pptr = cptr$ ;
16     $cptr = cptr.next$ ;
17  end
18 end
19 return("Failure: Reservation not found");

```

Fig. 5. Algorithm for Cancelling a Reservation

IV. PERFORMANCE EVALUATION

A. Simulation Setup

For the simulations, we used ndnSIM (Version 2.8) [17], which is an NS-3 based NDN simulator. We also used SUMO [18], a microscopic traffic simulation for urban mobility, to generate the traffic. We created a map of a part of the University Campus. The map size is 2000×2000 meters. We generated different parking places distributed over the map. The vehicle speed limit is set to 20 km/h, and we set the transmission range of the signal of all vehicles to 150 meters. We utilized the ndnSIM parking applications as a starting point and implemented our smart parking application. The parameters for the experimental setup are shown

in Table I. We assume that at each parking area, some parking spaces are occupied, and others are not.

TABLE I
SIMULATION STEP PARAMETERS

Maximum transmission range	150 m
Maximum transmission range of RSU	400 m
Number of vehicles	20, 30, ..., 100, 150
Vehicle speed	20 km/h
Simulation duration	200 ms
MAC type	IEEE802.11

B. Simulation Results

We evaluated the performance of an NDN-based smart parking application. Our evaluation consists of two parts. One is the evaluation of the function for finding parking lots with available parking spaces. The other is the evaluation of the parking reservation function of the system.

B1. Finding the Availability of Parking Space: We compared our approach with the original NDN system. There are two main differences. One is how sensor nodes send the data to the fog node in charge of the parking space. In the original NDN, the sensor data will be sent to the fog node only if an interest message is broadcast from the fog node to the network. In our push-based model, after initial advertisement from the fog node, the sensor nodes can periodically push the sensed data to the fog node without the need to receive further interest packets.

The other difference is how the vehicles make requests to find whether there are parking spaces available. In the original NDN, vehicles do not know any information about the fog node (an RSU or a selected parked vehicle). They will broadcast an interest message to request the availability of parking spaces. In our approach, vehicles are making the request to a specific geographical area and only those fog nodes in the geographical area may respond to the availability request.

We evaluate the performance using two metrics, response time and push overhead.

1) Response time. It is defined as the time interval from the time when the vehicle makes the request by sending an interest packet to the time when it receives the response (data packet). The number of vehicles varies from 20 to 150.

Figure 6 shows the response time when the number of vehicles varies from 20 to 150. We observed that the response time decreases when the number of vehicles increases. One reason is the caching effect of NDN because, with more vehicles, we can have a higher hit ratio at the cache at the intermediate nodes. The response time of our approach is less than that of the original NDN. Specifically, it achieves a 16.74% to

20.7% decrease when the number of vehicles increases from 20 to 150.

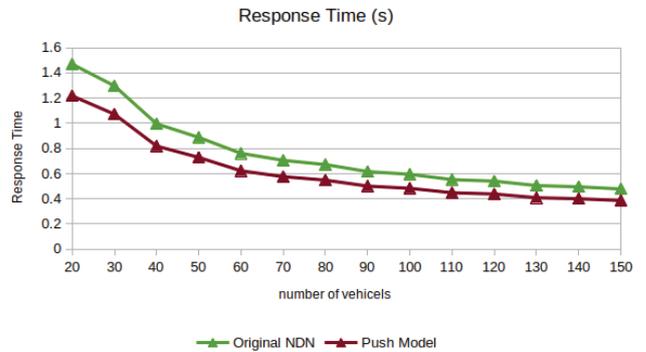


Fig. 6. Response time for different numbers of vehicles

2) Push-based Model Overhead: defined as the total number of periodic *advertisement* messages that are broadcast by the fog node over the total number of pushed data packets to the fog node. Figure 7 shows that the proposed push-based model achieves very low overhead, which decreases gradually as the number of vehicles increases. Establishing the connection by *advertisement* message helps in generating low overhead because we only need a few interest messages to be sent over the network. The push-based model outperforms the original NDN because we set a communication channel between the fog node and sensors to allow sensors to send periodically status information to the fog node.

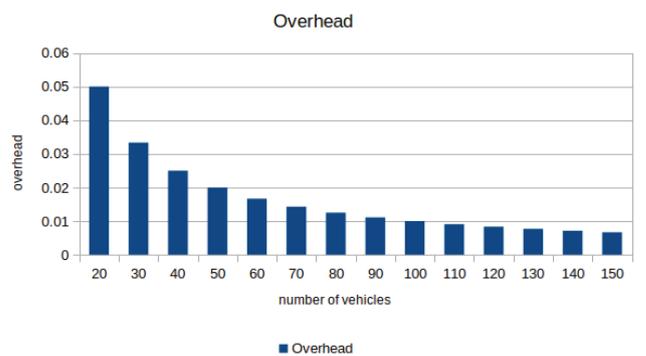


Fig. 7. Overhead caused by the proposed Push-Model

Reserving a Parking Space

We also evaluated the performance of the proposed reservation application by comparing it with a baseline approach, which allocates a parking space for a vehicle when it arrives at the parking lot without any prior reservation. The parameters for the experimental setup are shown in Table II. Figure 8 illustrates the map generated by SUMO. Parking Areas are also identified on the map.

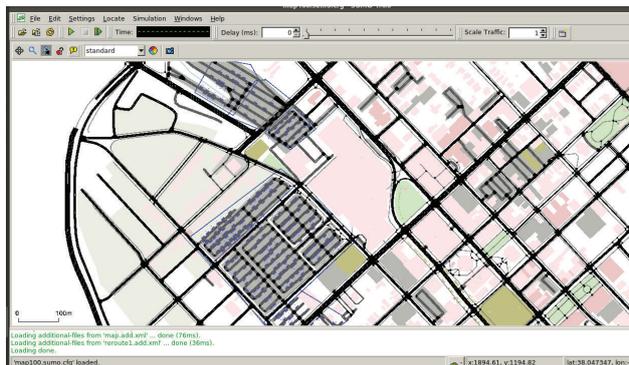


Fig. 8. SUMO map for a smart parking application

TABLE II
EXPERIMENT SETUP FOR PARKING SPACES

Parking slots	Total parking spaces
P1	25
P2	15
P3	30
P4	100
P5	60
P6	30

We evaluate the performance using two metrics, success rate, and the average time to park.

1) Success Rate to Find a Space. It is defined as the ratio of the number of vehicles that successfully find a vacant parking space over the total number of vehicles. Figure 9 illustrates that when the number of vehicles increases from 20 to 150, the success rate to find a space decreases. Our reservation system has a higher success rate compared with the baseline approach. In the baseline approach, the driver just drives to a parking lot, and if the parking lot is full, the task of finding space is considered failed even if the driver may go to other parking lots and find space later. When the number of vehicles is low, the difference between the baseline approach and our reservation system is small. However, the difference becomes more obvious when the number of vehicles increases. We observed the success rates of our approach are 9.53%, 26.98%, and 53.08% higher than the baseline approach when the numbers of vehicles are 70, 110, and 150, respectively. Our approach can do better because it allows drivers to reserve a parking place before the arrival time.

2) Average time needed to park a vehicle. It is defined as the average time required to find a vacant parking space for a vehicle. Figure 10 shows that the number of vehicles has an effect on the average time to park. As the number of vehicles increases, the average time needed to park a vehicle increases. From the figure, we can see that the average time to park for the reservation

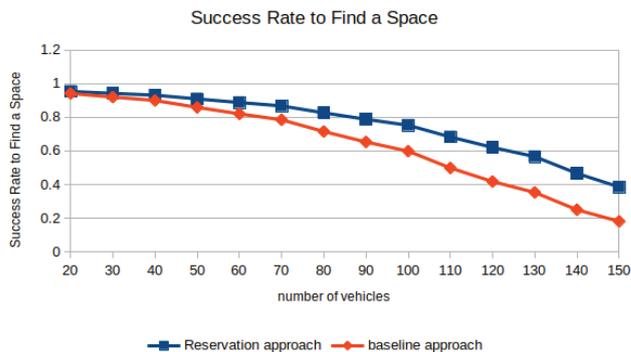


Fig. 9. Successful Rate of Finding a Space

approach outperforms is smaller than that of the baseline approach. More specifically, they are between 40.36% and 53.31% less than those of the baseline approach.

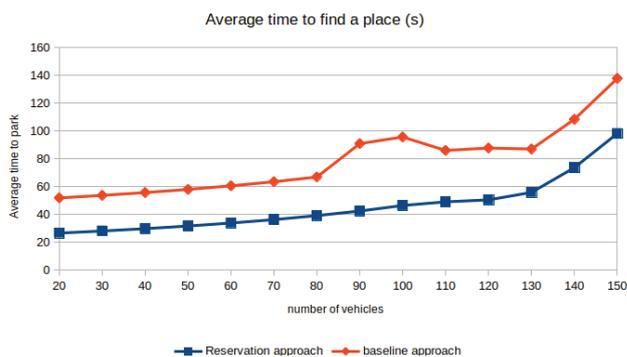


Fig. 10. Average time needed to park

CONCLUSION

We categorized different types of vehicular named data networking applications and provided an NDN naming scheme for each type. We considered a priority method to provide better service for some time-sensitive applications. We developed an NDN-based framework for the smart parking application. A fog cloud architecture is adopted so that sensor data can be collected locally and requests can be handled by distributed fog nodes to reduce the load on a central cloud server. The cloud aspect allows us to dynamically set up a new fog node from static vehicle to provide service for those parking spaces without dedicated servers or RSUs. The smart parking application enables drivers to query about the parking lots that have available parking spaces and make a reservation for a parking space for an interval in the future.

REFERENCES

- [1] Trista Lin, Herve Rivano, and Frederic Le Mouel. A survey of smart parking solutions. *IEEE transactions on intelligent transportation systems*, 18(12):3229–3253, 2017.
- [2] X. Wang and S. Cai. An efficient named-data-networking-based iot cloud framework. *IEEE Internet of Things Journal*, 7(4):3453–3461, 2020.
- [3] Khaoula Hassoune, Wafaa Dachry, Fouad Moutaouakkil, and Hicham Medromi. Smart parking systems: A survey. In *2016 11th International Conference on Intelligent Systems: Theories and Applications (SITA)*, 2016 11th International Conference on Intelligent Systems: Theories and Applications (SITA), pages 1–6. IEEE, 2016.
- [4] IEEE Electrical Insulation Society Staff Corporate Author. Smart parking solutions for urban areas. In *2013 IEEE 14th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2013 IEEE 14th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM), pages 1–6, [Place of publication not identified], 2013. IEEE.
- [5] Faheem, S.A. Mahmud, G.M. Khan, M. Rahman, and H. Zafar. A survey of intelligent car parking system. *Journal of Applied Research and Technology*, 11(5):714 – 726, 2013.
- [6] Thierry Delot, Nicolas Cenerario, Sergio Harri, and Sylvain Lecomte. A cooperative reservation protocol for parking spaces in vehicular ad hoc networks. 01 2009.
- [7] N Doulamis, E Protopapadakis, and L Lambrinos. Improving service quality for parking lot users using intelligent parking reservation policies. pages 1392–1397. IEEE, 2013.
- [8] ElMouatezbillah Karbab, Djamel Djenouri, Sahar Boulkaboul, and Antoine Bagula. Car park management with networked wireless sensors and active rfid. In *Electro/Information Technology (EIT), 2015 IEEE International Conference on, 2015 IEEE International Conference on Electro/Information Technology (EIT)*, pages 373–378. IEEE, 2015.
- [9] Redowan Mahmud, Ramamohanarao Kotagiri, and Rajkumar Buyya. *Fog Computing: A Taxonomy, Survey and Future Directions*, pages 103–130. Springer Singapore, Singapore, 2018.
- [10] W. Balzano and F. Vitale. Dig-park: A smart parking availability searching method using v2v/v2i and dgp-class problem. In *2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, pages 698–703, 2017.
- [11] Xueshi Hou, Yong Li, Min Chen, Di Wu, Depeng Jin, and Sheng Chen. Vehicular fog computing: A viewpoint of vehicles as the infrastructures. *IEEE transactions on vehicular technology*, 65(6):3860–3873, 2016.
- [12] Y. Xiao and Chao Zhu. Vehicular fog computing: Vision and challenges. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 6–9, 2017.
- [13] Chaogang Tang, Xianglin Wei, Chunsheng Zhu, Wei Chen, and Joel J. P. C Rodrigues. Towards smart parking based on fog computing. *IEEE access*, 6:70172–70185, 2018.
- [14] Yi Zhang, Chih-Yu Wang, and Hung-Yu Wei. Parking reservation auction for parked vehicle assistance in vehicular fog computing. *IEEE transactions on vehicular technology*, 68(4):3126–3139, 2019.
- [15] Flavio Bonomi, Rodolfo Milito, Jiang Zhu, and Sateesh Addepalli. Fog computing and its role in the internet of things. In *Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing*, MCC '12, page 13–16, New York, NY, USA, 2012. Association for Computing Machinery.
- [16] Ivan Stojmenovic, Sheng Wen, Xinyi Huang, and Hao Luan. An overview of fog computing and its security issues. *Concurrency and Computation: Practice and Experience*, 28(10):2991–3005, 2016.
- [17] I. Moiseenko S. Mastorakis, A. Afanasyev and L. Zhang. ndnSIM 2: An updated NDN simulator for NS-3, ndn, technical report ndn-0028, revision 2, 2016, 2016.
- [18] Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Laura Bieker. Recent development and applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements*, 5(3&4):128–138, December 2012.

A Dynamic Model for Circularity Assessment of Nutrient Recovery from Domestic Sewage

Anurag Bhambhani, Jan Peter van der Hoek, Zoran Kapelan

Abstract—The food system depends on availability of Phosphorus (P) and Nitrogen (N). Growing population, depleting Phosphorus reserves and energy intensive industrial nitrogen fixation are threats to their future availability. Recovering P and N from domestic sewage water offers a solution.

Recovered P and N can be applied to agricultural land replacing virgin P and N. Thus, recovery from sewage water offers a solution befitting a circular economy. To ensure minimum waste and maximum resource efficiency a circularity assessment method is crucial to optimize nutrient flows and minimize losses.

Material Circularity Indicator (MCI) is a useful method to quantify circularity of materials. It was developed for materials that remain within the market and recently extended to include biotic materials that may be composted or used for energy recovery after end-of-use. However, MCI has not been used in the context of nutrient recovery. Besides, MCI is time-static i.e. it cannot account for dynamic systems such as the terrestrial nutrient cycles.

Nutrient application to agricultural land is a highly dynamic process wherein flows and stocks change with time. The rate of recycling of nutrients in nature can depend on numerous factors such as prevailing soil conditions, local hydrology, presence of animals, etc. Therefore, a dynamic model of nutrient flows with indicators is needed for the circularity assessment.

A simple substance flow model of P and N will be developed with the help of flow equations and transfer coefficients that incorporates the nutrient recovery step along with agricultural application, the volatilization and leaching processes, plant uptake and subsequent animal and human uptake. The model is then used for calculating the proportions of linear and restorative flows (coming from reused/recycled sources). The model will simulate the adsorption process based on quantity of adsorbent and nutrient concentration in the water. Thereafter, application of the adsorbed nutrients to agricultural land will be simulated based on adsorbate release kinetics, local soil conditions, hydrology, vegetation, etc. Based on the model, the restorative nutrient flow (returning to the sewage plant following human consumption) will be calculated.

The developed methodology will be applied on a case study of resource recovery from waste water. In the aforementioned case study located in Italy, bio char or zeolite is to be used for recovery of P and N from domestic sewage through adsorption and thereafter, used as a slow-release fertilizer in agriculture.

Using this model, information regarding efficiency of nutrient recovery and application can be generated. This can help to optimize the recovery process and application of the nutrients. Consequently this will help to optimize nutrient recovery and application and reduce dependence of the food system on virgin extraction of P and N.

Keywords—Circular economy, dynamic substance flow, nutrient cycles, resource recovery from water.

Anurag Bhambhani, Technische Universiteit Delft, Netherlands (phone: +31-0620341381); e-mail: a.bhambhani@tudelft.nl).

Jan Peter van der Hoek, Technische Universiteit Delft, Netherlands; e-mail: j.p.vanderHoek@tudelft.nl).

Zoran Kapelan, Technische Universiteit Delft, Netherlands; e-mail: z.kapelan@tudelft.nl).

Floating Populations, Rooted Networks. Tracing the Evolution of Russeifa City in Relation to Marka Refugee Camp

Dina Dahood Dabash

Abstract— Refugee camps are habitually defined as receptive sites, transient spaces of exile and nondescript depoliticized places of exception. However, such arguments form partial sides of reality, especially in countries that are geopolitically challenged and rely immensely on international aid.

In Jordan, the dynamics brought with the floating population of refugees (Palestinian amongst others) have resulted in spatial after-effects that cannot be easily overlooked. For instance, Palestine refugee camps have turned by time into socioeconomic centers of gravity and cores of spatial evolution. Yet, such a position is not instantaneous. Amongst various reasons, it can be related, according to this paper, to a distinctive institutional climate that has been co-produced, by the refugees, host community and the state.

This paper aims to investigate the evolution of urban and spatial regulations in Jordan between 1948 and 1995, more specifically, state regulations, community regulations and refugee-self-regulation that all dynamically interacted that period. The paper aims to unpack the relations between refugee camps and their environs to further explore the agency of such floating populations in establishing rooted networks that extended the time and place boundaries. The paper's argument stems from the fact that the spatial configuration of urban systems is not only an outcome of a historical evolutionary process, but is also a result of interactions between the actors.

The research operationalizes Marka camp in Jordan as a case study. Marka Camp is one of the six "emergency" camps erected in 1968 to shelter 15,000 Palestine refugees and displaced persons who left the West Bank and Gaza Strip. Nowadays, camp shelters more than 50,000 refugees in the same area of land. The camp is located in Russeifa, a city in Zarqa Governorate in Jordan. Together with Amman and Zarqa, Russeifa is part of a larger metropolitan area that acts as a home to more than half of Jordan's businesses.

The paper aspires to further understand the post-conflict strategies which were historically applied in Jordan and can be employed to handle more recent geopolitical challenges such as the Syrian refugee crisis.

Methodological framework: The paper traces the evolution of the refugee-camp regulating norms in Jordan, parallel with the horizontal and vertical evolution of Marka camp and its surroundings. Consequently, the main methods employed are historical and mental tracing, Interviews, in addition to using available Aerial and archival photos of the Marka camp and its surrounding.

Keywords— Forced migration, Palestine Refugee camps, Spatial agency, Urban regulations

Dina Dahood Dabash is with Department of Architecture, KU Leuven University, Kasteelpark Arenberg 1 - box 2431, 3001 Leuven, Belgium (phone: 962-79-8485117; e-mail: archdina@gmail.com; dina.dabash@kuleuven.be).

To Improve or Not to Improve, Reflections from Jerash Urban Improvement Project, Jordan

Dina Dahood Dabash

Abstract— Palestine Refugee Camps have never been settings that can be overlooked; they even became (as physical settings) a cornerstone topic of negotiations whenever Palestinian matters are on the table (specifically in Jordan). Consequently, maintaining the familiar face of the camp with its dilapidated shelters and narrow streets, that rarely allowed its residents to extinguish a fire or evacuate a building safely, has become a fundamental method to protect the “right of the return” from the perspective of various stakeholders. When the Infrastructure and Camp Improvement Programme (ICIP) was established in 2007 as an additional UNRWA programme, some concerns were raised around the newly established section, mainly due to its direct impact on the “image” of the camp through a provision of a relatively nonconventional service that differs from what the Agency used to provide in the past seventy years.

By presenting the Urban Improvement Project in Jerash camp (UIP) -Jordan, this paper aims to contribute to the ongoing discussion around enduring the improvement of Palestine refugee camps “programmatically” in UNRWA or not. The UIP as a co-product by UNRWA and the camp’s community within one of the most vulnerable refugee camps in Jordan offers a remarkable opportunity to excerpt lessons that can contribute to the strategic shaping of the ICIP.

The paper concludes with five main uptakes mainly related to community engagement, power structures and UNRWA role in camps.

Keywords—Camp Improvement Programme, Jerash Camp, Palestine Refugee Camps, UNRWA.

Legal Issues Relating to Sustainable Buildings

Richard J. Sobelsohn

Abstract— Too often Sustainability efforts, although designed with the best intentions, neglect to focus on the legal implications of those designs. This can include without limitation, the myriad of contracts between the parties, compliance with building codes and other laws and regulations, and the application for certification and/or grants/loans. This paper will discuss some of these issues and how to obtain the desired goals, without sacrificing legal rights and remedies.

Keywords— Contracts, Risk, Remedies, Benefits.

Does Index-Based Livestock Insurance Change the Cash Saving Behaviour of Pastoralists in Southern Ethiopia?

Tnsue Gebrekidan Bezabh

Abstract— Because of climate change, agriculture in developing countries has become a riskier business. When weather shocks strike, in the absence of formal agricultural insurance markets, farmers often tap into their precautionary savings, which is the quickest form of ex-post self-insurance to buffer consumption and/or restock herds. However, it may not be enough to cope with systematic shocks. Recently, index-based livestock insurance is getting growing attention to deal with systematic agricultural risks. In this paper, we employ descriptive and multiple regressions to estimate the causal effects of index-based livestock insurance on saving behavior of the Borena pastoralists, which is less studied and has remained unclear yet, using three round household level panel data. Our estimations result show that cash saving behavior of the pastoral households is significantly and positively influenced by the insurance uptake. This finding suggests that the insurance can increase the likelihood of cash saving as it can enable insured households to off-take their herds at speculated market prices. Likewise, since the insured have risk-averting behavior and the insurance does not provide coverage against losses due to idiosyncratic risks, they may tend to save precautionary cash.

Keywords— index-based livestock insurance, saving, Borena.

Analyze and Visualize Eye-Tracking Data

Aymen Sekhri
aymensek@uef.fi

Emmanuel Kwabena
Frimpong
emmafrim@uef.fi

Bolaji Mubarak
Ayeyemi
mayeyemi@uef.fi

Aleksi Hirvonen
alekshi@uef.fi

Matias Hirvonen
matiasi@uef.fi

Tedros Tesfay
Andemichael
tedandem@uef.fi

ABSTRACT

Fixation identification, which involves isolating and identifying fixations and saccades in eye-tracking protocols, is an important aspect of eye-movement data processing that can have a big impact on higher-level analyses. However, fixation identification techniques are frequently discussed informally and rarely compared in any meaningful way. With two state-of-the-art algorithms, we will implement fixation detection and analysis in this work. The velocity threshold fixation algorithm is the first algorithm, and it identifies fixation based on a threshold value. For eye movement detection, the second approach is U'n' Eye, a deep neural network algorithm. The goal of this project is to analyze and visualize eye-tracking data from an eye gaze dataset that has been provided. The data was collected in a scenario in which individuals were shown photos and asked whether or not they recognized them. The results of the two-fixation detection approach are contrasted and visualized in this paper.

1. Introduction

Over the last decade, eye tracking has become increasingly popular as a doorway into observers' visual and cognitive processes. Eye-tracking has been used to examine behavior in areas including image scanning [1], arithmetic[2], driving[3], analogies, and reading. Fixations (pauses over informative regions of interest) and saccades are commonly used by researchers in these and other disciplines to describe eye movements (rapid movements between fixations). Natural settings, simple artificial stimuli, webpages, user interfaces, and, increasingly, information visualizations have all been studied using eye movement analysis. Eye-tracking has been widely utilized in human-computer interaction (HCI) to assess system usability and investigate the related subject of interface design [4], [5]. Duchowski [4] presents an overview of many eye-tracking applications in fields ranging from industrial engineering to marketing.

Fixation or gaze durations, saccadic velocities, saccadic amplitudes, and various transition-based parameters between fixations and/or regions of interest are common analysis metrics.

Fixation identification (or simply identification) is the process of translating raw eye-movement data points into fixation positions (and implicitly the saccades between them) on the visual presentation required for fixation and saccade research. Fixation identification considerably decreases the size and complexity of the eye-movement procedure by removing raw saccade data points and combining raw fixation points into a single representative tuple. This lowering is beneficial for at least two reasons. To begin with, a saccade may achieve little or no visual

processing [6], therefore the actual pathways used during saccades are frequently irrelevant for many research applications. Second, during fixations, tiny eye movements such as tremors, drifts, and flicks [7] are usually neglected in the higher-level analysis. As a result, fixation identification provides a simple way to reduce the complexity of eye-tracking data while keeping the most important qualities for studying cognitive and visual processing behavior.

A statistically-based description of observed eye movements is known as fixation identification. While it is commonly understood that fixations involve visual and cognitive processing [8], it is less clear when fixations begin and end. As a result, the identification problem remains a subjective process, despite the precision and flexibility of identification algorithms. As a result, comparing the created fixations to the subjective impressions of an observer is an excellent method of validating these algorithms.

In the eye-movement literature, notably in the research on the relationship between eye movements and higher-level cognition, fixation identification and its various implementations have gotten a lot of attention. Identification, on the other hand, is frequently a necessary aspect of eye-movement data processing, with significant consequences for later studies. A fixation is a collection of points with a distance between them less than a predetermined value and a temporal interval more than a predetermined period. It's been thought of as a collection of places where a subject's gaze has been concentrated[9]. In the literature, the minimum time and dispersion distance to define a fixation have been hotly contested. It has long been assumed that the minimum fixation time is more than 0.1 seconds [10]. The work the subject is doing determines the subject's minimum time. For tasks such as reading and visual search, the minimum fixation period is 0.225 seconds and 0.275 seconds, respectively. The average fixation duration for eye-hand coordination tasks was found to be 0.4 seconds [9]. In conclusion, the mean fixation period has been calculated to be 0.15 to 0.65 seconds [11]. Fixation dispersion angles have yet to be determined, but they are usually adjusted to less than 2° [9]. Apart from the widespread under-specification of identification algorithms, there has been little effort put into assessing and comparing alternative possibilities. In many circumstances, academics and practitioners lack sufficient information to determine which algorithm to apply in a given situation. This issue, expected, leads to the haphazard use of multiple methods, making it difficult to compare results acquired using various identification strategies.

In this paper, we used velocity-threshold identification (I-VT) and U'n' Eye, deep neural network algorithms to identify fixation on

the provided dataset. The result of both methods is compared, and we provide a review of metrics that can be used for quantitative comparison. Moreover, we discuss and visualize ways in which different metrics can be used to present the algorithm results and determine the effectiveness of the algorithms.

2. Method

2.1 Dataset

The data was collected in a circumstance where individuals have presented photos and asked to identify whether they recognized them. The dataset is made up of rows of raw gaze positions x,y recorded in time (one row per session and person). In the dataset, subjects have several rows/samples. The dataset is available in two formats: CSV (a text file with one row for each sample) and Matlab (a spreadsheet with one row for each sample). The following features are included in it:

- (x, y)
- (x, y, t)
- (x, y, t, v_x, v_y)
- $(x, y, t, v_x, v_y, a_x, a_y)$
- $(x, y, t, v_x, v_y, a_x, a_y, r_x, r_y)$ (raw data)

The values for x and y are 0,0 for the screen's center point, positive for the screen's right and lower side, and negative for the screen's left and upper side. Every row has a different number of values! (Because participants' reactions may take varying amounts of time.) Assume a 1000Hz sampling rate.

2.2 Velocity-Threshold Identification (I-VT)

Velocity-threshold fixation identification is the simplest of the identification processes to comprehend and apply (I-VT). Based on their point-to-point velocities, the I-VT method separates fixation and saccade locations[12]. Fixation velocity algorithms combine a duration criterion and an eye velocity-based stillness condition. Due to micro-movements in the eye and noise from the eye-tracker, the eye velocity is rarely at absolute zero. As a result, users of this technology must decide on a fixation velocity upper limit[13]. Fixation length is the average amount of time between fixations. Typically, the fixation duration is between 150 and 300 milliseconds [14]. The velocity profiles of saccadic eye movements show low velocities for fixations (i.e., 100 degrees per second) and high velocities for saccades (i.e., >300 degrees per second)[15]. I- The I-VT approach's pseudocode is shown in Table 1. I-VT requires one parameter, the velocity threshold. The point-to-point velocity threshold can be approximated using an acceptable angular velocity threshold [6] if angular velocities can

be determined (i.e., the distance between the eye and the visual stimuli is known).

Table 1: Pseudocode for the I-VT algorithm.

<pre> fixation_velocity(protocol,velocity_threshold,clustering_distance): Calculate point-to-point velocities for each point in the protocol. Label each point below velocity_threshold as a fixation point, otherwise as a saccade point. Collapse consecutive fixation points into fixation groups, removing saccade points. Map each fixation group to fixation at the centroid of its points. Collapse fixation points that the distance between them is below clustering_threshold, and combine them to fixation at the center of centroids. Return fixations, duration of each fixation </pre>
--

I-VT requires one parameter, the velocity threshold. The point-to-point velocity threshold can be approximated using an acceptable angular velocity threshold [6] if angular velocities can be determined (i.e., the distance between the eye and the visual stimuli is known). For example, Sen and Megaw [16] used a 20-degree-per-second barrier. When only point-to-point velocities are available, an appropriate velocity threshold value must be established based on data collection factors (e.g., sample frequency) and exploratory data analysis. We have added one parameter to the basic I-Vt algorithm, ie. Clustering distance. The modification was made to the velocity algorithm to improve its results. The modification is combining any two neighboring fixations points if the distance between them is less than a certain threshold. The reason for this modification is that the original velocity algorithm is sensitive to the velocity, so it gives many fixations (with small durations) in very close points, but clustering these close fixations will result in one fixation with a higher duration.

2.3 U'n'Eye deep neural network

U'n'Eye is a deep neural network model that can be used for the detection of saccades and other eye movements. It was constructed with a convolutional neural network (CNN) to automatically detect saccades at human-level performance accuracy. The application outperforms the present state of the art, according to prominent performance metrics, and will aid researchers in better understanding the neurophysiological [17]. In addition to saccades and microsaccades, the network was processed that governs saccade production and visual processing

able to predict other forms of eye movements, such as blinks and PSOs. U'n'Eye performed well when trained on a single type of data with only one coder's labels and when trained on two datasets with two coders' labels, according to the study. Prediction of eye movement can be possible with the already trained model, but it is recommended to train the network on your dataset to achieve optimal performance. U'n'Eye is free and accessible to the public, with a user-friendly interface and a web API that allows users to upload data and get categorization results. No parameter changes are necessary for training (e.g., learning rate, etc.) because the default parameters have been found to work well across datasets. An experimenter only requires a few hundred seconds of tagged data to train the network.

Predict saccades in recordings with pre-trained weights. Arrange the data into a matrix of *samples*timebins* or input them as a vector of length *timebins*.

$$(X, Y)$$

Input parameters:

X,Y : horizontal and vertical eye positions, {array-like, shape: (nsamples, nbins) or (nbins)}

Output:

Prediction: eye movement class prediction for each time bin, the same shape as input

Probability: softmax probability output of network, shape: (n_samples,classes,time) or (classes,time)}

The weight used for the model is "weights_dataset3" with the sampling frequency of 1kHz

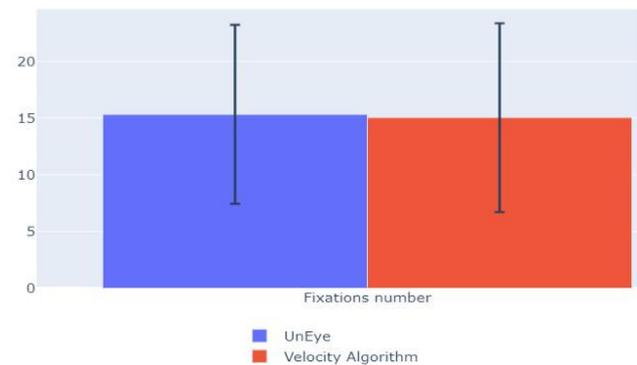
3. Results and Discussion

The results of velocity threshold identification and the U 'n'Eye deep learning model with various metrics that can be utilized for collective eye movement analysis over the dataset are shown and discussed in this part.

Mean fixation duration measures the average fixation time per subject. Considering table 2, the mean fixation duration for all subjects when the person identifies image ranges from 0.104s to 0.185s, and the mean fixation duration for all subjects when a person could not identify image ranges from 0.101s to 0.205s. Moreover, the maximum overall mean fixation duration in all subjects is 0.195s. According to different research results, the fixation duration should be between 150 and 300 milliseconds.

Therefore, subjects whose mean fixation threshold is not in the range of 150 and 300 milliseconds can be considered as involuntary (the eyes move there without a conscious decision). The standard deviation of fixation points for all subjects is presented in the table that defines how far do members of a group stray from the mean value of the group. Saccadic amplitude is the distance traveled by the eye between two fixation locations. The overall mean saccade amplitude for the subjects ranges from 1768.767 to 6182.264 and their overall standard deviation ranges from 842.032 to 3503.537. From this, you can comprehend there is a considerable difference between the amplitude saccade of the subjects.

Figure 1: comparison by number of fixations b/n algorithms



The above figure demonstrates the number of fixations identified by velocity threshold identification and the U 'n'Eye deep learning model. Both algorithms capture almost the same number of fixations.

Table 2: metrics used in velocity threshold identification algorithm

subject	MFD_true	MFD_SD_true	MFD_false	MFD_SD_false	MSA_true	MSA_SD_true	MSA_false	MSA_SD_false	MFD_overall	MFD_overall_SD	MSA_overall	MSA_overall_SD
S2	0.172	0.053	0.144	0.061	5435.837	3221.201	6621.338	3449.283	0.154	0.060	6182.264	3414.938
S4	0.157	0.049	0.167	0.052	7027.879	5607.174	4708.222	1520.309	0.164	0.051	5411.149	3503.537
S8	0.118	0.024	0.101	0.015	2889.662	749.630	3220.273	1130.792	0.111	0.022	3034.305	949.958
S10	0.104	0.012	0.129	0.030	4977.361	993.046	4571.209	1512.967	0.126	0.030	4605.055	1480.910
S16	0.185	0.011	0.199	0.035	4984.970	1972.415	2710.702	1033.678	0.195	0.032	3235.533	1624.150
S22	0.131	0.032	0.142	0.044	2830.419	741.837	1638.873	622.990	0.139	0.042	1956.619	842.032
S23	0.149	0.015	0.205	0.055	2024.028	936.453	1699.151	1043.371	0.193	0.055	1768.767	1030.065
S30	0.121	0.043	0.115	0.020	3590.710	703.209	5107.893	2562.173	0.116	0.025	4863.186	2428.384

Figure 2: comparison by mean fixation duration b/n algorithms

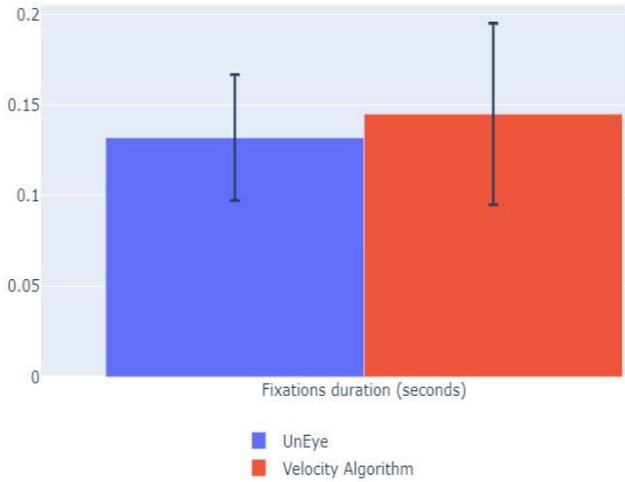
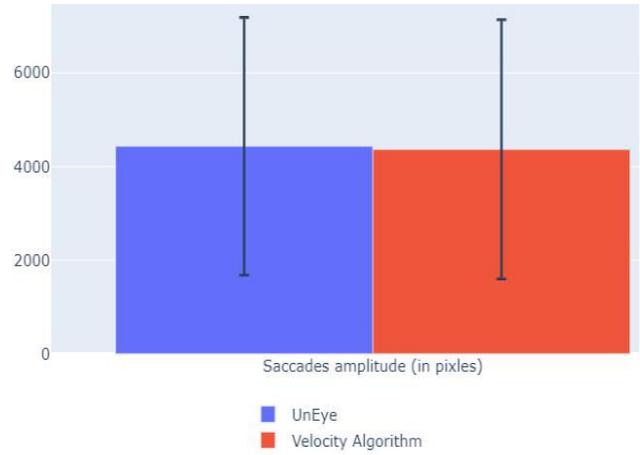


Figure 2 presents the mean fixation duration returned by velocity threshold identification and the U'nEye deep learning model. The mean fixation duration of the velocity threshold identification algorithm is a little bit higher than the deep learning model.

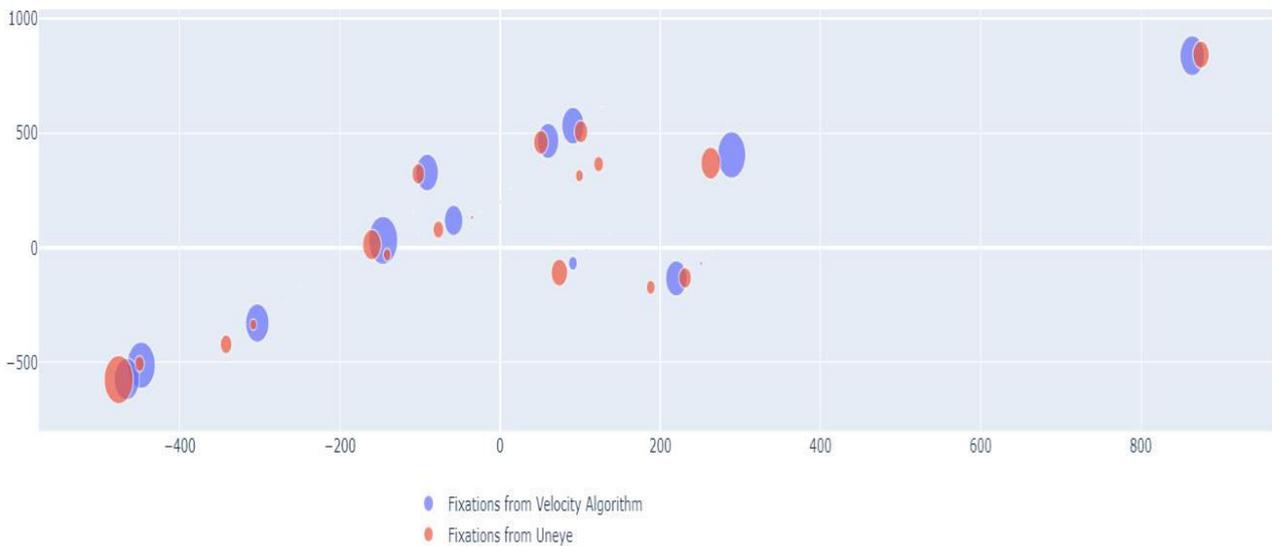
Figure 4: comparison by saccades amplitude b/n algorithms



The above figure demonstrates the saccade amplitude of velocity threshold identification and the U'nEye deep learning model. Almost the amplitude saccade of both algorithms is the same.

Figure 3: visualization of fixation identified by both methods in a two-dimensional plane

UnEye vs Velocity Algorithm for Sample30



The above figure demonstrates the position of the centroid of fixation points in a two-dimensional plane. As in the above figure, most of the fixation points identified by both algorithms are positioned almost in the same location. This implies that both methods work well in identifying the centroid of fixations.

Figure 5: comparison by mean fixation duration b/n recognized and non-recognized images

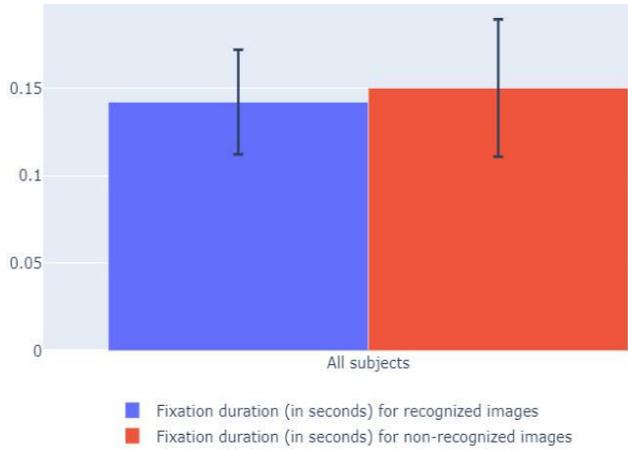
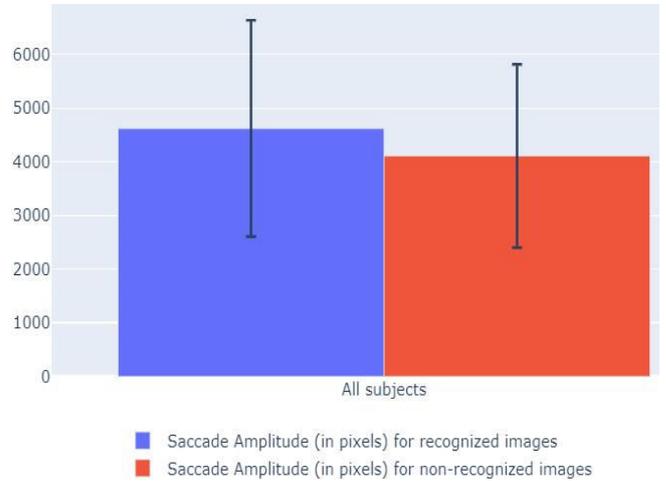


Figure 6: comparison by saccade amplitude b/n recognized and non-recognized images



Figures 5 and 6 present the mean fixation duration and saccade amplitude for recognized images and non-recognized images respectively. The mean fixation duration for both recognized images and non-recognized images is almost close to 1ms.

Figures 7 and 8 depict about mean fixation duration and saccade amplitude for recognized images and non-recognized images respectively of all subjects. Subject id-23 has the highest fixation duration compared to others.

Figure 7: visualization of mean fixation duration for recognized image and non-recognized images for every subject

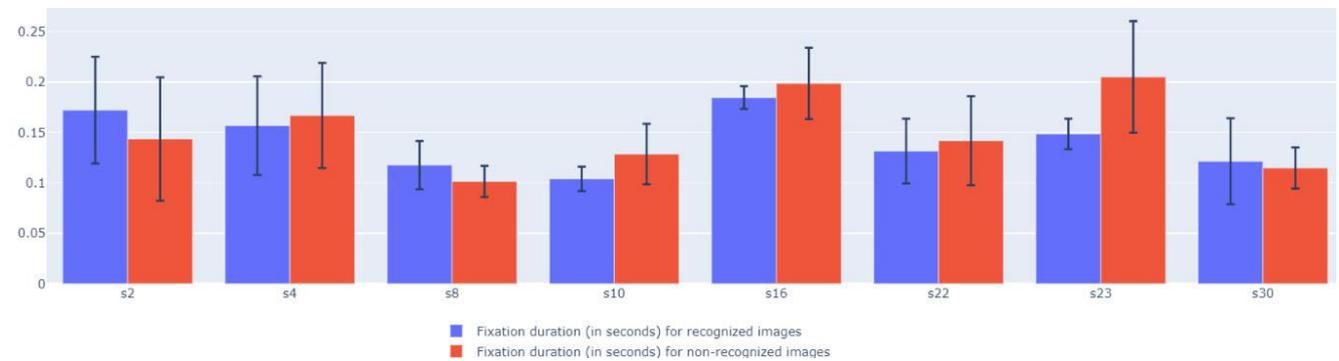
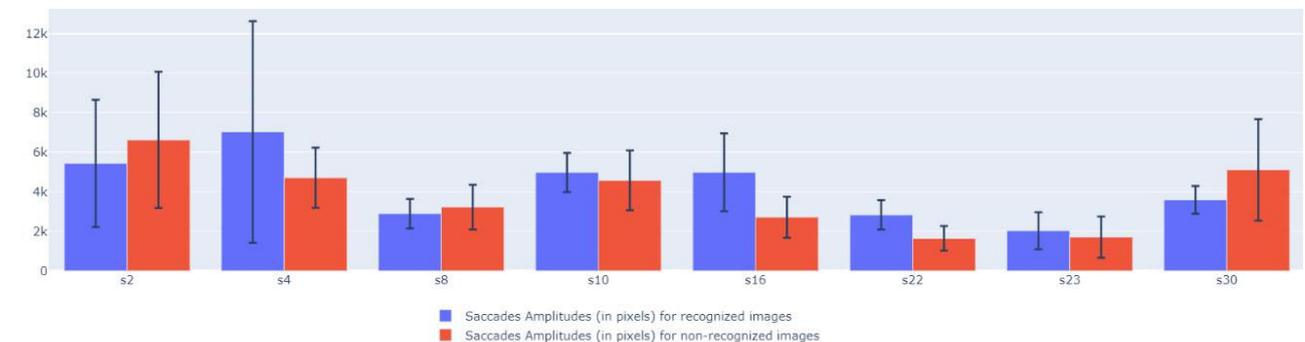


Figure 8: visualization of saccade amplitude for recognized image and non-recognized images for every subject



4. Conclusion

The results of eye movement analysis could help researchers better understand human cognitive and perceptual processes, as well as visual content design principles. Automated design predictions will be made in the future. In the conclusion, in this work, we use velocity threshold identification and U'nEye deep neural network model to capture fixation from the recorded eye-gaze dataset. We present the results of the algorithms with different metrics in a tabular and visual format. Moreover, both fixation identification algorithms are compared in terms of the number of fixations, mean fixation duration, and saccade amplitude. The Velocity-threshold fixation identification was calibrated with different threshold values to achieve better results. Velocity-threshold fixation identification has a higher mean fixation duration compared to the U'nEye deep neural network model and both algorithms capture almost the same number of fixations. The locations of centroids of the captured fixations points in both methods are placed closer to each other.

REFERENCES

- [1] D. Noton and L. Stark, "Scanpaths in saccadic eye movements while viewing and recognizing patterns," *Vision Research*, vol. 11, no. 9, pp. 929-938, Sep. 1971, doi: 10.1016/0042-6989(71)90213-6.
- [2] Suppes P, "Eye-movement models for arithmetic and reading performance."
- [3] M. F. Land and D. N. Lee, "Where we look when we steer," *Nature*, vol. 369, no. 6483, pp. 742-744, 1994, doi: 10.1038/369742a0.
- [4] A. T. Duchowski, "A breadth-first survey of eye-tracking applications," *Behavior Research Methods, Instruments, & Computers*, vol. 34, no. 4, pp. 455-470, 2002, doi: 10.3758/BF03195475.
- [5] J. H. Goldberg and X. P. Kotval, "Computer interface evaluation using eye movements: methods and constructs," *International Journal of Industrial Ergonomics*, vol. 24, no. 6, pp. 631-645, 1999, doi: [https://doi.org/10.1016/S0169-8141\(98\)00068-7](https://doi.org/10.1016/S0169-8141(98)00068-7).
- [6] A. F. Fuchs, "THE SACCADIC SYSTEM," 1971.
- [7] M. ALPERN, "Type of movement," *New York: Academic Press*, pp. 63-151, 1962.
- [8] P. A. C. Marcel Adam Just, *Using Eye Fixations to Study Reading Comprehension*. 1984.
- [9] K. Rayner, "Eye movements in reading and information processing: 20 years of research".
- [10] G. E. Manor BR, "Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks".
- [11] D. E. Irwin, "Memory for position and identity across eye movements.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 18, no. 2, pp. 307-317, 1992, doi: 10.1037/0278-7393.18.2.307.
- [12] C. J. Erkelens and I. M. L. C. Vogels, "The initial direction and landing position of saccades," in *Eye movement research: mechanisms, processes and applications*, J. M. Findlay, R. Walker, and R. W. Kentridge, Eds. Amsterdam: Elsevier, 1995, pp. 133-144. doi: 10.1016/S0926-907X(05)80012-1.
- [13] M. N. R. A. R. D. H. J. J. van de W. Kenneth Holmqvist, *Eye Tracking A Comprehensive Guide to Methods and Measures*. 2011.
- [14] B. A. Tom Tullis, *Behavioral and Physiological Metrics*. 2013.
- [15] D. D. , & G. J. H. Salvucci, "Identifying fixations and saccades in eye-tracking protocols.," 2000.
- [16] J. A. Hartigan, "Clustering algorithms," *New York: John Wiley & Sons.*, 1975.
- [17] J. B. H. N. Z. M. H. P. B. Marie E. Bellet, "Human-level saccade detection performance using deep neural networks," *Neurophysiology*, 2018.

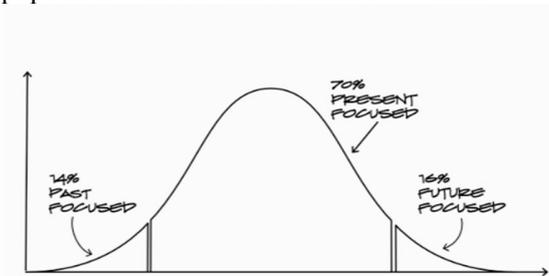
Future Focused Strategy: How to Avoid Getting Fracked

A. Dev Patnaik, B. Ryan Baum, C. Michelle Loret de Mola

Abstract— The world is changing. Look at these recent headlines. Tesla now has a market cap that is bigger than Ford or General Motors combined. Retailers are facing their own private Armageddon, meanwhile Amazon has started doing deliveries by drone. Soda consumption has fallen to a 30-year low, which is a world of hurt if you're Coke or Pepsi.



Most business leaders would agree that having an eye to the future is an essential part of strategy. The trouble is that according to psychologists, only about 16% of the US population have a future focused mindset.



70% of us are wired to be present-focused. This often shows up as acknowledgement that the world is changing and that the future is important but being consumed by the demands of meeting a quarterly target or delivering on near term objectives. Past-focused people see rapid change as “a blip” and might say things like “Cabs will never disappear; ride sharing is

just a fad.” Ironically, past-focused people are easier to deal with because they can be convinced with facts. Present-focused folks are more difficult because they agree with the premise but don't do anything to change their behavior.

The reality is that businesses that don't respond to rapid change get “fracked”. What do we mean by that? For example, a farmer living in the middle of the country might have lived there for decades and be a pretty good farmer. But unbeknownst to him, somebody is breaking shale rock a mile underground near the farm and it's leaching all sorts of chemicals up into the soil and it starts polluting everything underneath the land. The farmer wakes up one day, turns on the faucet and the water lights on fire. He has been fracked.

In the same way, businesses are getting fracked and being disrupted by external forces, exogenous forces, that have nothing to do with how good they are or how long they've been around.

Five strategies for staying future focused. *-Illustrative stories italicized*

1. Don't try to predict the future. *-Royal Dutch Shell and the origin of scenario planning*
2. Consider the unlikely possible. *-Donald Trump's surprise victory in 2016*
3. Envision changes in cultures and markets, not just technology. *-mobile phones vs the fall of the Soviet Union*
4. Find the future now. *-William Gibson, "The future exists today, it's just not widely distributed"*
5. Trade away everything that doesn't matter. *- Amazon & Kindle*

The world is changing. If leaders don't pay attention to that change, if they don't learn from that change, they're going to get fracked. But if they're future focused, and relentlessly think about what might be

A. Dev Patnaik is the Founder and CEO of Jump in San Mateo CA 94401. He also teaches Needfinding at Stanford. (email: dev@jumpassociates.com)

B. Ryan Baum is a Partner at Jump in San Mateo CA 94401. He also teaches Needfinding at Stanford. (email: ryan@jumpassociates.com)

C. Michelle Loret de Mola is a Director at Jump in San Mateo CA 94401. (email: michelle@jumpassociates.com)

possible, they can start planning for this kind of change long before it happens.

Keywords—Change, Fracked, Future-focused, Strategy

Multi-Robotic Partial Disassembly Line Balancing with Robotic Efficiency Difference via HNSGA-II

Tao Yin, Zeqiang Zhang, Wei Liang, Yanqing Zeng, Yu Zhang

Abstract—To accelerate the remanufacturing process of waste electronic products, this study designs a partial disassembly line with multi-robotic station to effectively dispose of excessive wastes. Multi-robotic partial disassembly line is a technical upgrade to the existing manual disassembly line. Balancing optimization can make the disassembly line smoother and more efficient. For partial disassembly line balancing with multi-robotic station (PDLBMS), a mixed-integer programming model (MIPM) considering the robotic efficiency differences is established to minimize cycle time, energy consumption and hazard index, and to calculate their global optimal values. Besides, an enhanced NSGA-II algorithm (HNSGA-II) is proposed to optimize PDLBMS efficiently. Finally, MIPM and HNSGA-II are applied to an actual mixed disassembly case of two types of computers, the comparison of the results solved by GUROBI and HNSGA-II verifies the correctness of the model and excellent performance of the algorithm, and the obtained Pareto solution set provides multiple options for decision-makers.

Keywords—Waste disposal, Disassembly line balancing, Multi-robot station, Robotic efficiency difference, HNSGA-II

I. INTRODUCTION

WITH the deterioration of environment and the shortage of natural resources, recycling waste products has become an important means to obtain raw materials for new products. Usually, the waste electronic products contain not only valuable parts that can be reused but also substances harmful to the environment. For example, CPU, memory and hard disk in computers are reusable, while batteries are harmful and need harmless treatment. Therefore, building a complete recycling system can reduce the impact of harmful substances on the environment and the waste of valuable resources. In this system, disassembly is a necessary process to separate parts from the structure of products. Large-scale resource recycling companies generally adopt disassembly line to dispose of wastes. At present, the research on disassembly line includes disassembly line design [1], disassembly line balancing (DLB) [2], and disassembly line sequence planning [3], etc. This study

This work is supported in part by the National Natural Science Foundation of China under Grant 51205328 and 51675450; in part by the Youth Foundation for Humanities and Social Sciences of Ministry of Education of China under Grant 18YJC630255; in part by the Sichuan Science and Technology Program under Grant 2019YFG0285; and in part by the Special Project of Postgraduate Academic Literacy Improvement Program under Grant 2021KCJS16. (Corresponding author: Zeqiang Zhang)

Tao Yin, Zeqiang Zhang, Wei Liang, Yanqing Zeng, and Yu Zhang are with the Technology and Equipment of Rail Transit Operation and Maintenance Key Laboratory of Sichuan Province, School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China (e-mail: amniyim@my.swjtu.edu.cn; zhangzq@home.swjtu.edu.cn; liangwei@my.swjtu.edu.cn; zyq2017200281@163.com; chaizhong1997@gmail.com).

belongs to the field of DLB.

DLB is a classical multi-objective combinatorial optimization problem [4]. Its optimization process is to assign the discrete tasks in the planned sequence to the sequential stations and make the multiple objectives optimal [5]. The existing research on DLB mainly focuses on single manned station mode, that is, only one worker in each station completes all tasks in this station. To improve the internal efficiency of stations, Cevikcan et al. put forward a multi-manned station disassembly line [6]. Considering the line efficiency and the worker safety, Yilin Fang et al. proposed a DLB with multi-robotic station (DLBMS) [7]. Subsequently, Yilin Fang, Quan Liu and others made an in-depth study on DLBMS [8]–[10]. However, all these studies employed complete disassembly mode. Different from the complete disassembly, only the valuable, hazardous and necessary parts are removed, and the remains are directly crushed for raw materials, this is called partial disassembly [11]. Partial disassembly can avoid invalid workload and reduce costs [12]. Thus, this study pays more attention to the partial DLBMS (PDLBMS). In reality, the efficiency of robots disassembling different tasks is also different due to the difference of robot models and equipped tools. To get a practical disassembly scheme, this study optimizes the PDLBMS considering robotic efficiency differences for the first time. The optimization objectives are cycle time of stations, energy consumption of robots, and hazard index of tasks.

In addition, DLB is an NP-hard problem [13]. For the products with small-scale tasks, the common methods include linear programming [14], nonlinear programming [15], and mixed-integer programming [5], etc. While facing the products with large-scale tasks, the mainstream methods are meta-heuristic algorithms, such as hummingbird algorithm [16], firefly algorithm [17], and whale optimization algorithm [18], etc. To effectively solve PDLBMS, a mixed-integer programming model (MIPM) which can calculate the single-objective optimal values is established, and an enhanced NSGA-II (HNSGA-II) is proposed. Finally, the MIPM and HNSGA-II are applied to an actual mixed disassembly case of two types of computers, and correctness of the model and excellent performance of the algorithm are verified by comparing their optimization results.

The main contributions of this study are as follows: 1. DLBMS is expanded from complete disassembly to partial disassembly; 2. Robotic efficiency difference is considered for the first time; 3. MIPM is established to calculate the single-objective global optimal values; 4. HNSGA-II is proposed to efficiently solve PDLBMS; 5. MIPM and

HNSGA-II are applied to an actual computer disassembly case.

The rest of this paper is organized as follows. Section II describes the new PDLBMS and establishes the MIPM. Section III introduces the HNSGA-II. Section IV employs the model and HNSGA-II to solve an actual mixed disassembly case of two types of computers, and gives the comparison of their optimization schemes. Section V concludes this study and discusses the future research work.

II. PROBLEM FORMULATION

A. Problem Description

Fig. 1 shows the schematic diagram of PDLBMS. Two different types of computers enter the disassembly line from the entrance, stations 1-3 are equipped with multiple robots to complete the corresponding tasks. Because of partial disassembly, the parts that do not need to be disassembled will flow out through the outlet of the line and be sent to the crushing workshop for raw materials. After investigation, it is found that most of the existing lines only disassemble a single product, while our designed line can disassemble many types of products.

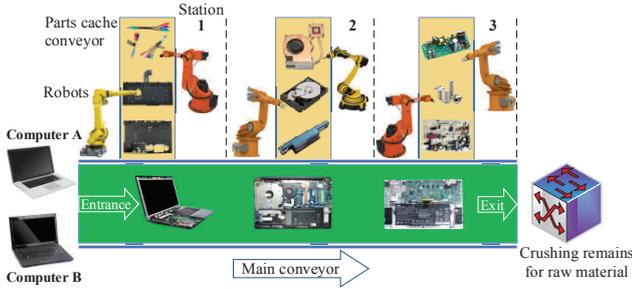


Fig. 1 Schematic diagram of PDLBMS

B. Notations

- a Product number.
- I Tasks set of products, its cardinality is N_a .
- W Station set, its cardinality is N_w .
- R Robot set, its cardinality is N_r .
- i, j Task number, $i, j \in I$.
- w Station number, $w \in W$.
- r Robot number, $r \in R$.
- RL_{max} Robot limitation in per station.
- s_i Starting time of task i .
- t_{ir} Working time of robot r disassembling task i .
- Tc Cycle time of stations. (s)
- OE_r Operational energy consumption of robot r . (kW·h)
- SE_r Standby energy consumption of robot r .
- E_w Total energy consumption of station w .
- B Large positive number.
- $P_a(i)$ Immediate predecessors set of task i in product a .
- d_i Demand attribute. 1, task i is demanded, or 0.
- h_i Hazard attribute. 1, task i is hazardous, or 0.
- m_{iwr} Task assignment variable. 1, task i is assigned to robot r in station w , or 0.
- n_{wr} Robot assignment variable. 1, robot r is assigned to station w , or 0.

- k_{ijsw} Task position variable. 1, tasks i and j are assigned to robot r in station w and j is behind i , or 0.
- s_w Station open variable. 1, station w is opened, or, 0.

C. Optimization Objectives

1) *Cycle time*: Optimizing cycle time can avoid robots waiting and parts stacking, and make the line smoother and more efficient, so it is regarded as the first objective:

$$\min f_1 = Tc \quad (1)$$

2) *Energy consumption*: As an important factor in the robotic disassembly line, energy consumption objective can be divided into two optimization indexes, namely peak energy consumption of stations and total energy consumption. Their expressions are as follows:

$$\min f_2 = \max(E_w) \quad (2)$$

$$\min f_3 = \sum_{w \in W} E_w \quad (3)$$

$$E_w = \sum_{r \in R} \left(\sum_{i \in I} OE_r \cdot m_{iwr} \cdot t_{ir} + SE_r \cdot (Tc - \sum_{i \in I} m_{iwr} \cdot t_{ir}) \right) \quad \forall w \in W \quad (4)$$

3) *Hazard index*: Removing the hazardous parts early can effectively avoid environmental pollution, so another objective is the hazard index:

$$\min f_4 = \sum_{i \in I} \sum_{w \in W} \sum_{r \in R} m_{iwr} \cdot (s_i + t_{ir}) \cdot h_i \quad (5)$$

D. Constraints of PDLBMS

The optimization process of PDLBMS needs to meet the disassembly mode constraints, precedence relationship constraints, cycle time constraints, task time constraints, task assignment constraints, station configuration constraints, and robot configuration constraints.

1) *Disassembly mode constraints*: Because of partial disassembly, except for the demanded, hazardous, and necessary tasks, other tasks will not be disassembled. The expressions of partial disassembly are as follows:

$$\sum_{w \in W} \sum_{r \in R} m_{iwr} \leq 1 \quad \forall i \in I \quad (6)$$

$$\sum_{w \in W} \sum_{r \in R} m_{iwr} = 1 \quad \forall i \in \{i | h_i + d_i \geq 1, i \in I\} \quad (7)$$

2) *Precedence relationship constraints*: Some tasks must be disassembled in order due to the connection relationships or spatial position constraints between parts, this is called precedence relationship. In partial disassembly, if the immediate succeeding task is disassembled, its immediate preceding tasks must be disassembled. While if the immediate preceding task is disassembled, its immediate succeeding tasks may not be disassembled. This constraint can be expressed below:

$$1 \leq \sum_{w \in W} \sum_{r \in R} m_{iwr} + B \cdot \left(1 - \sum_{w \in W} \sum_{r \in R} m_{jwr} \right) \quad \forall j \in I, \forall i \in P_a(j) \quad (8)$$

The precedence relationship constraint is reflected in the timeline that the starting time of the immediate succeeding task must be later than the complete time of its immediate preceding tasks. The expression is shown below:

$$\begin{aligned} & \sum_{w \in W} \sum_{r \in R} m_{jwr} \cdot s_j + B \cdot (1 - \sum_{w \in W} \sum_{r \in R} m_{jwr}) \\ & \geq \sum_{w \in W} \sum_{r \in R} m_{iwr} \cdot (s_i + t_{ir}) \quad \forall j \in I, \forall i \in P_a(j) \end{aligned} \quad (9)$$

3) *Cycle time constraints*: First, the working time of each task cannot exceed one cycle time:

$$\sum_{i \in I} m_{iwr} \cdot t_{ir} \leq Tc \quad \forall w \in W, \forall r \in R \quad (10)$$

In the process of assigning tasks to stations, the starting time of task i disassembled by the current robot in the current station must be large than the sum of the starting time of the current station and the total working time of all tasks assigned to the current robot earlier than task i . This constraint can be expressed below:

$$\begin{aligned} & Tc \cdot (\sum_{w \in W} \sum_{r \in R} m_{jwr} \cdot w - 1) + \sum_{w \in W} \sum_{r \in R} \sum_{i \in I, i \neq j} k_{ijwr} \cdot t_{ir} \\ & \leq \sum_{w \in W} \sum_{r \in R} m_{jwr} \cdot s_j + B \cdot (1 - \sum_{w \in W} \sum_{r \in R} m_{jwr}) \quad \forall j \in I \end{aligned} \quad (11)$$

In addition, the complete time of every task assigned to the current station must be less than the complete time of the current station. The expression is as follows:

$$\sum_{w \in W} \sum_{r \in R} m_{iwr} \cdot (s_i + t_{ir}) \leq Tc \cdot (\sum_{w \in W} \sum_{r \in R} m_{iwr} \cdot w) \quad \forall i \in I \quad (12)$$

4) *Task time constraints*: Taking the starting time of the robot assigned to the first station as the starting point of the timeline, so the starting time of any task should be non-negative, and the constraint expression is as follows:

$$\sum_{w \in W} \sum_{r \in R} m_{iwr} \cdot s_i \geq 0 \quad \forall i \in I \quad (13)$$

For any two tasks assigned to the same robot in the same station, the latter task must wait until the former task is completed. This ensures that a robot can only disassemble one task at once:

$$\begin{aligned} & B \cdot (1 - k_{ijwr}) + B \cdot (1 - \sum_{w \in W} \sum_{r \in R} m_{jwr}) + \sum_{w \in W} \sum_{r \in R} m_{jwr} \cdot s_j \\ & \geq \sum_{w \in W} \sum_{r \in R} m_{iwr} \cdot (s_i + t_{ir}) \quad \forall i, j \in I, i \neq j, \forall w \in W, \forall r \in R \end{aligned} \quad (14)$$

5) *Task assignment constraint*: If two tasks are assigned to the same robot in a station, these two tasks cannot be assigned to other robots or other stations. This can be constrained below:

$$\begin{aligned} & m_{iwr} + m_{jwr} \leq 1 + (k_{ijwr} + k_{jiwr}) \\ & \forall i, j \in I, i < j, \forall w \in W, \forall r \in R \end{aligned} \quad (15)$$

$$\begin{aligned} & 0.5 \cdot (m_{iwr} + m_{jwr}) \geq k_{ijwr} + k_{jiwr} \\ & \forall i, j \in I, i < j, \forall w \in W, \forall r \in R \end{aligned} \quad (16)$$

6) *Station configuration constraints*: Although the number of stations is initially given as N_w , it varies during the optimization process. When all tasks are assigned to one station, only the first station needs to be opened. When there are many tasks in the given stations, they need to be all opened. The constraint of the number of opened stations is as follows:

$$1 \leq \sum_{w \in W} s_w \leq N_w \quad (17)$$

The number of tasks that can be assigned to each station

ranges from 0 to N_a . When tasks are assigned to a station, the station is opened; otherwise, the station is closed. This constraint can be expressed below:

$$s_w \leq \sum_{i \in I} \sum_{r \in R} m_{iwr} \leq s_w \cdot N_a \quad \forall w \in W \quad (18)$$

In addition, according to the layout requirements of the disassembly line, stations should be opened sequentially:

$$s_w \leq s_{w-1} \quad \forall w \in W, w \neq 1 \quad (19)$$

7) *Robot configuration constraints*: In the given N_r robots, some robots can be employed or not:

$$\sum_{w \in W} n_{wr} \leq 1 \quad \forall r \in R \quad (20)$$

When a task is assigned to a robot in a station, the robot must be employed in the same station:

$$\sum_{i \in I} m_{iwr} - B \cdot n_{wr} \leq 0 \quad \forall w \in W, \forall r \in R \quad (21)$$

There is an upper bound and a lower bound on the number of robots employed in each station. The station without robots will not be opened. The number of robots assigned in an opened station ranges from 1 to RL_{max} due to space limitations. This can be constrained below:

$$0 \leq \sum_{r \in R} n_{wr} \leq RL_{max} \quad \forall w \in W \quad (22)$$

III. PROPOSED HNSGA-II

NSGA-II is a classic multi-objective meta-heuristic algorithm with its simple structure and excellent performance [19]. Due to the problem characteristics of PDLBMRs, the crossover and mutation operations of NSGA-II need to be improved. To pursue better solutions, a spur strategy is proposed to enhance NSGA-II. The enhanced algorithm is called HNSGA-II. The structure of HNSGA-II includes encoding and decoding, new solution generation operation (crossover, mutation, and spur strategy), and update population.

A. Encoding and Decoding

In the optimization process, apart from the necessary disassembly task sequence (DT), it also needs to construct the station numbering sequence (SN) and robot numbering sequence (RN) to present the open status of stations and the employed status of robots, respectively. The DT can be generated according to the precedence relationships of products. The SN size is the number of total tasks N_a , and its elements are the station numbers sorting in ascending order. The RN size is the total given robots N_r , its elements are the random arrangement of multiple zeros and station numbers, its index is the robot number, and the element zero indicates that the robot corresponding to the index is not employed. An example of the three sequences meeting the requirements is shown in Fig. 2. The red annotations "d" and "h" denote the demanded and hazardous attributes of tasks.

Decoding consists of two parts. One is to assign tasks and robots to stations. The assignment result of DT and RN in Fig. 2 is shown in Fig. 3. Because of partial disassembly, the tasks {5,11,3,9} are not disassembled and station 3 is not opened.

The tasks and robots assigned to station 1 are $\{1,6,2\}$ and $\{3,7,8\}$, and those assigned to station 2 are $\{4,7,8,10\}$ and $\{2,5,9\}$.

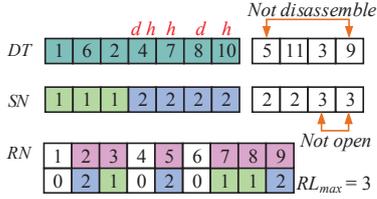


Fig. 2 An example of DT , SN , and RN

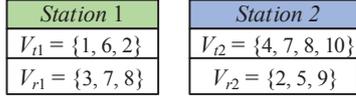


Fig. 3 The first part of decoding

Based on the results in Fig. 3, another part of the decoding is to assign the tasks in each station to the robots in the corresponding stations. Taking station 1 as an example to explain the assignment approach: Task 1 will be assigned to the robot which has the shortest disassembly working time in robot set $\{3,7,8\}$. If the time disassembling task 1 by robots 3,7,8 are the same, the robot with the least operational energy consumption will be selected to disassemble task 1. The assignment of other tasks is the same as that of task 1, and also needs to obey the constraints in the previous section.

B. New Solution Generation Operation

Similar to the original genetic algorithm, the three sequences DT , SN , RN are simultaneously performed the crossover and mutation operations to generate three new sequences. It is worth noting that: a. The new DT needs to meet the precedence relationship constraints; b. The new SN needs to be sorted in ascending order; c. The new RN needs to meet the limitation of RL_{max} . When the newly generated sequences do not meet the relevant constraints, repair method should be used to improve the crossover and mutation operations.

In addition, to obtain better solutions, the current non-inferior solution set in each iteration and the four single-objective optimal solutions in this set are selected as parents to be performed the crossover operation, the selection method is called spur strategy. Because the parents are excellent, the new solutions generated by the spur strategy are also considered to be excellent. Thus, the spur strategy is considered as an effective method to enhance the original NSGA-II.

C. Update Population

Non-dominated sorting approach and crowding distance of the original NSGA-II are employed to update the population. Besides, an external storage E is designed to screen the required number of non-inferior solutions, and the screening method is the crowding distance.

IV. APPLICATION TO COMPUTER DISASSEMBLY LINE

The established MIPM and the proposed HNSGA-II are applied to an actual mixed disassembly case of two types of

computers. The model is developed by the exact solver GUROBI, and the HNSGA-II is programmed by MATLAB 2014a. Their running environment is Win10 system with an Intel (R) Core (TM) i5-9400 2.9 GHz and 8 GB RAM. By comparing the optimization results of these two methods, correctness of the model and superiority of the algorithm are verified.

A. Data Preparation

The precedence relationships of computer A and computer B come from the literature [20] and are shown in Fig. 4. The number of tasks in computers A and B is 8 and 10, so the total number is 18. Because the efficiency of 16 given robots is different, the time for 16 robots to disassemble the 18 tasks is also different, and the disassembly time is shown in TABLE I.

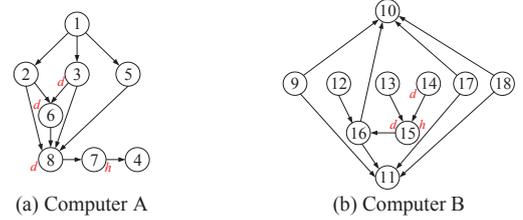


Fig. 4. Precedence relationships of two computers

TABLE I
WORKING TIME OF 16 ROBOTS DISASSEMBLING 18 TASKS

Tasks	Robots															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	2	11	14	10	6	11	14	3	14	3	3	3	3	4	10	7
2	10	6	10	8	9	7	7	8	15	6	2	6	8	15	7	14
3	15	10	3	15	12	11	4	10	7	13	4	14	13	4	14	11
4	14	4	6	3	13	14	13	4	9	10	14	11	15	5	5	11
5	2	8	10	7	6	7	6	10	12	9	6	12	9	11	13	6
6	11	2	9	10	6	13	6	14	12	6	9	9	13	4	16	3
7	7	5	5	4	7	4	5	5	8	4	9	4	10	3	6	6
8	4	6	5	15	11	9	3	4	7	6	10	15	13	11	12	5
9	10	9	13	11	14	11	10	3	6	4	7	14	2	2	11	7
10	8	5	11	15	13	12	8	11	3	6	8	8	9	16	8	14
11	9	7	4	11	14	6	13	2	8	4	8	14	6	10	9	4
12	14	12	12	2	14	6	9	6	4	14	16	7	6	7	15	4
13	4	8	8	6	4	15	7	16	8	16	3	7	16	12	15	9
14	8	3	8	4	14	11	3	13	15	5	8	13	5	15	12	10
15	3	7	12	11	15	3	5	5	9	12	13	3	9	7	7	4
16	7	9	8	3	13	11	10	12	15	3	10	15	5	12	3	11
17	12	14	2	11	14	5	8	4	3	11	6	2	10	10	10	14
18	3	6	3	9	3	13	15	11	16	11	15	6	16	13	10	13

Energy consumption of robots is divided into operational energy consumption (OE_r) and standby energy consumption (SE_r) per unit time. The energy consumption data of 16 robots are shown in TABLE II.

TABLE II
ENERGY CONSUMPTION DATA OF 16 ROBOTS

Robot	1	2	3	4	5	6	7	8
OE	6.32	6.57	5.92	7.24	6.63	6.40	9.66	7.00
SE	0.63	0.66	0.59	0.72	0.66	0.64	0.97	0.70
Robot	9	10	11	12	13	14	15	16
OE	6.90	7.96	5.34	6.03	8.62	7.88	6.00	9.22
SE	0.69	0.80	0.53	0.60	0.86	0.79	0.60	0.92

In addition, the number of the given stations N_w is 4, and the robot limitation in per station RL_{max} is 3.

B. Optimization Results and Analysis

1) *Optimization results:* The algorithm parameters are set: population size $M=300$, total number of iterations $N=500$, and external storage size $N_E=10$. After running the HNSGA-II 10 times, one of the obtained optimal disassembly scheme sets is listed in TABLE III.

TABLE III
AN OPTIMAL DISASSEMBLY SCHEME SET OBTAINED BY HNSGA-II

No.	Disassembly schemes																			
	Must disassemble										Not disassemble									
1 ^(b)	DT	13	1	2	14	3	6	5	8	15	7	12	4	9	17	16	18	10	11	
	SN	1	1	1	2	2	2	3	3	4	4	4	4	4	4	4	4	4	4	
	RN	3	2	2	3	0	4	3	0	0	4	1	1	2	4	0	1			
2 ^(c)	DT	1	13	2	5	3	14	6	15	8	7	9	12	18	16	17	11	10	4	
	SN	1	1	1	1	2	2	2	3	3	4	4	4	4	4	4	4	4	4	
	RN	1	2	2	0	3	2	0	3	1	4	1	3	0	4	4	0			
3 ^(a)	DT	1	13	5	14	2	3	15	6	8	7	4	12	16	9	17	18	10	11	
	SN	1	1	1	1	2	2	2	3	4	4	4	4	4	4	4	4	4	4	
	RN	1	2	2	1	3	0	0	3	0	0	1	2	4	4	3	4			
4	DT	14	1	13	3	5	2	15	6	8	7	4	12	9	16	18	17	10	11	
	SN	1	1	1	1	1	1	1	2	2	3	3	3	3	3	3	4	4	4	
	RN	1	2	0	4	0	0	1	2	2	3	1	3	4	0	3	4			
5	DT	14	1	13	3	5	2	15	6	8	7	4	18	12	9	16	17	11	10	
	SN	1	1	1	1	1	1	1	2	2	3	3	3	3	3	3	3	3	4	
	RN	1	2	3	0	0	0	1	2	0	4	1	3	4	3	2	4			
6	DT	13	14	1	15	2	3	5	6	8	7	18	9	12	16	4	17	10	11	
	SN	1	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	4	
	RN	1	1	4	0	4	3	2	0	3	0	1	0	4	2	3	2			
7	DT	14	13	1	15	2	3	6	5	8	7	17	12	16	4	9	18	10	11	
	SN	1	1	1	1	1	1	2	2	2	2	2	2	2	3	4	4	4	4	
	RN	2	2	0	1	4	3	0	3	0	3	1	1	4	2	4	0			
8	DT	13	14	1	2	15	3	6	5	8	7	18	9	12	17	16	4	11	10	
	SN	1	1	1	1	1	1	1	1	2	2	3	3	3	3	3	3	4	4	
	RN	1	1	3	0	4	3	4	2	0	4	1	2	2	0	0	3			
9	DT	14	13	1	2	15	3	5	6	8	7	18	9	4	12	17	16	10	11	
	SN	1	1	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	4	
	RN	1	1	3	3	0	4	0	2	2	4	1	0	3	2	0	4			
10 ^(d)	DT	13	14	1	15	2	3	5	6	8	7	18	9	4	12	17	16	10	11	
	SN	1	1	1	1	1	1	1	1	2	2	2	2	2	3	3	3	4	4	
	RN	1	1	3	0	4	3	2	0	3	0	1	0	4	2	4	2			

The Pareto solution set corresponding to the 10 optimal schemes in TABLE III and the four single-objective global optimal values solved by GUROBI are shown in TABLE IV. The bold numbers in TABLE IV indicate the single-objective global optimal values.

TABLE IV
THE RESULTS BY GUROBI AND HNSGA-II

Method	No.	f_1	f_2	f_3	f_4	Time/s
GUROBI	1 ^(a)	5	-	-	-	197.65
	2 ^(b)	-	51.79	-	-	262.90
	3 ^(c)	-	-	177.61	-	123.30
	4 ^(d)	-	-	-	23	485.77
HNSGA-II	1 ^(b)	5	51.79	188.93	36	115.03
	2 ^(c)	5	52.61	177.61	31	
	3 ^(a)	5	82.29	189.56	26	
	4	7	139.62	212.08	25	
	5	7	139.62	212.96	24	
	6	9	117.23	213.59	24	
	7	9	118.6	204.55	25	
	8	11	132.69	193.91	25	
	9	11	132.69	195.55	24	
	10 ^(d)	11	132.69	199.39	23	

It is observed from TABLE IV that the HNSGA-II can obtain a Pareto solution set in a single calculation, and the set includes four single-objective global optimal values, which proves the correctness of the model and HNSGA-II. Besides, it is found

that the time of obtaining a Pareto solution set is 115.03s, which is less than the time of obtaining four single-objective global optimal values by GUROBI. This indicates that the efficiency of HNSGA-II for PDLBMS is higher than GUROBI, and further shows the superiority of HNSGA-II.

2) *Disassembly scheme analysis:* Fig. 5 and Fig. 6 show the Gantt diagrams of disassembly schemes corresponding to the four single-objective global optimal values (a-d) obtained by GUROBI and HNSGA-II. Thereinto, w represents the opened station, r represents the employed robots, the green, pink, yellow and transparent rectangles represent the demanded, hazardous, both demanded and hazardous, and normal tasks. From Fig. 5 and Fig. 6, it can be found that except for the same scheme of optimal hazard index f_4_min , schemes of the other three optimal objectives obtained by HNSGA-II are all superior to those of GUROBI. This shows the superiority of HNSGA-II from the quality of solutions. The reason for the results is that GUROBI can only optimize a single objective one by one, while HNSGA-II can optimize four objectives simultaneously.

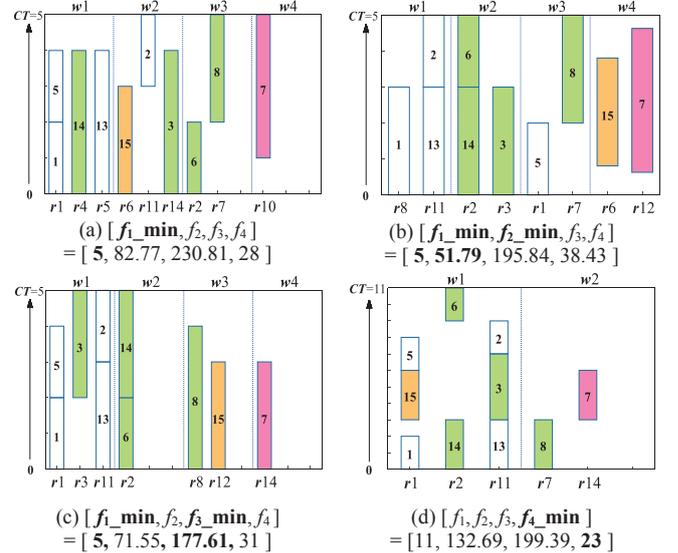


Fig. 5 Four single-objective global optimal schemes by GUROBI

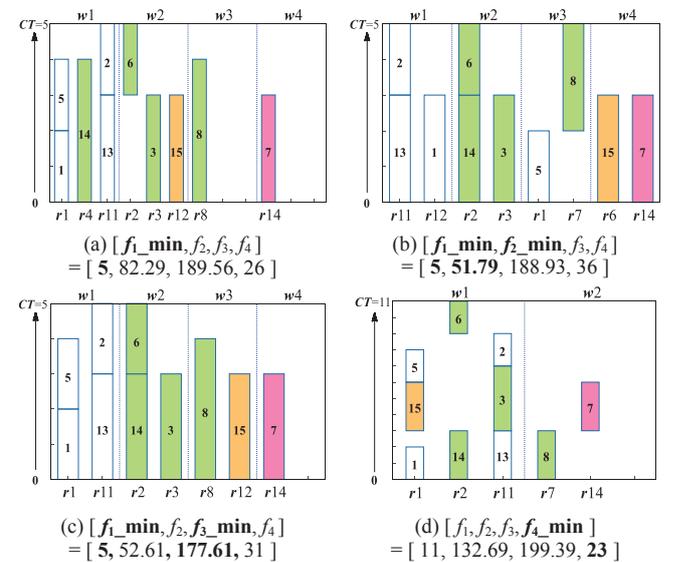


Fig. 6 Four single-objective global optimal schemes by HNSGA-II

It is worth mentioning that the Pareto set containing several non-inferior solutions provides rich options for decision-makers. When the decision-makers need to select one optimal scheme, weighting the Pareto solution set is an effective screening method.

V. CONCLUSIONS AND FUTURE RESEARCH WORK

The correctness of PDLBMRS model and the superiority of HNSGA-II are verified by optimizing the mixed disassembly case of two types of computer, which further indicates the established model can perfectly express the objectives and constraints of PDLBMRS.

Future research work: 1. PDLBMRS will be expanded from the straight line to the U-shape; 2. Considering the end of life state of products will make the PDLBMRS more practical; 3. Each station with multiple robots and multiple workers will be a new disassembly mode in DLB.

REFERENCES

- [1] G. Lowe and R. Bogue, "Design for disassembly: A critical twenty first century discipline," *Assem. Autom.*, vol. 27, no. 4, pp. 285–289, Oct. 2007.
- [2] K. Wang, X. Li, L. Gao, and P. Li, "Modeling and balancing for disassembly lines considering workers with different efficiencies," *IEEE Trans. Cybern.*, to be published.
- [3] X. Guo, S. Liu, M. Zhou, and G. Tian, "Disassembly sequence optimization for large-scale products with multiresource constraints using scatter search and petri nets," *IEEE Trans. Cybern.*, vol. 46, no. 11, pp. 2435–2446, Nov. 2016.
- [4] L. Zhu, Z. Zhang, Y. Wang, and N. Cai, "On the end-of-life state oriented multi-objective disassembly line balancing problem," *J. Intell. Manuf.*, vol. 31, no. 6, pp. 1403–1428, Dec. 2019.
- [5] T. Yin, Z. Zhang, Y. Zhang, T. Wu, and W. Liang, "Mixed-integer programming model and hybrid driving algorithm for multi-product partial disassembly line balancing problem with multi-robot workstations," *Robot. Comput. Integr. Manuf.*, to be published.
- [6] E. Cevikcan, D. Aslan, and F. B. Yeni, "Disassembly line design with multi-manned workstations: a novel heuristic optimisation approach," *Int. J. Prod. Res.*, vol. 58, no. 3, pp. 649–670, Mar. 2019.
- [7] Y. Fang, H. Wei, Q. Liu, Y. Li, Z. Zhou, and D. T. Pham, "Minimizing energy consumption and line length of mixed-model multi-robotic disassembly line systems using multi-objective evolutionary optimization," in *Proc. 14th Int. Manuf. Sci. Eng. Conf. MSEC 2019*, Erie, PA, USA, 2019, pp. 1–10.
- [8] Y. Fang, Q. Liu, M. Li, Y. Laili, and D. T. Pham, "Evolutionary many-objective optimization for mixed-model disassembly line balancing with multi-robotic workstations," *Eur. J. Oper. Res.*, vol. 276, no. 1, pp. 160–174, Dec. 2018.
- [9] Q. Liu, Y. Li, Y. Fang, Y. Laili, P. Lou, and D. T. Pham, "Many-objective best-order-sort genetic algorithm for mixed-model multi-robotic disassembly line balancing," in *Proc. 11th CIRP Conf. IPS*, Zhuhai, China, 2019, pp. 14–21.
- [10] H. Ming, Q. Liu, and D. T. Pham, "Multi-robotic disassembly line balancing with uncertain processing time," in *Proc. 11th CIRP Conf. IPS*, Zhuhai, China, 2019, pp. 71–76.
- [11] L. Zhu, Z. Zhang, and C. Guan, "Multi-objective partial parallel disassembly line balancing problem using hybrid group neighbourhood search algorithm," *J. Manuf. Syst.*, vol. 56, no. 6, pp. 252–269, Jul. 2020.
- [12] K. Wang, X. Li, and L. Gao, "Modeling and optimization of multi-objective partial disassembly line balancing problem considering hazard and profit," *J. Clean. Prod.*, vol. 211, pp. 115–133, Feb. 2019.
- [13] S. M. McGovern and S. M. Gupta, "A balancing method and genetic algorithm for disassembly line balancing," *Eur. J. Oper. Res.*, vol. 179, no. 3, pp. 692–708, Jun. 2007.
- [14] K. Igarashi, T. Yamada, S. M. Gupta, M. Inoue, and N. Itsubo, "Disassembly system modeling and design with parts selection for cost, recycling and CO2 saving rates using multi criteria optimization," *J. Manuf. Syst.*, vol. 38, pp. 151–164, Jan. 2016.
- [15] M. L. Bentina, O. Battaia, A. Dolgui, and S. J. Hu, "Second order conic approximation for disassembly line design with joint probabilistic constraints," *Eur. J. Oper. Res.*, vol. 247, no. 3, pp. 957–967, Dec. 2015.
- [16] T. Yin, Z. Zhang, and J. Jiang, "A Pareto-discrete hummingbird algorithm for partial sequence-dependent disassembly line balancing problem considering tool requirements," *J. Manuf. Syst.*, vol. 60, pp. 406–428, Jul. 2021.
- [17] L. Zhu, Z. Zhang, and Y. Wang, "A Pareto firefly algorithm for multi-objective disassembly line balancing problems with hazard evaluation," *Int. J. Prod. Res.*, vol. 56, no. 24, pp. 7354–7374, Dec. 2018.
- [18] Y. Zhang, Z. Zhang, C. Guan, and P. Xu, "Improved whale optimisation algorithm for two-sided disassembly line balancing problems considering part characteristic indexes," *Int. J. Prod. Res.*, to be published.
- [19] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.
- [20] S. M. McGovern and S. M. Gupta, "Ant colony optimization for disassembly sequencing with multiple objectives," *Int. J. Adv. Manuf. Technol.*, vol. 30, no. 5–6, pp. 481–496, Sep. 2006.

Resilient Manufacturing: Using Augmented Reality to Improve Training and Operating Practices of EV's Battery Assembly

L.C. Moreira* and M. Kauffman*

Abstract— This paper outlines the results of an experimental research on deploying an emerging augmented reality (AR) system for real-time task assistance (or work instructions) of highly customised and high-risk manual operations. The focus is on human operators' training effectiveness and performance and the aim is to test if such technologies can support enhancing the knowledge retention levels and accuracy of task execution to improve health and safety (H&S). A new AR enhanced assembly method is proposed and experimentally tested using a real industrial process as case study for electric vehicles' (EV) battery module assembly. The experimental results revealed that the proposed method improved the training practices and performance through increases in the knowledge retention levels from 40% to 84%, and accuracy of task execution from 20% to 71%, when compared to the traditional paper-based method. The results of this research validate and demonstrate how emerging technologies are advancing the choice for manual, hybrid or fully automated processes by promoting the XR-assisted processes, and the connected worker (a vision for Industry 4 and 5.0), and supporting manufacturing become more resilient in times of constant market changes.

Keywords—Augmented reality, extended reality, connected worker, XR-assisted operator, manual assembly 4.0, industry 5.0, smart training, battery assembly

I. INTRODUCTION

After a decade of a stagnant productivity rate, the arrival of the 4th Industrial Revolution (I4.0) is expected to create up to USD3.7 trillion in value to global manufacturing by 2025 [1]. The novel generation of manufacturing facilities fostered by I4.0 widely adopts different technologies to digitalise manufacturing [2]. Manufacturing digitisation brings the goal of higher resilience that can be achieved by moving from a hierarchical structure to an integrated or interconnected structure [3]. Transforming manufacturing resilience could add £26bn of productivity value to the economy [4].

The current global scenario is represented by fast changes in customer requirements, which ultimately impact on products' complexity, volume of production and quicker time to market. The paradigm shift from mass production to mass customisation have greatly impacted on the complexity of products and processes [5, 6], but also in the workforce and

businesses as a whole; if businesses are not able to quickly respond to market changes, they can face severe consequences and lose competitive advantage.

In this scenario, manufacturing resilience represents a common goal amongst manufacturers, who often struggles to understand what it means and how to increase their resilience levels. Improving resilience in manufacturing is more challenging in the presence of manual operations due to the limitations in human adaptation and learning [7]. How best adapt manual operations poses unique challenges for manufacturers that seek for effective workforce (re)training and doing right-first-time solutions [8].

In this research, we aim to study how AR and computer vision technologies can enhance manufacturing resilience through optimising work instructions for manual assembly operations. For that, we propose an AR enhanced assembly (AREA) method to assist operators in both training and operations in a real industry application of EV battery module assembly as case study for testing and validation. The new method was developed using an XR-assistance technology, also called light guide systems or projection-based AR. Experimental studies involving 62 participants have been carried out for data collection. The results are analysed and compared with the traditional paper-based method.

This paper is organised as follows: Section 2 presents the research background. The research idea is presented in Section 3. The research methodology is provided in Section 4. Section 5 shows the case study details, followed by the results and analysis in Section 6. Finally, the conclusions are given in Section 7.

II. BACKGROUND

Mass customisation has emphasised the important role played by human operators and manual assembly processes in increasing flexibility of manufacturing systems at all functions [8]. In addition, manual operations are crucial for the social sustainability attained by the manufacturing sector [8,9]. Nevertheless, in a mass customised environment, human operators are overwhelmed by an enormous component variety and change in specifications and quality requirements, having to adapt their tasks from product to product with limited learning opportunities [10, 11]. As a result, leading to high stress levels, high absenteeism and turnover rates, which represent a major bottle neck to achieving manufacturing resilience.

Manufacturing companies are constantly confronting the challenges of competitive markets and lack of know-how of

*Authors' affiliation:
Institute for Advanced Manufacturing and Engineering (AME), Centre for Clean Growth and Future Mobility (CGFM)
Coventry University
Coventry, United Kingdom
Email: lorenacmoreira1@gmail.com

application of technology innovations within their manufacturing environments [7]. Implementing new technologies to improve manufacturing resilience has been challenging for several reasons, including fast speed of technological change, uncertainties and risks around the factory of the future vision (or lack of a strategic road map), and people barriers to technology adoption.

Based on this scenario, Figure 1 summarises key factors that can be controlled and targeted to improve manufacturing resilience levels. The factors have been split into: i) key challenges and impacts on *processes* (or manufacturing activities); ii) key indicators of manufacturing systems that *businesses* must seek to improve; and, iii) key challenges and impacts on the *people* (or workforce).

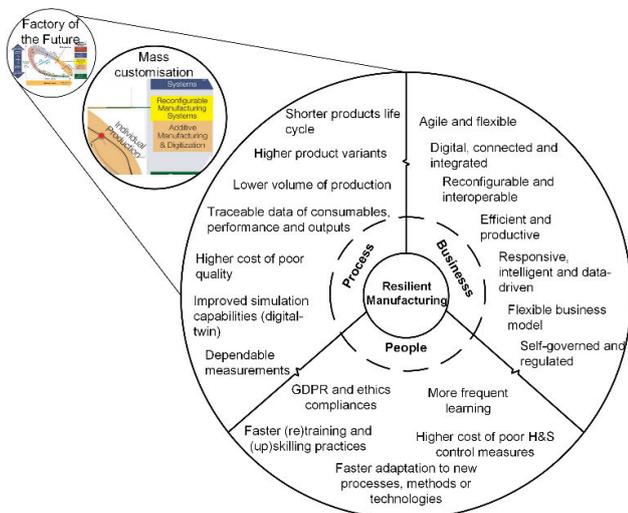


Fig.1 Impacts of mass customisation and the key factors on business, process and people to achieve resilient manufacturing.

This paper addresses the *process* and *people* challenges faced by a manufacturer of EV battery modules represented by the high costs of poor quality and high H&S risks due to the hazardous voltage, ranging from 60 volts to approximately 600 volts, during the manual assembly process. While a fully automated process could resolve those issues, there was no business case for an automated solution given the low volume and high number of product variants, resulting on the choice for a manual assembly operation. However, manual operations raised additional challenges due to health and safety (H&S) to the human operators and their vulnerability to errors, which could lead to delays, defects, poor product quality, and severe accidents in the workplace.

Given this scenario, the manufacturer was faced with two challenges:

- How to improve the effectiveness and H&S of the training and operating practices for battery modules assembly; and
- How to advance the work instructions method to support the worker/trainee in a safe and timely manner and reduce the process' vulnerability to human error.

Based on the aforementioned, this paper undertakes a people-centred approach to tackle the major *process* challenges to

ultimately, improve *business* resilience indicators. Consequently, the research question addressed in this paper can be defined as: how can emerging technologies help advancing (re)training and (up)skilling practices, and improve the performance and H&S of workers?

III. RESEARCH IDEA

Firstly, this paper proposes that extended reality (XR) such as AR, and other core emerging technologies such as computer vision can advance a new range of process configuration, so called, XR-assisted processes, to help addressing the trade-offs between level of automation and resilience gains (see Figure 2). In this new concept, manufacturing resilience has 4 main forces: Product Customisation indicator, Flexibility indicator, Productivity indicator and Quality indicator.

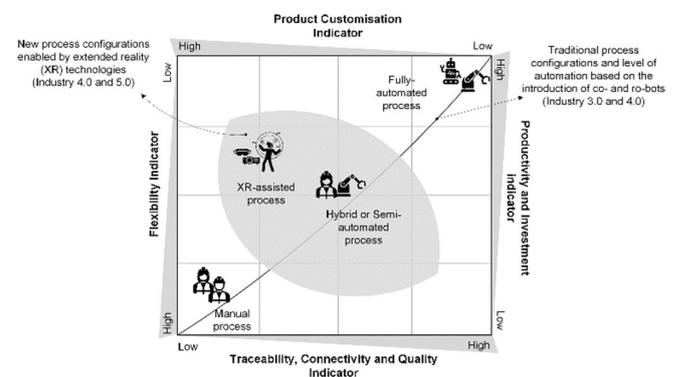


Fig. 2. Trade-offs between level of automation and key resilient manufacturing factors: introducing XR-assisted process.

The graph shown in Figure 2 illustrates that XR are transforming manual operations into XR-assisted hybrid processes, which can support in increasing key resilience indicators such as traceability, flexibility, productivity and quality without having a major impact on costs.

Therefore, the research question of this paper is addressed by proposing the use of AR and computer vision technologies (also known as light-guide or projection-based technology), to develop a new method that provides real-time assistance to the worker/trainee during the assembly of EV's battery modules.

The proposed AR enhanced assembly (AREA) method will provide real-time and interactive instructions to the worker/trainee by using features such as pick-by-light, process visualisation and real-time quality verification and validation [11, 12]. This way, offering a more advanced training experience compared to traditional methods such as paper-based work instructions. In addition, the AREA system is coupled with a flexible and user-friendly platform to create and reconfigure standard operating procedures with minimal effort.

The expected benefits of using the AREA method include to improve the training effectiveness and operator performance to reduce human error; this way, such method will promote improved capabilities impacting on higher productivity, responsiveness, agility and flexibility. The next section presents the methodology of this research.

IV. METHODOLOGY

This research developed an experimental study using a real-world manufacturing application as case study. A five-step approach was used (see Figure 3) to develop the research idea, and experiments were planned for data collection. The data was analysed using pre-defined rules and the results were discussed.



Fig.3 Five-step research approach.

Details of experimental plan and data analysis

An experimental plan is necessary to test the proposed AR method for its ability to improve the training practices with considering workers/trainees knowledge retention levels and accuracy of task execution. Consequently, the criteria and methods of analysis are stated as follows:

- Criteria 1: training effectiveness, which will be determined by the knowledge retained by the participants on the process steps after 24h from the training.
- Criteria 2: health and safety (H&S), determined by the number of errors made by the operator when performing the assembly process (or the accuracy of task execution). Experiments using the traditional paper-based method were carried out for comparison sake.

The experiments were performed with 62 participants, who were equally split into two groups, for data collection. Each group was assigned with one method of training: AR and paper-based. During the experiment, each participant was given 5 minutes to learn the assembly process using the respective training method. After 24hs from the training, the participants responded to a questionnaire with 10 questions related to the process steps of the assembly process. Finally, their responses were scored according to the following rules: A) Wrong answers = 0 points, B) Partially correct answers = 1 point, and C) Correct answers = 2 points. The total score of each participant was converted into a proportion considering the maximum points each participant could have achieved (i.e., 20 points). Finally, the results were analysed.

The following section presents the case study to test and validate the proposed approach.

V. CASE STUDY: AR-ASSISTED TRAINING FOR EV'S BATTERY MODULES ASSEMBLY

Problem Definition

The engineers of an automotive company are working on a new project to manufacture highly complex high-density battery packs for EVs. The products comprise a high variety of components and requirements. A fully automated solution is not financially viable and the current training and operating method using paper-based work instructions pose major challenges due

to its high-vulnerability to human errors, which can imply in severe H&S issues and high costs of poor quality. Consequently, there's a need for an advanced system to improve the training and operating capabilities of this manual assembly operation.

Selection of the technology

After screened several emerging technologies, an AR-assistance system comprised of projection-based AR, computer vision and a manufacturing operations management (MOM) software was selected. Its capabilities and features to provide real-time and interactive work instructions were appealing to minimise the vulnerability of the process to human error as well as to provide an advanced training experience for the workers/trainees.

The AR system works based on a set of process steps designed within the MOM software (a digital twin of work instructions). For that, the computer vision system uses object and motion recognition, based on RGB, infrared and depth cameras to train the MOM for all the parts and process requirements at each steps. Then, the projection-based AR feature is trained to provide real-time light-guided assistance such as picking/placing by light and further information such as written or visual instructions (e.g., text, image or video). Once the process is initialised (in training or operation mode), the augmented work instructions are given at each process step, while the computer vision captures the operator's actions in real-time and the MOM uses those inputs to trigger the next action autonomously. The system recognises when a mistake is made by the operator and a set of actions can be defined accordingly, e.g., stop the process until it's done correctly or call for help.

Development of the AR enhanced assembly (AREA) method for EV's battery modules

To develop the new AREA training and operating method using the selected technology, firstly, the standard operating procedures (SOP) or work instructions, which included 10 process steps, were developed. Secondly, the process steps were created in the MOM software, which required the SOP, the CAD parts for visualisation, and the requirements and tolerances. After that, the pick/put-by-light features and the quality and H&S verification and validation checks were defined. Finally, the AREA method was tested by expert engineers before running the experimental trials. Figure 4 illustrates the process to develop the AREA method.

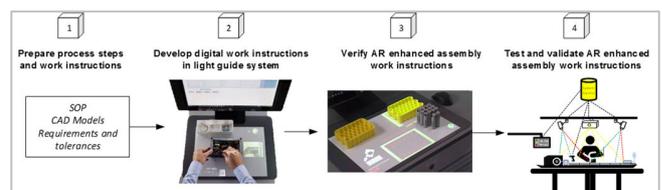


Fig.4 Development of the AR enhanced assembly method for battery modules.

Similar process was followed to develop the paper-based method, which replicates the same process steps but presents the work instructions in the traditional paper format.

The experimental trials were carried out based on the experimental plan provided in Section 4. The data collected is analysed and the results are presented in the next section.

VI. RESULTS AND ANALYSIS

The results from the experimental trials for the participants performance using the AREA and the paper-based methods are shown in Figures 5 and 6.

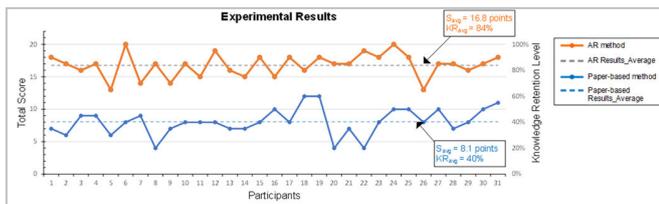


Fig.5 Experimental results for the two groups of participants based on AR and Paper-based methods, where S_{avg} and KR_{avg} stands for average of total score and average of knowledge retention level, respectively.

Figure 5 shows the total scores and knowledge retention levels of the participant. The results in this graph reveal that the participants who were trained using the AR enhanced assembly method achieved a better performance compared to the participants who received the paper-based training. Furthermore, the AREA method achieved an average of 84% (16.8 out of 20 points) of knowledge retention, more than the double of the average of the paper-based group (40%, 8.1 out of 20 points).

A further assessment was carried out considering the accuracy of task execution (Figure 6), where the participants received scores depending on their performance at each process step. The graph in Figure 6 shows a histogram with the responses from the groups trained using the AREA and paper-based methods.

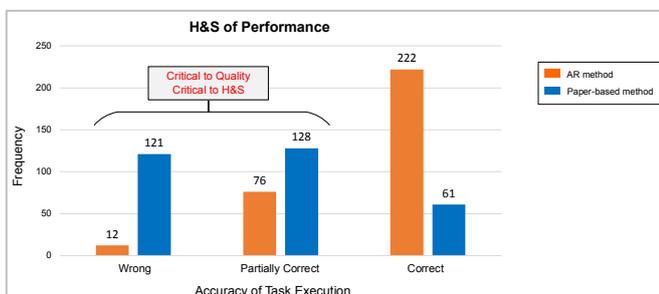


Fig.6 Experimental results based on accuracy of task execution and H&S.

The experimental results show that 80% of the tasks performed by the participants trained using the paper-based

method presented mistakes, which 40% (121 tasks) were performed wrongly. Whereas the results from the AREA method show that 70% of the tasks were performed correctly, and only 4% (12 tasks) were totally wrong.

Given the scenario where operational mistakes will lead to high costs of poor quality and high H&S risks, the experimental results of this research demonstrate the importance of deploying emerging technologies such as XR to develop more effective and robust training and operating practices. As a result, processes are less vulnerable to human error, the workforce is supported to cope with the challenges of mass customised production environments and businesses improve their resilience levels.

VII. CONCLUSION

This research work stemmed from the need to support an automotive manufacturing supplier working on a new battery for electric vehicles (EVs). Consequently, this research had a direct impact on a manufacturing company by providing the evidence to develop a solution that will enhance their training and operating practices, and ultimately, improve their resilience levels.

The results of this research validate crucial benefits of deploying augmented reality systems to support companies in becoming more resilient through improved training and operating practices of manual assembly operations.

The proposed AR enhanced assembly (AREA) method improved the training outputs of EV's battery modules assembly by increasing the knowledge retention levels from 40% to 84%, and by increasing the accuracy of task execution from 20% to 71%. Such improvements are critical for processes with high number of product variants, high costs of poor quality and high H&S risks, such as mass customised manufacturing environments. The AREA method results supported minimising human errors and, consequently, maximising quality, productivity and workers' H&S during manual assembly processes. Ultimately, providing positive impacts on business productivity, agility and flexibility.

The results of this research demonstrate how AR technologies are advancing the choice for manual, hybrid or fully automated processes when there isn't a financial viability of implementing full automation. The main findings can be used as recommendation for manufacturers to re-think their (re)training and (up)skilling practices and manual operations capabilities.

Although it is out of scope of this paper, the proposed solution using the AR system also demonstrated great potential to improve process traceability, connectivity and workers wellbeing. Such aspects will be considered in future research. We can conclude that given the current challenges faced by manufacturers, such emerging technologies bring great capabilities to overcome the challenges of human factors, and businesses can therefore become more resilient to both internal and external changes, yet securing jobs.

REFERENCES

- [1] Schmitz, C., Tschiesner, A., Jansen, C., Hallerstedde, S. and Garms, F., 2019. Industry 4.0: Capturing value at scale in discrete manufacturing. In Technical report. McKinsey & Company.
- [2] Moreira, L.C., Li, W.D., Lu, X. and Fitzpatrick, M.E., 2019. Supervision controller for real-time surface quality assurance in CNC machining using artificial intelligence. *Computers & Industrial Engineering*, 127, pp.158-168.
- [3] Hankel, M. and Rexroth, B., 2015. The reference architectural model industrie 4.0 (rami 4.0). *ZVEI*, 2(2), pp.4-9.
- [4] (Online) RS Components: available at < https://uk.rs-online.com/web/generalDisplay.html?id=footer1/release/210308_resilienceindex_uk> Accessed in April 2021.
- [5] Medini, K., Andersen, A.L., Wuest, T., Christensen, B., Wiesner, S., Romero, D., Liu, A. and Tao, F., 2019. Highlights in customer-driven operations management research. *Procedia Cirp*, 86, pp.12-19.
- [6] Oh, J.W., Na, H. and Choi, H., 2017. Technology Trend of the additive Manufacturing (AM). *Journal of Korean Powder Metallurgy Institute*, 24(6), pp.494-507.
- [7] Ganji, E.N., Shah, S. and Coutroubis, A., 2018, June. Mass Customization and Technology Functions: Enhancing Market Responsiveness. In 2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC) (pp. 1-6). IEEE.
- [8] Romero, D., Stahre, J. and Taisch, M., 2020. The Operator 4.0: Towards socially sustainable factories of the future.
- [9] Breque, M., De Nul, L. and Petridis, A., 2021. Industry 5.0: towards a sustainable, human-centric and resilient European industry. Luxembourg, LU: European Commission, Directorate-General for Research and Innovation.
- [10] Pilati, F., Faccio, M., Gamberi, M. and Regattieri, A., 2020. Learning manual assembly through real-time motion capture for operator training with augmented reality. *Procedia Manufacturing*, 45, pp.189-195.
- [11] Torres, Y., Nadeau, S. and Landau, K., 2021. Classification and quantification of human error in manufacturing: a case study in complex manual assembly. *Applied Sciences*, 11(2), p.749.
- [12] de Souza Cardoso, L.F., Mariano, F.C.M.Q. and Zorzal, E.R., 2020. A survey of industrial augmented reality. *Computers & Industrial Engineering*, 139, p.106159.

Dr Lorena C. Moreira works for the Institute for Advanced Manufacturing and Engineering, Coventry University, UK. She has a BSc, MRes and PhD in Manufacturing and Systems Engineering, and her areas of interest are social and environmental sustainability and how emerging technologies such as AI, XR and computer vision can support developing intelligent systems for multiple manufacturing KPIs, including carbon footprint, worker wellbeing and productivity.

Overview of Adaptive Spline Interpolation

Rongli Gai, Zhiyuan Chang, Xiaohong Wang, Jingyu Liu

Abstract—In view of various situations in the interpolation process, most researchers use self-adaptation to adjust the interpolation process, which is also one of the current and future research hotspots in the field of CNC(Computerized Numerical Control) machining. In the interpolation process, according to the overview of the spline curve interpolation algorithm, the adaptive analysis is carried out from the factors affecting the interpolation process. The adaptive operation is reflected in various aspects, such as speed, parameters, errors, nodes, feed rates, random period, sensitive point, step size, curvature, adaptive segmentation, adaptive optimization, etc. This paper will analyze and summarize the research of adaptive imputation in the direction of the above factors affecting imputation.

Keywords—Adaptive algorithm, CNC machining, interpolation constraints, spline curve interpolation.

I. INTRODUCTION

INTERPOLATION is the basis for generating tool paths in CNC machine tool processing. The interpolation accuracy directly affects the processing quality of the CNC system, and the interpolation speed directly determines the processing efficiency of the CNC machine tool. Therefore, the interpolation algorithm is the core of the entire CNC system. Because the sudden situation in the interpolation process will affect the processing process, and the manual intervention will make the processing efficiency low, the adaptive interpolation algorithm has become a current research hotspot.

The adaptive interpolation algorithm [1]-[3] is the main practical method to deal with unexpected situations in the interpolation process. In the interpolation process, there are many factors that affect the interpolation. Due to the different emphasis, the adaptive operation is mainly carried out in the main aspects such as speed, parameters, error, and curvature. With the in-depth exploration of researchers, other influencing factors risk correction is also possible through adaptive algorithms. Through the analysis of the existing adaptive interpolation algorithms, this paper summarizes and compares the interpolation algorithms of various adaptive types, shows the advantages and disadvantages of various methods, and then points out the research and development trend of adaptive interpolation algorithms.

II. OVERVIEW OF SPLINE INTERPOLATION ALGORITHM

A. Spline Curve

There are many definitions of splines. It is generally believed that splines are connected by a series of k-order polynomial

curves at specific data points. Each data point connects two polynomial curves. It is continuous, and it is the first and second derivatives are also continuous [4]-[8]. Splines are defined piecewise, and the shape of the curve is generally determined by a set of control points or data points [9], [10].

Taking the cubic spline curve as an example, on the parameter interval $[u_i, u_{i+1}]$, the general cubic spline curve equation $S(u)$ can be expressed as [11]:

$$S(u) = A_i + B_i(u - u_i) + C_i(u - u_i)^2 + D_i(u - u_i)^3 \quad (1)$$

In the formula, u is the parameter of the spline curve, A_i, B_i, C_i, D_i is the coefficient matrix of the spline curve, and its dimension is equal to the dimension of the processed curve. When calculating the value of A_i, B_i, C_i, D_i , in addition to using the coordinate values corresponding to u_i and u_{i+1} , it is generally calculated by the continuous characteristics of the first-order and second-order derivatives at the endpoints of the spline curve [12].

In the field of CNC machining, commonly used spline curves include Bezier curve, B-spline curve and NURBS(Non-Uniform Rational B-Splines) curve. Because the shape control function of NURBS curve is particularly powerful and flexible, and it can accurately represent analytical curves and free-form surfaces, NURBS curves are set by the International Organization for Standardization (ISO) as the data exchange standard for CAD(Computer Aided Design) and CAM(Computer Aided Manufacturing). The NURBS curve is developed on the basis of the Bezier curve and the ordinary B-spline curve, and is well compatible with the two parametric curves, and has now become a unified curve expression.

Assuming that there are $n + 1$ data points $P_0, P_1, P_2 \dots P_n$, a NURBS curve of order k is defined as:

$$C(u) = \frac{\sum_{i=0}^n w_i P_i N_{i,k}(u)}{\sum_{i=0}^n w_i N_{i,k}(u)} \quad (2)$$

In (2), u is the parameter of the NURBS equation, $P_i (i = 0, 1 \dots n)$ is the data point or control point of the NURBS curve, and w_i is the weight corresponding to the control point

Rongli Gai, is an associate professor, works in the school of information of Dalian University and is engaged in the research of machine learning and motion control algorithms, Dalian, China (corresponding author, phone: 136-1095-9380; e-mail: gairli@sict.ac.cn).

Zhiyuan Chang, is a graduate student studying at Dalian University (e-mail: 18263468648@163.com).

P_i , where $w_0 > 0, w_n > 0$, and other $w_i \geq 0$; $N_{i,k}(u)$ is the basis function of the k-order NURBS curve, determined by the node vector $U = (u_i, u_{i+1}, \dots, u_{i+k})$ the basis function satisfies the Cox-de Boor recurrence:

$$\begin{cases} N_{i,1}(u) = \begin{cases} 1 & u_i \leq x < u_{i+1} \\ 0 & \text{Otherwise} \end{cases} \\ N_{i,k}(u) = \frac{u - u_i}{u_{i+k} - u_i} N_{i,k-1}(u) + \frac{u_{i+k+1} - u}{u_{i+k+1} - u_{i+1}} N_{i+1,k-1}(u) \end{cases} \quad (3)$$

Here you need to define $\frac{0}{0} = 0$.

B. Spline Curve Interpolation

A spline curve [13] refers to a curve formed by connecting polynomial curve segments. The boundary of each segment satisfies a specific continuous condition, and its shape is generally determined by a set of control points. Because the spline curve can accurately and uniformly represent the analytical curve and the free curve, it is stipulated by the International Organization for Standardization as the data exchange standard of CAD/CAM. According to STEP-NC (the numerical control system data standard ISO14649)[14]-[16], the spline curve interpolation makes the processing information of the three-dimensional geometric model conforming to the STEP standard directly as the input of the numerical control system, that is, the parameter curve is directly interpolated.

In spline interpolation, chord is usually used to approach arc, and its principle is shown in Fig. 1.

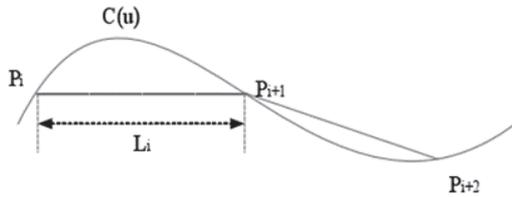


Fig. 1 Schematic diagram of spline interpolation

In spline interpolation, chords are usually used to approximate arcs [17]. $C(u)$ is the parameter curve, u is the interpolation parameter, and L_i is the interpolation step. Suppose the current interpolation point is $P_i = C(u_i)$, the next interpolation point is $P_{i+1} = C(u_{i+1})$, the ideal instantaneous speed of the tool at point P_i is $V_{ds}(u_i)$, and the interpolation period is T , according to the geometric relationship:

$$L_i = V_{ds}(u_i) \cdot T = \|C(u_{i+1}) - C(u_i)\| = \|P_{i+1} - P_i\| \quad (4)$$

The arc length l_i from parameter u_i to u_{i+1} is:

$$l_i = \int_{u_i}^{u_{i+1}} \|C'(u)\| du \quad (5)$$

The arc length is approximated by the chord length, ideally $L_i = l_i$, that is

$$V_{ds}(u_i) \cdot T = \int_{u_i}^{u_{i+1}} \|C'(u)\| du \quad (6)$$

Due to the nonlinear relationship between arc lengths and parameters of NURBS curves and B-spline curves, it is difficult to calculate the exact value of G during real-time spline interpolation, and only approximate values can be obtained. This approximation will cause the difference between the theoretical feed-rate and the actual feed-rate during spline interpolation, that is, feed-rate fluctuations.

After calculating the next parameter u_{i+1} , the actual feed rate V_i is:

$$V_i = \frac{|C(u_{i+1}) - C(u_i)|}{T} \quad (7)$$

According to the above feed information [18], the next interpolation parameter u_{i+1} can be obtained, and then the actual interpolation point P_{i+1} can be obtained. Due to the nonlinear relationship between the arc length and the parameters of the spline curve, the exact value of u_{i+1} often cannot be obtained during interpolation, but only approximate values can be obtained. This results in a discrepancy between the ideal feed-rate and the actual feed-rate, that is, feed-rate fluctuations. The fluctuation of the feed rate will not only lead to the reduction of the machining accuracy of the workpiece, but may even cause chattering and affect the machining quality. The size of the speed fluctuation can be measured by the speed volatility δ_i [19]:

$$\delta_i = \frac{\Delta V_i}{V_{ds}(u_i)} \times 100\% = \left(1 - \frac{\|C(u_{i+1}) - C(u_i)\|}{V_{ds}(u_i)T}\right) \times 100\% \quad (8)$$

Through the understanding of spline curve interpolation, it can be found that there are many factors affecting the interpolation process, and there are many directions worthy of in-depth study on the point of adaptation [20],[21].

III. ADAPTIVE INTERPOLATION METHOD

Common adaptive interpolation methods mainly include speed look-ahead adaptation, parameter adaptation, error adaptation, and other direction adaptations. These adaptive methods are analyzed below.

A. Speed Forward-Looking Adaptive

The speed look-ahead control [22], [23] is to realize the preprocessing of the program by monitoring the machining trajectory, so as to carry out the acceleration analysis and deceleration area discrimination of each linkage axis in advance,

and realize the smooth transition at the turning point of the block or the deceleration point, to ensure excellent acceleration and deceleration control to avoid machine impact and ensure processing quality.

a. Speed Calculation of Continuous Micro Program Segment

Because of its simple expression, small amount of calculation and strong applicability, small segment path is the most widely expressed form of machining path of complex curve and surface parts such as impeller, blade and mold. With the rapid development of high-speed and high-precision machining technology, the machining process of complex curved surface parts takes milling instead of grinding as the development goal, and puts forward higher requirements for micro segment path machining.

Zhengjie et al. [24] inserted the bisector of the transition angle at the corner of the line segment for transition, adaptively divided the number of prospective segments, and planned forward and reverse speeds to make the speed transition between micro-line segments smooth and further improve the surface quality of the workpiece. Haihong et al. [25], [26] adopted an adaptive look-ahead control algorithm that optimizes the connection speed between trajectory segments for the current situation that most of the actual machining parts trajectories are composed of straight lines and circular arc trajectories. By judging and adjusting the accessibility, the speed between adjacent track segments can be smoothly and quickly connected, reducing the frequent start and stop of the feed axis in a large number of tiny track segments, and improving the acceleration and deceleration efficiency and machining accuracy.

Bolin et al. [27] segmented the curve by finding the breakpoints of G^0 (The position at the tool path connection point is continuous) continuous, G^1 (At the connection of line segments, two adjacent tool paths have a common tangent direction) discontinuous and G^1 continuous, G^2 (At the connection of line segments, two adjacent tool paths have common curvature) discontinuous and key points with large curvature. However, there may be speed trajectory planning intersection between the current interpolation adjacent sensitive point area and the next interpolation adjacent sensitive point area. Yongqiao [28] used a G^2 continuous B spline global smoothing algorithm interpolation technology, and realized the real-time smoothing and corner speed optimization of continuous small line segments in the forward-looking window based on the local smoothing of continuous small line segments and the optimization of corner speed based on the distribution of corner errors. and cross-segment adaptive speed planning. Jialan et al. [29] adopted the Cardinal spline curve transition method to establish a corner curve transition vector model to determine the corner transition curve that satisfies the conditions and then determine the maximum transfer speed to achieve high-speed connection between micro-segments.

Based on the S-type acceleration and deceleration scheme, Ren et al. [30] found the high-curvature points of the discrete

machining path through cubic spline curve fitting and estimated the optimal speed at the high-curvature points, and controlled the machining path in blocks, effectively. It solves the problem of overcut easily occurring at high curvature points in traditional machining. Guangqiang et al. [31] performed pre-deceleration processing before reaching the area with large curvature or sudden change of curvature, and increased the speed planning to ensure the machining accuracy under the condition that the pre-deceleration does not exceed the acceleration limit.

With the development of numerical control technology, the linear interpolation technology based on small line segment path has been quite mature, but with the rapid development of high-speed machining, the traditional interpolation technology of small line segment can no longer meet the needs of high-speed machining. The higher the surface machining accuracy requirements, the shorter the line segments generated by the tool path, each of which is at the millimeter level or lower, which seriously restricts the improvement of the machining speed and makes the development of high-speed interpolation technology for micro-line segment paths facing a series of problems:

- 1) Linear high-order continuous smoothing problem: The tangent direction at the corners of adjacent line segments may suddenly change, and the curvature is discontinuous, resulting in sudden changes in acceleration at the corners and causing machine tool vibration.
- 2) Matching motion constraints of each axis of CNC machine tools: Only using the maximum speed and acceleration constraints cannot fully describe the performance characteristics of the machine tool.

b. Speed Optimization at the Transfer of Adjacent Program Segments

The first-order discontinuity at the corner of the linear tool path leads to the fluctuation of the feed rate and the sudden change of the acceleration at the corner. The look-ahead function was first used to solve this dilemma. By pre-reading a certain number of program segments, the micro-line segment path is analyzed and processed in advance, and then the point of the speed that needs to be processed is determined to maximize the connection speed of the line segments, thereby increasing the feed speed.

Yingqi et al. [32] used the Simpson adaptive integration method to calculate the length of the curve and adaptively segment the curve. Although this method improves the feed rate and reduces the processing time, it ignores the constraint of the rotation angle on the speed. Wei et al. [33] used backtracking and re-interpolation to predict at the speed-sensitive point and determine the deceleration point in this area, and then replace the previous adaptive storage speed with the re-interpolation speed in the curvature-sensitive area. By differentiating the ADAMS differential equation, Renping et al. [34], [35] estimated the correction iteration accuracy and retrospectively reconstructed the speed of the interpolation point, accurately predicted the position of the deceleration point, and made the feed of each speed sensitive point. The speed is maximized as

far as possible under the premise of satisfying the machining accuracy. Haixia et al. [36] proposed an adaptive interpolation algorithm (S-A-NURBS) for symmetric combined sinusoidal function acceleration and deceleration control, which can effectively solve the problem of starting and stopping at the inflection point, smooth the transfer speed of adjacent sections, and reduce losses.

The speed optimization and smoothing processing at the transition of adjacent program segments can be divided into two types: the accurate passing method of the path connection point and the smooth transition method of the path connection point.

1) The accurate passing method of the path connection point:

This method refers to accurately passing through each given path connection point during forward-looking speed smoothing and interpolation. The advantage of this method is that the interpolation accuracy is high and the interpolation error is not introduced. The disadvantage is that when planning the speed of the path segment, because the beginning and end speed of the optimized path segment is not zero, it is impossible to complete the interpolation of the path segment according to the expected acceleration and deceleration strategy in an integer interpolation cycle. To solve this problem, Wang et al. [37] presented an integer optimization model and solved it by a simplified method. Literature [38], [39] gives two strategies to adjust the feed speed curve for linear acceleration and deceleration, so that the adjusted path segment margin is exactly equal to the path increment of one interpolation cycle, so that the path segment can complete interpolation in an integral number of interpolation cycles. Li et al. [40] proposed a variable period interpolation method. According to the margin of less than one interpolation period at the end of the path segment, the interpolation period is adjusted accordingly. The variable interpolation period is provided by the RTX(Real Time Executive) real-time subsystem [41]. However, due to the influence of the timer precision, the feed rate at the end of the path segment is prone to fluctuations. In addition, this method is limited by the length of the path segment. When the length of the path segment is short (such as several microns to tens of microns), the connection speed of the path segment will be limited to a lower value (related to the interpolation period, the interpolation period The longer it is, the lower the connection speed is limited), so the actual feed speed is still difficult to increase sufficiently.

2) The smooth transition method of the path connection point:

This method performs local fairing on the pre-reading path segment to realize cross segment continuous interpolation. It does not need to accurately pass through the given path connection point, which can overcome the problem that the interpolation period is not an integer, and can also avoid the restriction of the path segment length on the feed speed. There are two key points in this method. One is to use parametric curves to locally smooth the pre-read path segments. Commonly used parametric curves are Bezier curve [42], [43], PH(Pythagorean-Hodograph) curve [40], [44], and B Spline [45] etc. The second is to optimize the

connection speed of the path segment after partial smoothing according to the allowable geometric error and machine tool characteristics, and plan the speed according to the desired acceleration and deceleration strategy. Yau et al. [46], [47] optimized the joint speed according to the corner error caused by the servo lag. Tsai et al. [48], Zhang et al. [489 and Ye et al. [50] optimized the bridging speed according to the maximum allowed acceleration and path curvature. Ping et al. [51] based on the given maximum geometric error, the smoothing error and the interpolation error are evenly distributed, and the connection speed is optimized.

c. Forecast before Deceleration Point

The acceleration part of the interpolation process runs according to the normal acceleration law, even if it does not reach the maximum allowable speed, it will not affect the machining accuracy, but for the deceleration process, the deceleration must be monitored. If the speed at the speed sensitive point cannot be reduced in time as required, the machining accuracy will be affected. In the acceleration and deceleration planning, the inaccurate prediction of the deceleration point has always been a major difficulty in the acceleration and deceleration control planning.

Guoliang [52] introduced the concept of forward and reverse bidirectional interpolation to solve the problem that the deceleration starting point and the length of the curve cannot be accurately obtained during the interpolation process. Hepeng [53] proposed a bidirectional interpolation strategy based on the speed planning method of time bidirectional rounding. Lei et al. [54] performed reverse interpolation for the determined velocity abrupt point within the prospective range by using the local optimal solution of the abrupt point velocity based on the bow height error and cubic polynomial acceleration and deceleration. Liu Nian et al. [55],[56] iterated the NURBS parameters through Steffensen algorithm to realize the adaptive planning of speed. On this basis, Dong [57] adopted the interpolation algorithm calculated by Steffensen type parameters, and adaptively adjusted the maximum jerk according to the curvature information to implement high-precision interpolation to reduce the velocity fluctuation rate. Xia et al. [58] used NURBS symmetry to predict the deceleration point, combined with real-time step size adaptive control to estimate the interpolation parameters by the simplified Adams differential equation algorithm, and finally used the prediction-correction to control the precision.

d. Other Types of Speed Look-Ahead Control

After years of hard work by many scholars at home and abroad, the research on adaptive curve interpolation has made great progress. Adaptive interpolation methods to meet different machining conditions have been proposed successively, and each adaptive interpolation algorithm has its own unique consideration for the determination of the feed rate.

Peng et al. [59], [60], based on the dual-core architecture system, studied and analyzed the optimization principle and linear assembly-level optimization of the IQ Math library, and

optimized the adaptive forward-looking algorithm. significance. Jianfeng et al. [61] imposed constraints on geometric accuracy, machine tool drive characteristics and cutting characteristics, and proposed a customized FSLP(Feed Speed Customization Algorithm Based on Linear Programming) algorithm for speed planning, which could effectively control the swing speed of the tool shaft and limit the change in swing speed, greatly improving the surface quality of side milling. On the basis of RTCP(Rotational Tool Center Point) linear encryption, Haitao [62] established a calculation model of limit feed rate under nonlinear error constraints, and established an adaptive feed rate prediction model under nonlinear error-machine tool dynamics and thermal properties-joint speed constraints. The interpolation method effectively reduces the vibration and impact of the machine tool.

Adaptive speed forward-looking involves knowledge in many fields, such as complex tool pose planning, machine tool kinematics and dynamics, machining error control, high-speed and high-precision interpolation algorithm and intelligent algorithm optimization. Although researchers have made some research achievements in the direction of adaptive interpolation, due to limited ability and energy, there are still many contents worthy of further research and discussion in the research field of optimal feed speed planning of NC machining.

TABLE I
COMPARISON OF SPEED LOOK-AHEAD ADAPTIVE METHODS

Method	Advantage	Disadvantage
Speed calculation of continuous Micro Program Segment	Simple expression, small amount of calculation, strong applicability	The precision of the joint is not enough, resulting in linear high-order continuous smoothing problem
Speed optimization at the transfer of adjacent program segments	Increased feed rate, reduced machining time, and high interpolation accuracy	The acceleration between segments is prone to sudden change, and the feed rate is unstable
Forecast before deceleration point	High estimation accuracy, fast convergence speed, and reduced calculation amount	The calculation process is complicated and the algorithm operation time is long
Other types of speed look-ahead control	low speed fluctuations	The stability is poor, the calculation amount of acceleration and deceleration control is too large, and most of them are limited to the given tool path

The comparison of various methods in the speed look-ahead adaptation is shown in Table I.

B. Parameter Adaptation

The concept of adaptive parametric curve interpolation has greatly enriched the connotation of feed rate control in parametric curve interpolation. Researchers can establish different feed rate planning models considering the required constraints according to different working conditions, and finally get feed speed profile curve to meet different requirements.

Zhimei [63], [64] put forward the concept of parameter intermediate point, using the distance from the parameter

intermediate point to the interpolation straight line as the interpolation error, and adjusting the interpolation error in real time by changing the parameter increment. Shiming [65] used a parameter adaptive algorithm to optimize the cubic NURBS curve to determine the node vector, and then based on the chord-cut iteration method to iterate the predicted nodes to ensure the normal range of volatility. Zhiwei [66] proposed an adaptive determination method for small line segment velocity planning parameters based on statistics to solve the problem that it is difficult to determine the optimal machining parameters for CNC systems to process different workpieces. The Gear estimation correction interpolation algorithm of Hengjun et al. [67] obtains the estimated step size through parameter value estimation and the preprocessing matrix method, and then uses the adaptive ideal step size to iteratively correct the parameter values. It solves the shortcomings of poor real-time performance and large speed fluctuations. Zhibing [68] used the improved Admas differential equation method to estimate the parameters, and realized the adaptive correction of the parameters according to the constraints of chord height error, velocity and acceleration.

C. Error Adaptation

After the real-time error estimation of the contour is obtained, the estimated contour error needs to be used for contour control. There are currently two main classes of methods for contour control: cross-coupling controller (CCC) [69]-[71] and task coordinate method [71]-[74]. The basic principle of the cross-coupling controller is: taking the ideal contour information of each axis and the actual machining position information as the comparative data, the real-time contour error is estimated by using the contour error estimation algorithm for this group of data, and then calculated by the cross-coupling controller, and decoupled to each axis with the cross coupling coefficient, so as to control the contour error as a whole. The principle of the task coordinate method is: the estimated contour error is decomposed along the normal and tangential directions at the desired position point or specified point, and then the control laws for these two variables can be independently designed. Since the normal component is often considered to be approximately equal to the contour error, the control law bandwidth for the normal error is higher than that for the tangential error when designing the controller.

Zhiwei [75] combined the contour error adaptive interpolation algorithm with the path length adaptive speed planning algorithm. Zuhao [76] combined the advantages of equal parameter method and equal error method, established the adaptive error model of trisection method, and put forward the optimized error adaptive arc interpolation algorithm to obtain high machining accuracy. Yitian [77] proposed a real-time adaptive contour error estimation method (ACEE), which made the number of error estimation points more than the number of servo control points, and improved the accuracy of contour error estimation. The contour error estimation model proposed by Yang [78], [79] combined the dual fuzzy variable universe adaptive control algorithm to design the contour error controller of the contour optimal circle approximation method, which

specifically avoided the estimation dependent on the tracking error. This method can effectively reduce the machining contour error. Based on the existing contour error calculation model, Henan [80] proposed a new contour error calculation model—the calculation model of four-point double circle weighted approximation, The model also does not depend on the tracking error value and contour shape, and the effect of contour control for large curvature is more obvious.

D. Other Direction Adaptations

Yiqiao [81] proposed a cubic spline interpolation model for node adaptive selection in the problem of the change of error data. Aiming at the problem of redundant interpolation and large approximation error in the equidistant parameterization method, Zhen [82] designed an adaptive interpolation algorithm based on NURBS curve contour curvature radius, and Qian [83] solved the equidistant curve by adaptive discrete method Self-inflicted problem. Hui [84] proposed a method to find the optimal interpolation period, and proposed an adaptive random period interpolation control method from the perspective of changing the interpolation period, which provided a new method for eliminating the frequency superposition problem in the case of a large interpolation period. way. Jipeng [85] proposed a NURBS curve adaptive fitting algorithm---FKTP(Feature Point Recognition and The Knot Placement Technique) for dense point sequences in space, and then based on FKTP performed a speed adaptive adjustment strategy for targeted two-way rendezvous to ensure the continuity of the left and right motion parameters of the rendezvous point. Shujie et al. [86] adaptively select the interpolation algorithm based on the complexity of the interpolation calculation for different interpolation accuracy and curve-related parameters, and dynamically select an efficient calculation method.

TABLE II
COMPARISON OF ADAPTIVE SPLINE INTERPOLATION METHODS

Method	Advantage	Disadvantage
Speed forward-looking adaptive	Low speed fluctuation and strong applicability	Mutation is easy to occur at the high-order continuity
Parameter adaptation	Strong real-time performance and strong convergence	There are high requirements for determining the first few parameters, error compensation is required, and the algorithm has limitations
Error adaptation	High machining accuracy and strong stability	Large amount of calculation, poor real-time performance, poor effect at high speed and large curvature
Other direction adaptation	The degree of time optimization is strong, and the convergence speed is accelerated.	The applicable direction is limited and the technology is not perfect

The comparison of various adaptive imputation methods is shown in Table II.

IV. ADAPTIVE DEVELOPMENT IN NEW DIRECTION

Hu et al. [87] used a non-dominated genetic algorithm, namely NSGA-II, based on the constraints on the tangent vector and control points, without knowing the nodes in advance, and

the obtained interpolation curve is similar to the given tangent vector and data polyline, which is better than other method is more natural. In the research of robot trajectory planning, Wei [88] analyzed the basic trajectory planning algorithms in different trajectory planning spaces, and proposed a 4-5-4 polynomial hybrid interpolation algorithm based on this trajectory planning algorithm. Aiming at the problem that the standard PSO optimization algorithm is easy to fall into the local optimum and difficult to obtain the global optimum, this interpolation algorithm proposes an improved scheme that combines the standard PSO optimization algorithm with the natural selection mechanism and the adaptive inertia weighting factor. Based on the above two schemes, he proposed a scheme combining the improved PSO optimization algorithm and the 4-5-4 polynomial hybrid interpolation algorithm to realize the optimal trajectory planning of joint point to point.

V. CONCLUSION AND PROSPECT

In summary, adaptive spline interpolation has become one of the research hotspots in the field of CNC machining. In the computer numerical control system, the interpolation algorithm is the basis for generating the machining trajectory, and its pros and cons largely determine the machining efficiency and precision of the numerical control system. Based on various factors affecting the interpolation in the process of interpolation, this paper summarizes various adaptive interpolation methods, and expounds the characteristics of each method. It can be seen from the analysis that the adaptive spline curve interpolation currently has the following trends:

First, the adaptive interpolation method for the purpose of speed look-ahead is the mainstream of research. Since the feed rate is constrained and restricted by many factors, the complementary self-adaptation of many restrictive factors can be studied in depth in the horizontal direction, and the research direction of multi-axis and multi-dimensional space self-adaptive interpolation can be approached in the vertical direction.

Second, we can carry out divergent innovation research according to the factors affecting adaptive interpolation by innovating and improving the basic kinematics basis and kinematics related models.

Third, it can be cross applied with algorithms in robot trajectory planning or research methods in other fields, such as the combination with machine learning and neural network. In recent years, the cross application of knowledge between subject directions is frequent. We can find suitable adaptive interpolation methods in experiments and explore new research directions.

REFERENCES

- [1] Lin Wang, Jianfu Cao, "A look-ahead and adaptive speed control algorithm for high-speed CNC equipment," *The International Journal of Advanced Manufacturing Technology*, vol. 63, pp. 705-717, 2012.
- [2] Zhang Lixian, Sun Ruiyong, Gao Xiaoshan, et al, "High speed interpolation for micro-line trajectory and adaptive real-time look-ahead scheme in CNC machining," *Science China Technological Sciences*, Zhang, Lixian et al. "High speed interpolation for micro-line trajectory and adaptive real-time look-ahead scheme in CNC machining." *Science*

- China Technological Sciences, vol. 54, pp. 1481-1495, 2011.
- [3] Haina, Zhong et al, "Velocity control method of 5-axis simultaneous computerized numerical control system," *Computer Integrated Manufacturing Systems*, 2007.
 - [4] Yunan Cao, Tianmiao Wang, Youdong Chen, Hongxing Wei, "Application of S acceleration and deceleration before interpolation in CNC foresight," *Journal of Beijing University of Aeronautics and Astronautics*, pp. 594-599, 2007.
 - [5] Heng Ye, Minghong Xie, "Adaptive forward planning algorithm for continuous micro-line segment flexible acceleration and deceleration," *Journal of Huaqiao University (Natural Science Edition)*, vol. 42, no. 1, pp. 1-8, 2021.
 - [6] Dierckx, Paul, "Curve and surface fitting with splines," *Monographs on numerical analysis*, 1996.
 - [7] Späth, Helmuth, "One-dimensional spline interpolation algorithms," 1995.
 - [8] Späth, Helmuth, "Two-dimensional spline interpolation algorithms," 1993.
 - [9] Salomon, David, "The Computer Graphics Manual," *Texts in Computer Science*, 2011.
 - [10] Piegl, Les A. and Wayne Tiller, "The NURBS book (2nd ed.)," 1997.
 - [11] Fazhong Shi, "Computer-aided geometric design and non-uniform rational B-splines: CAGD & NURBS," *Beijing University of Aeronautics and Astronautics Press*, 1994.
 - [12] Yu Shilin, Zhu Guobao, "Mathematical Processing of List Curves in CNC Machining Programming," *Journal of Wuhan Jiaotong University*, 2000.
 - [13] Zhang Xiaohui et al, "Research on Smooth Compression Interpolation Algorithm in NC Machining," *Chinese Journal of Mechanical Engineering*, vol. 47, no. 5, pp. 156-162+169, 2011.
 - [14] International Standards Organization. 15014649. Part1: Overview and fundamental PrineiPles, 2003.
 - [15] Suh, S. H. et al, "Architecture and implementation of a shop-floor programming system for STEP-compliant CNC," *Comput. Aided Des*, vol. 35, pp. 1069-1083, 2003.
 - [16] Shin, Seungjun et al, "Reincarnation of G-code based part programs into STEP-NC for turning applications," *Comput. Aided Des*, vol. 39, pp. 1-16, 2007.
 - [17] Feibiao Pang, "Research on high-quality machining technology for complex contour and curve parts," *Xi'an University of Architecture and Technology*, 2020.
 - [18] Yanwei Wu, "Research on spline interpolation algorithm and speed optimization of CNC system," *Beijing Jiaotong University*, 2020.
 - [19] Mingrui Yin, "Research on adaptive speed planning method based on dynamic characteristics of feed system," *Shandong University of Technology*, 2020.
 - [20] Shoujin Xu, "Research on forward-looking adaptive interpolation algorithm based on NURBS curve," *Zhejiang Sci-Tech University*, 2016.
 - [21] Lin Chen, Xufeng Huang, Meng Liu, Shanshan Yuan, Feixiang He, Jian Yi, Haihong Pan, "Continuous Trajectory Look-Ahead Algorithm for Integrated Multiple Constraints Optimization," *Chinese Journal of Mechanical Engineering*, vol. 55, no. 13, pp. 151-159, 2019.
 - [22] Guang Yang, "Research and implementation of key algorithms for forward-looking direct interpolation of NURBS curve in high-speed machining CNC system," *Zhejiang University*, 2019.
 - [23] Anjiang Cai, Jie Wang, Lixia Li, Dan Zhao, "Five-axis linkage tool path non-uniform rational B-spline interpolation technology," *Science Technology and Engineering*, vol. 17, no. 9, pp. 153-158, 2017.
 - [24] Zhengjie Chen, Dongbiao Zhao, Keqinag Li, "An adaptive forward-looking smooth interpolation algorithm for micro-line segment velocity," *Mechanical Design and Manufacturing Engineering*, vol. 43, no. 3, pp. 41-44, 2014.
 - [25] Haihong Pan, Jian Yi, Meng Liu, "Research on the forward-looking interpolation algorithm of adaptive jerk NURBS curve," *Combined Machine Tool and Automated Machining Technology*, no. 11, pp. 54-59+64, 2020.
 - [26] Haihong Pan, Zengqi Yang, Lin Chen, Haitao Dong, Bingqiong Hunag, Huaqing Tan, "An Adaptive Prospective Control for Optimizing the Convergence Speed between Trajectory Segments," *Chinese Journal of Mechanical Engineering*, vol. 51, no. 5, pp. 151-159, 2015.
 - [27] Bolin Dong, Wenjie He, "Research on regional velocity trajectory planning of adjacent sensitive points in NURBS interpolation," *Mechanical Design and Manufacturing*, no. 6, pp. 177-180+184, 2020.
 - [28] Yongqiao Jin, "Research on trajectory optimization modeling and forward-looking interpolation technology for high-speed machining of tiny line segments," *Shanghai Jiaotong University*, 2015.
 - [29] Jialan Yuan, Yan Huang, Rongli Gai, "Adaptive look-ahead interpolation algorithm based on Cardinal spline curve," *Combined Machine Tool and Automated Machining Technology*, no. 10, pp. 1-5, 2018.
 - [30] Kun Ren, Jianzhong Fu, Zichen Chen, "Research on a new algorithm for speed look-ahead control in high-speed machining," *Journal of Zhejiang University(Engineering Science)*, no. 11, pp. 1985-1988, 2006.
 - [31] Guangqiang He, Dong Yu, Xiaohui Zhang, "Five-axis NURBS curve interpolation method based on speed prospective control," *Combined Machine Tool and Automation Machining Technology*, no. 5, pp. 38-41+45, 2011.
 - [32] Yingqi Li, Guirong Wang, "NURBS interpolation control method for polynomial acceleration and deceleration," *Combined Machine Tool and Automated Machining Technology*, no. 5, pp. 73-77, 2018.
 - [33] Wei Tang, Guanghui Cai, Hengfeng Qin, "Real-time look-ahead algorithm for NURBS curve parameter interpolation," *Computer Application Research*, vol. 33, no. 2, pp. 454-457, 2016.
 - [34] Renping Ye, Dehuai Zeng, "Research on Adaptive Real-time Prospective Interpolation Algorithm Based on NURBS Curve," *Machine Tool and Hydraulics*, vol. 44, no. 13, pp. 46-52, 2016.
 - [35] Renping Ye, "Research on motion modeling and simulation of five-degree-of-freedom serial manipulator," *Shenzhen University*, 2016.
 - [36] Haixia Song, "Research on modeling and trajectory planning of stone engraving robotic arm," *Huaqiao University*, 2019.
 - [37] Wang, Y. et al. "An Optimal Feed Interpolation Algorithm for High-Speed Five-Axis Machining," *Materials Science Forum*, vol. 532-533, pp. 873-876, 2006.
 - [38] Feng, Jing Chun et al, "An Optimal Feed Interpolation with Jerk-Limited Acceleration for Five-Axis Grinding," *Key Engineering Materials*, vol. 359-360, pp. 204-209, 2007.
 - [39] Li, Yuyao et al, "Variable-period feed interpolation algorithm for high-speed five-axis machining," *The International Journal of Advanced Manufacturing Technology*, vol. 40, pp. 769-775, 2009.
 - [40] Li Zhiwei, and Zeng Yanming, "CNC Machining Technology of Large Guide Vane Mixed-Flow Pump Blades," *Drainage and Irrigation Machinery*, 2006.
 - [41] Bi Qingzhen et al, "A Practical Continuous-Curvature Bezier Transition Algorithm for High-Speed Machining of Linear Tool Path," *ICIRA*, 2011.
 - [42] Bi Qingzhen et al, "Analytical curvature-continuous dual-Bezier corner transition for five-axis linear tool path," *International Journal of Machine Tools & Manufacture*, vol. 91, pp. 96-108, 2015.
 - [43] Shi, Jing et al, "Development of Real-Time Look-Ahead Methodology Based on Quintic PH Curve with G2 Continuity for High-Speed Machining," *Applied Mechanics and Materials*, vol. 464, pp. 258-264, 2013.
 - [44] Jahanpour, Javad et al, "High-speed contouring control with NURBS-based C2 PH spline curves," *The International Journal of Advanced Manufacturing Technology*, vol. 49, pp. 663-674, 2010.
 - [45] Zhao, Huan et al, "A real-time look-ahead interpolation methodology with curvature-continuous B-spline transition scheme for CNC machining of short line segments," *International Journal of Machine Tools & Manufacture*, vol. 65, pp. 88-98, 2013.
 - [46] Yau, Hong-Tzong and Junbin Wang, "Fast Bezier interpolator with real-time lookahead function for high-accuracy machining," *International Journal of Machine Tools & Manufacture*, vol. 47, pp. 1518-1529, 2007.
 - [47] Yau, Hong-Tzong et al, "Development and implementation for real-time lookahead interpolator by using Bezier curve to fit CNC continuous short blocks," *IEEE International Conference on Mechatronics*, 2005, pp. 78-83.
 - [48] Tsai, Meng-Shiun et al, "Development of a real-time look-ahead interpolation methodology with spline-fitting technique for high-speed machining," *The International Journal of Advanced Manufacturing Technology*, vol. 47, pp. 621-638, 2010.
 - [49] Yeh, Syh-Shiuh and Hsin-Chuan Su, "Implementation of online NURBS curve fitting process on CNC machines," *The International Journal of Advanced Manufacturing Technology*, vol. 40, pp. 531-540, 2009.
 - [50] Ye, Peiqing et al, "Interpolation of continuous micro line segment trajectories based on look-ahead algorithm in high-speed machining," *The International Journal of Advanced Manufacturing Technology*, vol. 37, pp. 881-897, 2008.
 - [51] He Jun, You Youpeng, and Wang Huaming, "Ferguson spline transition algorithm for high-speed machining of micro-segments," *China Mechanical Engineering*, no. 17, pp. 2085-2089, 2008.
 - [52] Guoliang Yin, "Research on NURBS curve interpolation and speed look-ahead algorithm in motion control system," *Harbin Institute of Technology*, 2019.
 - [53] Hepeng Ni, "Research on Robot Milling Trajectory Planning and Flutter

- Stability,” Shandong University, 2019.
- [54] Lei Zhang, “Research and implementation of speed-adaptive cubic polynomial look-ahead interpolation algorithm for high-speed CNC machining,” Chang’an University, 2017.
- [55] Nian Liu, “Research and implementation of real-time NURBS curve interpolation algorithm for embedded CNC system,” Nanjing University of Aeronautics and Astronautics, 2016.
- [56] Nian Liu, Dongbiao Zhao, Jianbo Yang, “NURBS curve interpolation algorithm based on Steffensen iteration method,” *Machinery and Electronics*, vol. 34, no. 4, pp. 3-7, 2016.
- [57] Dong Wei, Shuyou Zhang, Xiaojian Liu, “High-precision low-velocity fluctuation interpolation algorithm for non-uniform rational B-spline curves,” *Journal of Zhejiang University (Engineering Science)*, vol. 50, no. 11, pp. 2215-2223, 2016.
- [58] Xia Qin, Dezhao Li, Hua Deng, “Research on Position Interpolation Algorithm of Industrial Robot Based on NURBS Curve,” *Manufacturing Automation*, vol. 40, no. 4, pp. 67-72, 2018.
- [59] Peng Peng, Yi Hu, Dong Yu, Zhiyin Chen, Xiaohu Lu, “Design and optimization of adaptive speed look-ahead algorithm based on IQMath,” *Combined Machine Tool and Automated Machining Technology*, no. 7, pp. 49-52, 2014.
- [60] Peng Peng, “Research and Development of Algorithm Optimization Mechanism for High Performance Embedded CNC System,” Graduate School of Chinese Academy of Sciences (Shenyang Institute of Computing Technology), 2014.
- [61] Jianfeng Zhou, “Research on the feedrate customization method for five-axis machining of complex surfaces,” Dalian University of Technology, 2014.
- [62] Zhimei Cao, “Three-point circle method parameter adaptive interpolation algorithm,” *Mechatronic Engineering Technology*, vol. 50, no. 5, pp. 35-39, 2021.
- [63] Zhimei Cao, “Research on Adaptive Interpolation Algorithm of Non-circular Curve Parameters,” *Manufacturing Automation*, vol. 36, no. 8, pp. 4-8, 2014.
- [64] Shiming Shi, “Research on Path and Trajectory Planning Algorithm for Robot Offline Programming,” Guangdong University of Technology, 2017.
- [65] Zhiwei Su, “Research on speed planning method of three-axis CNC machining based on SAT identification speed limit interval,” Huazhong University of Science and Technology, 2018.
- [66] Hengjun Liu, Dongbiao Zhao, Zhenying Qin, “Research on NURBS curve interpolation algorithm based on Gear prediction correction method,” *Mechanics and Electronics*, vol. 35, no. 1, pp. 3-7, 2017.
- [67] Zhibing Qu, “Design and development of industrial robot embedded controller based on SERCOS III,” Southeast University, 2016.
- [68] Zhiwei Lin, “Realization of NURBS curve adaptive interpolation in high-speed CNC system and research on acceleration and deceleration control strategy,” University of Science and Technology of China, 2014.
- [69] Koren, Yoram, “Cross-Coupled Biaxial Computer Control for Manufacturing Systems,” *Journal of Dynamic Systems Measurement and Control-transactions of The Asme*, vol. 102, pp. 265-272, 1980.
- [70] Koren, Yoram and C. C. Lo, “Variable-Gain Cross-Coupling Controller for Contouring,” *CIRP Annals*, vol. 40, pp. 371-374, 1991.
- [71] Yeh, Syh-Shiuh and Pau-Lo Hsu, “Theory and Applications of the Robust Cross-Coupled Control Design,” *Journal of Dynamic Systems Measurement and Control-transactions of The Asme*, vol. 121, pp. 524-530, 1999.
- [72] Chiu, George T.-C. and Masayoshi Tomizuka, “Coordinated Position Control of Multi-Axis Mechanical Systems,” *Journal of Dynamic Systems Measurement and Control-transactions of The Asme*, vol. 120, pp. 389-393, 1998.
- [73] Chiu, George T.-C. and Masayoshi Tomizuka, “Contouring control of machine tool feed drive systems: a task coordinate frame approach,” *IEEE Trans. Control. Syst. Technol.*, vol. 9, pp. 130-139, 2001.
- [74] Sencer, Burak et al, “Modeling and Control of Contouring Errors for Five-Axis Machine Tools—Part I: Modeling,” *Journal of Manufacturing Science and Engineering-transactions of The Asme*, vol. 131, pp. 031006, 2009.
- [75] Zuhao Weng, “Research and design of efficient and flexible numerical control algorithm with continuous jerk and error adaptive characteristics,” Shanghai Jiaotong University, 2015.
- [76] Yitian Cai, “Research on Adaptive Contour Error Estimation and Contour Control Based on Real-time Ethernet,” Shanghai Jiaotong University, 2016.
- [77] Yang Chen, “Research on Contour Accuracy Control Method of Numerical Control System,” Tianjin University, 2016.
- [78] Yang Chen, Taiyong Wang, Jingchuan Dong, Bo Li, Zhiqiang Yu, “Contour error control based on optimal circle approximation method for contours,” *Computer Engineering and Science*, vol. 38, no. 6, pp. 1281-1285, 2016.
- [79] Henan Xu, “Research on Contour Control Method of CNC Machine Tool Feed Servo System,” Tianjin University, 2017.
- [80] Yiqiao Huang, “Efficient detection and comprehensive compensation of geometric errors of five-axis CNC machine tools,” Shanghai Jiaotong University, 2016.
- [81] Zhen Sang, “Research on machining path generation method of multi-axis single-channel double-head woodworking CNC machine tool,” Tianjin University, 2017.
- [82] Qian Wang, “Research on tool radius compensation algorithm of CNC system supporting NURBS interpolation,” Harbin Institute of Technology, 2017.
- [83] Hui Luo, “Research on Motion Control Method of Open CNC System,” Shanghai University of Engineering Technology, 2016.
- [84] Jipeng Yuan, “Research on Adaptive NURBS Curve Fitting and Its Velocity Planning Algorithm,” Shandong University, 2019.
- [85] Shujie Sun, Hu Lin, Liaomo Zheng, “A NURBS Curve Prospective Interpolation Algorithm for Backward Interpolation,” *Journal of Computer Aided Design and Graphics*, vol. 26, no. 9, pp. 1543-1549, 2014.
- [86] Shujie Sun, Hu Lin, Limo Zheng, Feng Liu, “NURBS Curve Interpolation Algorithm with Adaptive Computational Complexity,” *Small and Microcomputer Systems*, vol. 35, no. 4, pp. 895-899, 2014.
- [87] Liangchen Hu, Wensheng Zhang, “NSGA-II approach for proper choice of nodes and knots in B-spline curve interpolation,” *Computer-Aided Design*, vol. 127, pp. 102885, 2020.
- [88] Wei Ren, “Research on Trajectory Planning of Six Degrees of Freedom Industrial Robot,” Anhui Engineering University, 2020.

LACGC: Business Sustainability Research Model for Generations Consumption, Creation, and Implementation of Knowledge: Academic and Non-Academic

Satpreet Singh

Abstract— This paper introduces the new LACGC model to sustain the academic and non-academic business to future educational and organizational generations. The consumption of knowledge and the creation of new knowledge is a strength and focal interest of all academics and Non-academic organizations. Implementing newly created knowledge sustains the businesses to the next generation with growth without detriment. Existing models like the Scholar-practitioner model and Organization knowledge creation models focus specifically on academic or non-academic, not both. LACGC model can be used for both Academic and Non-academic at the domestic or international level. Researchers and scholars play a substantial role in finding literature and practice gaps in academic and non-academic disciplines. LACGC model has unrestricted the number of recurrences because the Consumption, Creation, and implementation of new ideas, disciplines, systems, and knowledge is a never-ending process and must continue from one generation to the next.

Keywords—Academics, Consumption, Creation, Generations, Non-Academics, Research, Sustainability

Satpreet Singh is CEO at Ardass Corporation, Manteca, CA 95336 USA
(phone: 209-400-0005; fax: 209-923-4619; e-mail:
satpreetsingh@ymail.com).

Diagnosis and Resolution of Intermittent High Vibration Spikes at Exhaust Bearing of Mitsubishi H-25 Gas Turbine using Shaft Vibration Analysis and Detailed Root Cause Analysis

Fahad Qureshi, Faheem Ahmed

Abstract— This paper provides detailed study on the diagnosis of intermittent high vibration spikes at exhaust bearing (Non-Drive End) of Mitsubishi H-25 gas turbine installed in a petrochemical plant in Pakistan. The diagnosis is followed by successful root cause analysis of the issue and recommendations for improving the reliability of machine. Engro Polymer and Chemicals (EPCL), a Chlor Vinyl complex has a captive power plant consisting of one combined cycle power plant (CCPP), having two gas turbines each having 25 MW capacity (make: Hitachi) and one extraction condensing steam turbine having 15 MW capacity (make: HTC). Besides, one 6.75 MW SGT-200 1S gas turbine (make: Alstom) is also available. In 2018, the organization faced an issue of intermittent high vibration at exhaust bearing of one of H-25 units having tag GT-2101 A which eventually led to tripping of machine at configured securities. Since machine had surpassed 64,000 running hours and major inspection was also due so bearings inspection was performed. Inspection revealed excessive coke deposition at labyrinth where evidence of rotor rub was also present. Bearing clearance was also at upper limit and slight babbitt (soft metal) chip off was observed at one of its pads so it was preventively replaced. The unit was restated successfully and exhibited no abnormality until October 2020 when these spikes reoccurred leading to machine trip. Recurrence of the issue within two years indicated that root cause was not properly addressed so this paper furthers the discussion on in-depth analysis of findings and establishes successful root cause analysis which captured significant learnings both in terms of machine design deficiencies and gaps in operation & maintenance (O & M) regime. Lastly, revised O& M regime along with set of recommendations are proposed to avoid recurrence.

Keywords— Exhaust side bearing, Gas turbine, Rubbing, Vibration.

I. INTRODUCTION

ENGRO Polymer and Chemicals Limited has 61.5 MW rated capacity of power generation to meet inhouse load demands. Out of four turbines (three gas turbines, one steam turbine) at site, reliability of two Mitsubishi H25 gas turbine (design specifications shown in table 1) is critical for plant sustainable operations. These two units were commissioned in 2009 and remains in operation throughout year except planned maintenance shutdowns. Both turbines have surpassed 100,000 running hours as of 31st December 2021, but major inspection (MI) has not been performed as recommended by OEM (see

details in table 2) while combustion inspection (CI) and hot gas path inspection (HGPI) have been performed.

TABLE I
MITSUBISHI H-25 DESIGN SPECIFICATIONS

Parameter	Specification
Power Output	22,530 kW (@ 47 °C, 45 % RH)
Turbine Inlet Air Pressure	13.68 kg/cm ² (abs)
Turbine Inlet Temperature	1,290 °C
Exhaust Temperature	576 °C
Compressor	17 Stages, Axial Type
Combustor	Can Annular
Turbine	3 Stages
Shaft Speed	7,258 r/min
Fuel Type	Natural Gas

TABLE 2
OEM MAINTENANCE PLAN FOR H25

Activity	Due Hours
Combustion Inspection	16,000
Hot Gas Path Inspection	32,000
Major Inspection	64,000

A. Machine Construction

Mitsubishi H-25 is a variant of GE Frame V with similar power capacity (see figure 1). Turbine is coupled to generator via a reduction gearbox having final output rpm of 1500. Generator is installed at cold end (air inlet) of turbine whereas hot end (exhaust) is free and flue gases are routed towards heat recovery steam generator (HRSG).

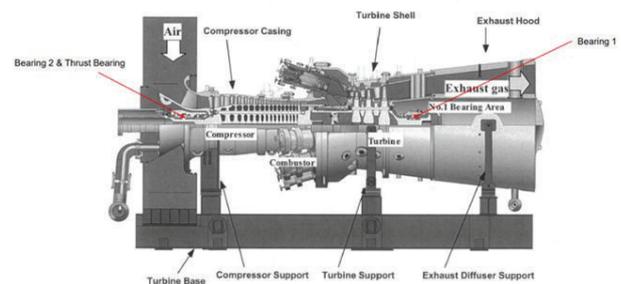


Fig. 1 Mitsubishi H-25 construction

Fahad Qureshi is Rotary Equipment Engineer at Engro Polymer and Chemicals Ltd, Karachi 75000, Pakistan (phone: +92-332-2826058; email: fqureshi@engro.com)

Faheem Ahmed is Section in Charge of Rotary Equipment at Engro Polymer and Chemicals Ltd, Karachi 75000, Pakistan (email: faheemahmed@engro.com)

Load coupling (rigid) is used between turbine and gearbox having spacer length of 800 mm whereas diaphragm coupling is used between generator and gearbox having 900 mm spacer length. Turbine is radially supported by two journal bearings having tilting pad configuration, one is installed at cold end (compressor stub shaft) having 90 mm bearing width (Bearing 2) whereas the other one is installed at its hot end (turbine stub shaft) having 110 mm bearing width (Bearing 1). For axial support, one thrust bearing (active & inactive) is installed at its cold end. See figure 2.

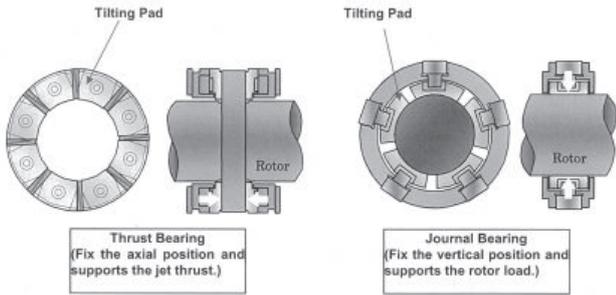


Fig. 2 H-25 bearings construction

B. Monitoring & Protection

Machine is equipped with Mark VI controller having Bently Nevada 3300 system for vibration monitoring. Each journal bearing is equipped with two proximity probes having 45 degrees right and left configuration. One key phaser is installed at turbine stub shaft flange (exhaust side). See figure 3.

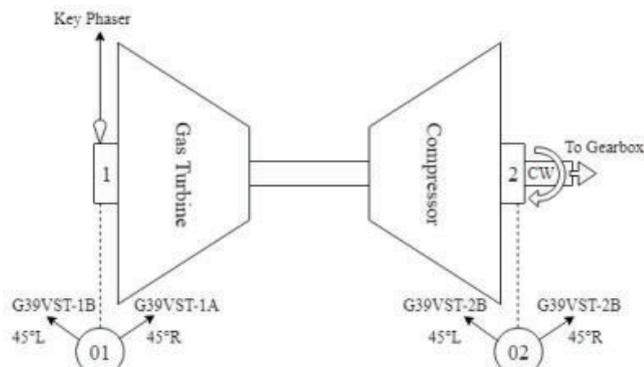


Fig. 3 H-25 vibration instrumentation

Alarm and Trip values with voting logic is shown in table 3.

TABLE 3
H-25 VIBRATIONS ALARM AND TRIP LOGICS

Alarm	105 μm
Trip	155 μm
Trip Actuation	Any one out of four in alarming range and one in trip range

There is no instrumentation available for monitoring of rotor position however all bearings (radial + thrust) oil drain temperatures are measured but that doesn't reflect accurate condition compared to metal temperatures. There is no trip on oil temperatures but alarms only.

II. FAILURE EVENT

On 22nd October 2020, during periodic hot washing activity on one of the H-25 unit having tag GT-2101 A, vibration at bearing 1 X probe reached to 102 μm . Immediately washing was aborted and load reduced which stabilized the vibration. Prior to this, one event was observed on same 1X probe on 17th October 2022 when vibration reached to 73 μm . On rated load, vibrations at bearing 1X & 1Y normally runs below 50 μm so this situation was considered abnormal. These vibration events were similar to the events experienced back in October 2018 [1] but its recurrence within two years was questionable and incomprehensible.

A. Sequence of Events

Due to plant full load operation, it was not possible to shutdown turbine for detail inspection until proper root cause was established so turbine was kept in operation while intermittent vibration events were reoccurring with progressive amplitude at same location (bearing 1X). See details in table 4. Meanwhile, its tripping setpoint was increased to 165 μm (no change in delay timer) which was within permissible limits as defined in [2].

TABLE 4
GT-2101 A VIBRATION EVENTS AT BEARING 1X

S. No	Date	Running Load (MW)	Max. Vibration (μm)	Action Taken
1	17/10/2020	21.8	73	No action taken
2	22/10/2020	22.2	102	Washing aborted and vibration stabilized.
3	26/10/2020	23.2	85.4	No action taken
4	28/10/2020	22.5	83.2	01 MW load reduced
5	20/11/2020	6	118	02 MW load increased
6	25/11/2020	4.5	83	01 MW load increased
7	10/12/2020	21.8	156.3	04 MW load reduced but turbine tripped
8	10/12/2020	0	143.7	No action taken
9	22/12/2020	21.8	72	No action taken
10	23/12/2020	21.5	97.3	0.5 MW load increase
11	23/12/2020	21.8	55	No action taken
12	24/12/2020	21.6	72	No action taken
13	24/12/2020	22	82	No action taken
14	24/12/2020	21.7	131.1	01 MW load increased
15	27/12/2020	22.5	127.9	0.6 MW load increased
16	11/01/2021	16.5	166.1	2 MW load increased but turbine tripped

Machine remained in operation in vulnerable condition for almost eight days until January 11th, 2021, when machine tripped for second time and root cause with plan was established for rectification.

III. DIAGNOSIS

To analyze the issue and devise way forward, it was decided to analyze machine parameters (oil pressure, oil temperature, air flow etc.) trend and utilize ADRE SXP and DSPI 408 for

shaft vibration analysis [3] since GE system 1 is not available on this machine for diagnosis.

A. Preliminary Analysis

Before starting investigation, one thing was evident that the vibration pattern looked exactly similar to events experienced in 2018 and at that time following corrective actions were taken:

- 1) 10,300 liters out of 12,000 liters oil was replaced due to high varnish content.
- 2) Bearing 1 & 2 were replaced due to slightly high radial clearance against OEM recommended range.
- 3) Bearing 1 labyrinths and housing cleaning was performed.

So, recurrence of issue within two years raised questions on

what went wrong and missed.

All the machine parameters extracted from Mark VI were trended against ambient temperature and load, but no correlation could be built. Only thing common between these two events was cold season. In both 2018 and 2020, vibration spikes initiated in October end when ambient temperature starts decreasing, and lube oil header temperature went below 60 °C so instructions were made to keep it above 60 °C but vibration spikes continued.

Steady state vibration data acquired through ADRE captured two vibration events on 26th October 2020. These were recorded at 21:53 hours and 23:32 hours with amplitude up to 23.3/25.1 um pp and 67.6/85.4 um pp respectively on Turbine bearing 1. During the peaks at bearing 1, it was observed that vibration reduced in amplitude on Turbine bearing 2 to 43.4/45.7 um pp

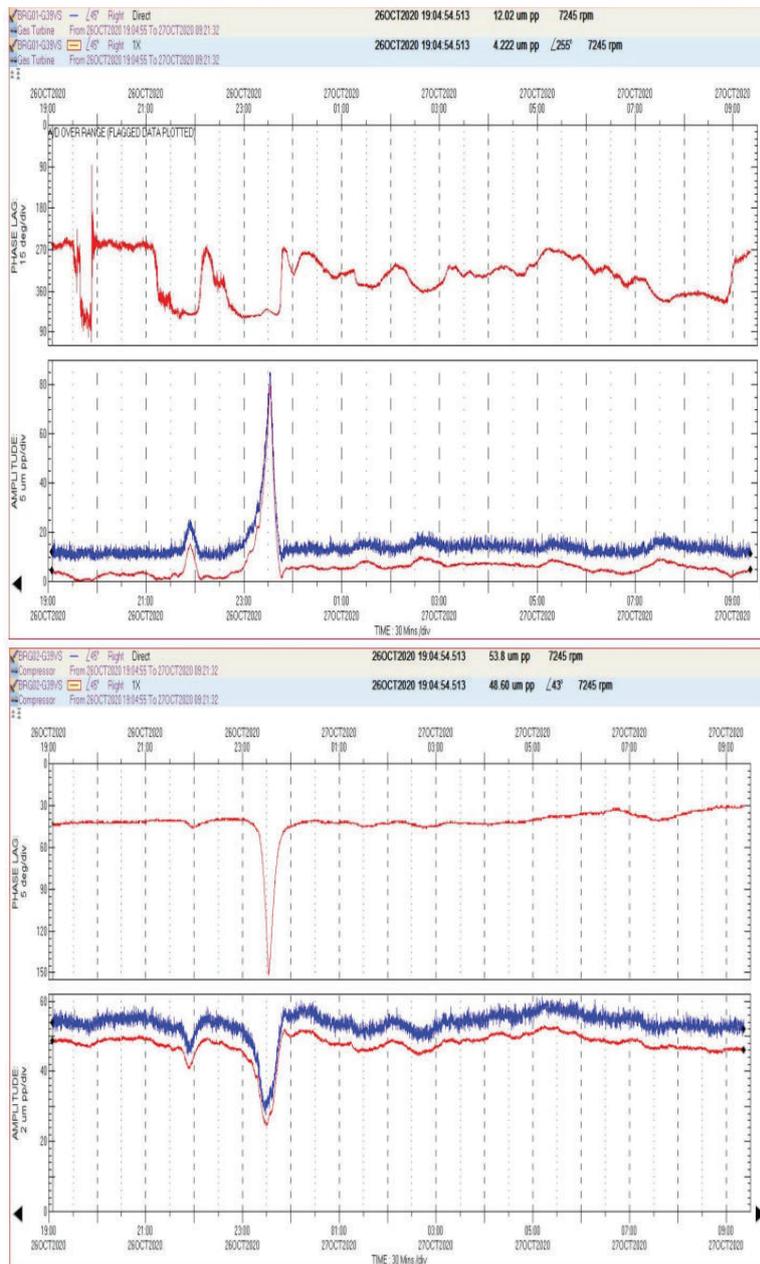


Fig. 4 (a) Time base trend of Bearing 1 & 2 amplitude and phase angle

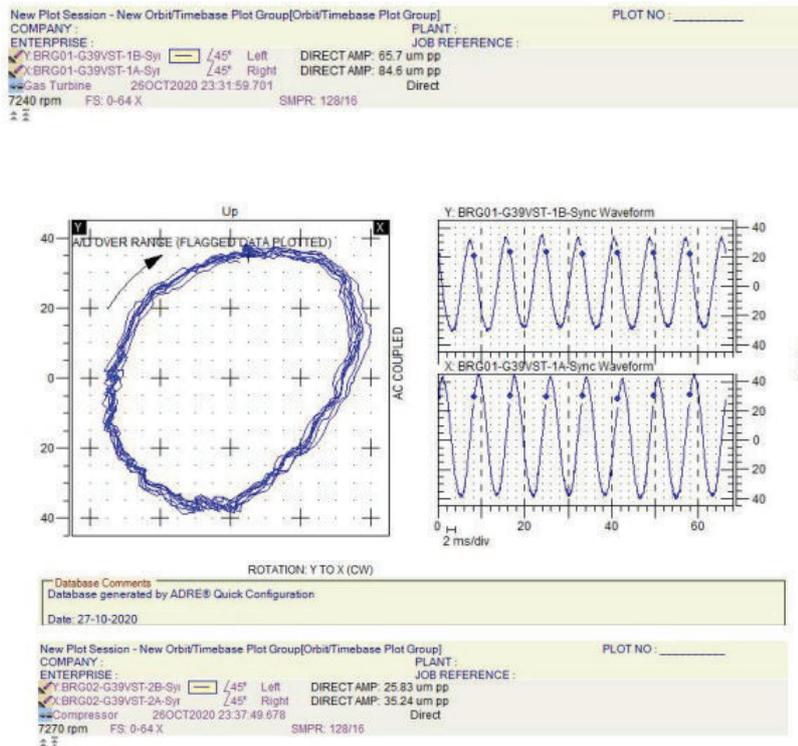


Fig. 4 (b) Orbit plots of Bearing 1 & 2

and 23.2/27.7 $\mu\text{m pp}$ respectively. However, no vibration change was observed at Generator bearings.

Duration of first vibration peak was observed to be 45 minutes (vibration increase started at 21:44 and normalized at 22:03 hours) and that of second peak was observed to be 56 minutes (vibration increase started at 22:48 and normalized at 23:44). Almost all vibration change was observed at frequency of 1x Turbine rpm while significant change in 1x phase up to 20° and 113° respectively was observed at Turbine bearing 1 and 2 which most probably corresponds to symptoms of light rubbing [4]. See Fig. 4a.

Orbit plots during vibration peaks at Turbine bearing 1 indicated maximum amplitude corresponding to vibration peak while at Turbine bearing 2 orbit collapsed to a minimum amplitude and also found slightly flatter which most probably indicates light rubbing (Fig. 4b).

Based on the rotor dynamics theory, analysis of the steady state data depicted a partial rub condition as it occurs mostly once per revolution of the rotor (1X), which further

corresponded to normal tight rub condition. This was proven by the fact that rotor was moving in unconstrained condition and only during the instantaneous contact with stationary component (seal), its motion became constrained resulting in decrease and then increase in vibration as contact was broken. [5]. Periodic nature of the vibration events ruled out the possibility of unbalance, misalignment or rotor crack.

During vibration spikes, machine operating parameters like inlet air temperature & pressure, exhaust duct temperature & pressure, oil temperature & pressure, bearings temperatures, IGV position and load remained constant therefore vibration spike was not related to change in operational conditions.

Hence, vibration peaks observed at Turbine bearings 1 and 2 with dominant 1x rpm response and 1x phase change are most probably suspected due to some light rubbing event taking place either in Turbine bearings or bearing oil seals.

Rectification of this issue required bearing inspection which was not possible to perform in 2020 so turbine had to be kept running. Parameters were analyzed so correlation could be

deduced triggering these vibration events. Most of the events occurred in late evening and midnight when inner barrel (Fig. 5a) temperature reduced by 15-20 °C. This inner barrel has circulating air for which two bearings cooling fans (S-2120 A/B) are installed. Its specification is enlisted in table 5.

TABLE 5
BEARING COOLING FAN SPECIFICATIONS

Parameter	Specification
Type	Centrifugal
Inlet Air Temperature	3 to 120 °C
Inlet Air Pressure	Ambient Pressure (kPa)
Discharge Pressure	0.462 KPa @ 15 °C
Discharge Flow	410 m ³ /min

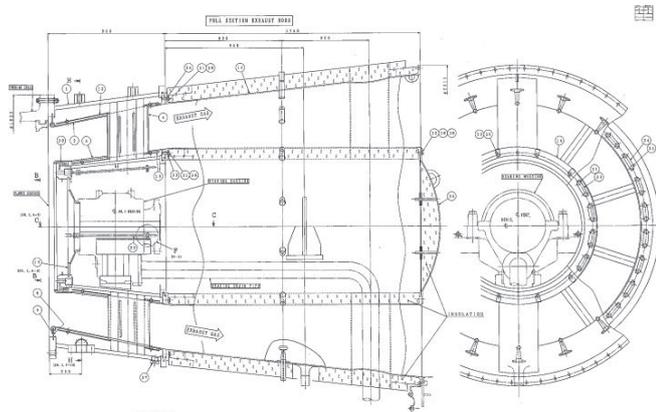


Fig. 5 (a) Exhaust hood inner barrel where bearing 1 housing is located

To eliminate inner barrel temperature variation, an inhouse fabricated steam coil heater was installed at its inlet as shown in Fig. 5b which worked well in controlling temperature variation.



Fig. 5 (b) Steam coil heater installed at inlet of bearing cooling fan

Post installation of heater, following were observations:

- December 26th, 2020, no vibration spike.
- December 27th, 2020, vibration spike (127.9 μm) observed at bearing 1X.
- December 28th, 2020, no vibration spike.

Though frequency of vibration events reduced but not eliminated so December 27th, 2020, event was analyzed (Fig. 6) which concluded that these vibration spikes are completely random and can best be controlled though load management (human factor).

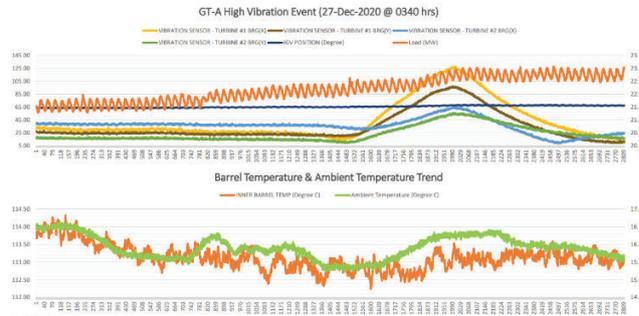


Fig. 6 Bearing 1 vibrations trend against load, IGV position, ambient and barrel temperature

B. Root Cause Analysis (RCA)

It was confirmed that the vibration events were triggered due to light rubbing at either bearing or seals but its recurrence within two years was intelligible as almost 86 % oil was replaced, and new bearing was installed in 2018. Simple oil analysis performed on monthly intervals reported no issue however advanced oil analysis was performed for complete assurance. Its results were received as shown in Fig. 7 which showed neither abnormality nor any recommendation so formation of thick layer of coke was questionable within 16,000 hours.

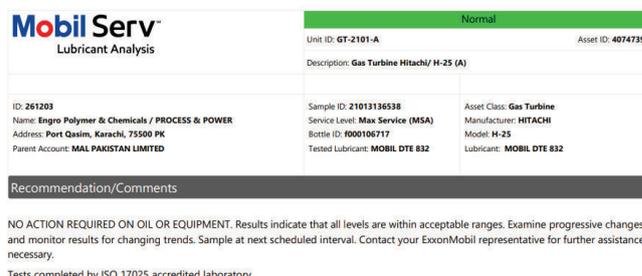


Fig. 7 Advanced oil analysis results of GT-2101 A performed on 08th Jan 2021

Instrumentation health (vibration probes) was also checked but there was no issue as confirmed by both inspection and vibration trend which was gradual. Usually in case of instrumentation malfunction, rapid fluctuations are observed which was not in this case.

Finally, a detailed sequence of events and activities around this turbine over period of last one month was enlisted to identify any irregularity. Surprisingly, on October 7,2020, on west side of GT-2101 A (Fig.9a), sand excavated from project

construction activity at other unit was dumped which remained there for seventeen days until removed on October 18th, 2020.

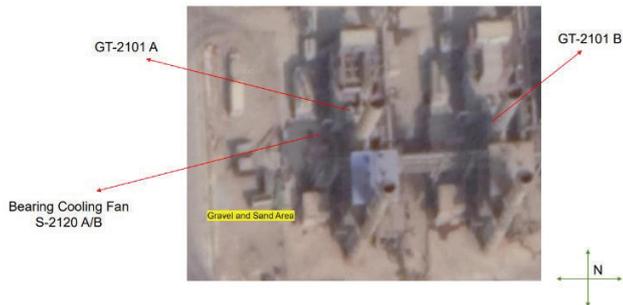


Fig. 9 (a) Panoramic top view of GT-2101 A/B captured from google satellite image.

During this period, surrounding air quality was compromised leaving vulnerable condition for turbine compressor and bearing cooling fan. Bearing cooling fan housings were checked immediately after this observation and their sealing was found damaged resulting in unfiltered air passing through as shown in Fig. 9b and 9c.



Fig. 9 (b) Bearing cooling fan filter seat sealing condition

But as per cooling and sealing air diagram provided by OEM, bearing cooling fan is just designed for ventilation of inner barrel so its air quality should have no impact on bearing. However, by design there is sealing air provided on bearing 2 from compressor 6th stage whereas there is no sealing air on



Fig. 9 (c) Bearing cooling fan silencer inside condition

bearing 1. This deviation was marked to OEM during 2018 vibration event, but no feedback was received. This deviation was highly probable resulting in coke formation as bearing 1 is located at hot end and with absence of sealing air, oil could travel and burn specially if labyrinth is damaged.

During this entire episode, a separate discussion was in place with OEM on high 3rd stage wheel space temperatures as it was never replaced. A technical letter was issued to EPCL by Mitsubishi on this situation which revealed that bearing cooling fan air is not just for inner barrel ventilation but for sealing of bearing 1 and cooling of 3rd stage wheel after space location as shown in Fig. 8. This was missing from OEM cooling and sealing air diagram (a major loophole) resulting in its potential impact unidentified since years and not addressed in 2018 vibration event where oil aging (Coking) was assumed to be the root cause.

1. Conclusion

With the enlightenment of bearing 1 sealing and 3rd stage after wheel space cooling path, it became clear that unfiltered air (high dust) remained flowing for seventeen days when loose sand was dumped in the vicinity at west side of GT-2101 A resulting in buildup of deposits at turbine seals which reduced clearance and rotor momentarily rubbed triggering vibration spikes.

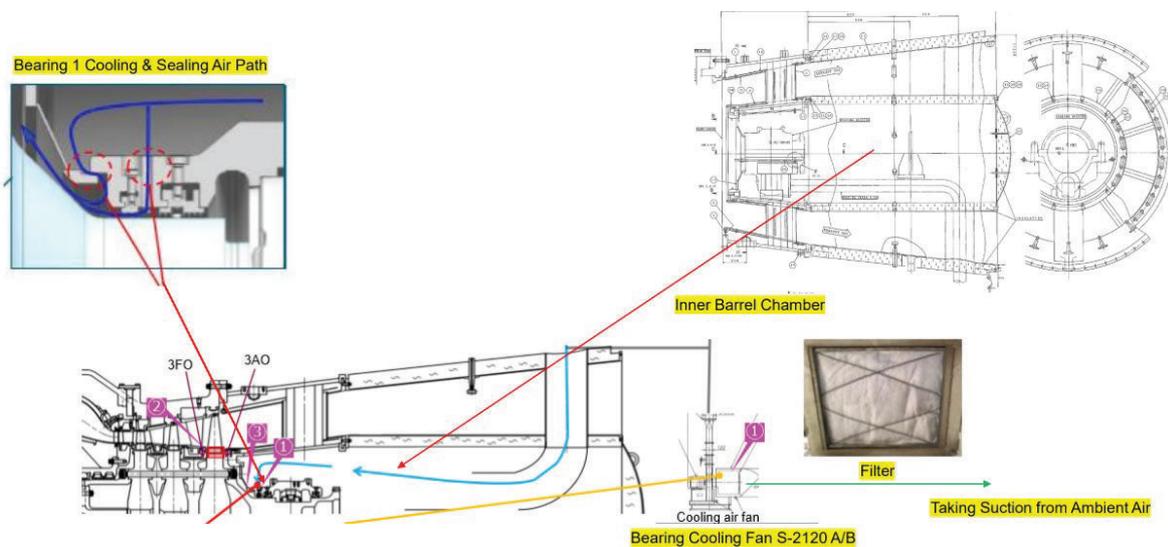


Fig. 8 Cooling and Sealing air path of Bearing 1 housing

IV. RESOLUTION

The root cause was already established for recurrence of this intermittent rubbing and now only opportunity was awaited which was not possible in year 2020 due to production requirements. Machine tripped on high vibration on 11th January 2021 and was finally taken for bearing 1 inspection.

A. . Inspection

Since thrust bearing temperatures and vibrations at bearing 2 were normal so they were not inspected. After cooling of turbine and removal of exhaust hood casing, access to bearing 1 housing was achieved. Inner conditions of the barrel in vicinity of bearing showed depositions of dust as shown in Fig. 10



Fig. 10 Top view of bearing 1 housing showing deposits of dust at half joint bolts.

Close inspection of the sealing and cooling air path showed clogging due to which air flow was restricted resulting in both increased 3rd stage aft wheel space (3AO) temperatures (Fig.11a) and intermittently rotor rubbed at forward side of bearing 1 labyrinth seal as shown in Fig. 11b



Fig. 11 (a) Mating joint of exhaust cone and bearing 1 housing from where air flows to cool 3rd stage wheel aft space.

Thick deposits of black color apparently carbonized oil were formed on forward side labyrinth as shown in Fig.11c. Deposits thickness varied across 360 degrees and showed clear sign of rotor rub due to reduced clearance. Exactly similar condition was faced in 2018 vibration events. Sample of these deposits was collected and send to inhouse lab for analysis of its constituents to validate hypothesis.

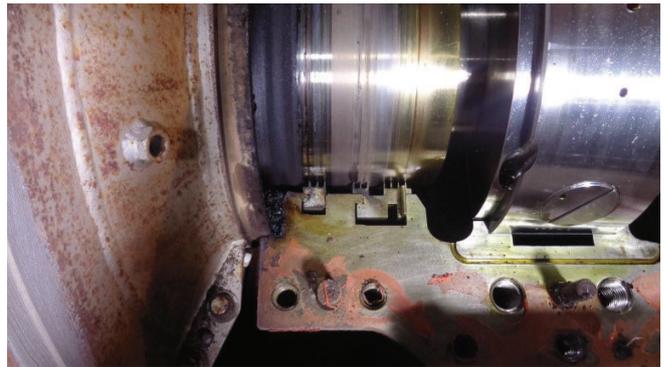


Fig.11 (b) Bearing 1 and rotor with housing top cover removed, showing deposits at forward side labyrinth nearest to 3rd stage wheel.



Fig. 11 (c) Bearing 1 housing top cover forward side labyrinth seal where rotor was rubbing due to thick deposits.

Bearing 1 as found radial clearance measured was 0.38 mm which was slightly higher than OEM recommended range (0.21-0.37 mm) however tilting pads & babbitt condition was healthy as verified through non-destructive testing (dye penetrant test + ultrasonic test). There was no surface or sub-surface crack however bearing was preventively replaced with new one and radial clearance measured was 0.33 mm, within OEM recommended range. Please note both as found and as left bearing clearances were measured without crush of bearing 1 housing top half.

After side of bearing 1 labyrinth had no deposits since there is no sealing air. All labyrinths were cleaned and found intact (Fig 12). As found clearances measured were within the OEM recommended limits so same labyrinths reused.

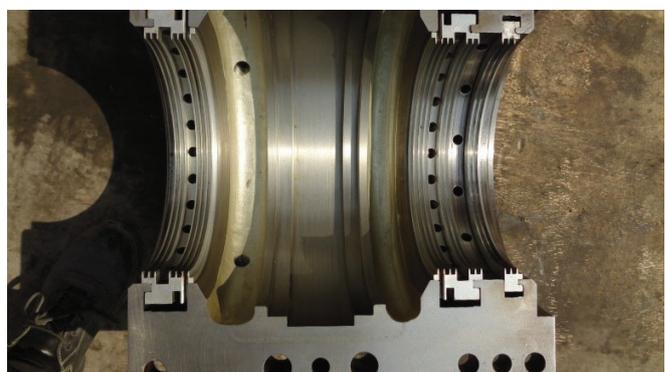


Fig. 12 Bearing 1 housing top cover labyrinths after cleaning

Bearing 1 protection seal was also inspected which is designed to prevent hot gas air flow to bearing during startup & shutdown and found to be healthy.

Analysis of the deposits showed presence of sand as predicted. Please see table 6.

TABLE 6
BEARING 1 DEPOSITS LAB ANALYSIS

Constituent	Percentage
Organic matter	39.86
Fe2O3	12.7
Silica as insoluble material	32.56
Aluminum	1.48
Calcium	6.79
Magnesium	1.91
Nickel	Nil
Strontium	Nil
Barium	Nil

B. Test Run

After completion of maintenance activity, turbine was handed over to operations for cold washing followed by startup. Machine startup successfully carried out on January 20th, 2021, at 16:42 hours and it ramped up smoothly till achieving full rpm. After analyzing its parameters for ten minutes, go ahead was given to sync it with grid. Load on turbine was gradually increased and reached 21 MW on 22nd January 2021 at 02:14 hours. Max vibration observed was 64.41 μ m at bearing 1X whereas at 1Y it was 63.03 μ m. Though there was no intermittent vibration spike observed but both bearing 1 & 2 vibration values were comparatively higher however it was in satisfactory zone as per ISO standard 7919-3/10816-7 [6]. The probable reason for this change could be attributed to change in rotor position thereby impacting alignment as it was not corrected and bearing 2 clearance was not measured.

Steady state ADRE data was collected for vibration analysis and following were key observations:

- 1) No vibration spike with phase change was observed which was evident before bearing inspection activity.
- 2) Spectrum plots of turbine bearing 1 were compared for current and previous data of 24- Dec-2020. As previously, the dominant vibration was observed at 1x frequency and 1x amplitude and no sub synchronous frequency was visible, a typical condition of rub. Please see Fig. 13a & 13b
- 3) Turbine bearing 1 orbit amplitude was found larger corresponding to higher overall vibration and bearing 2 orbit was found more elliptical (showing heavy preload) as shown in Fig. 14a & 14b however there was no symptom of rubbing.
- 4) As previously, the dominant vibration was observed at 1x frequency, and 1x amplitude at turbine bearing 1 was found significantly higher corresponding to overall increased vibration.

Besides vibration, one major improvement post inspection was significant reduction in 3rd stage wheel after space temperatures as shown in table 7. It recorded 26.15 % & 23 % drop in 3AO1 and 3AO2 respectively which will increase creep life consumption time of 3rd stage components.

TABLE 7
3RD STAGE WHEEL SPACE AFT TEMPERATURES @ 21 MW

Parameter	Before Maintenance	Post Maintenance
3AO1	412	304.25
3AO2	414	318.46

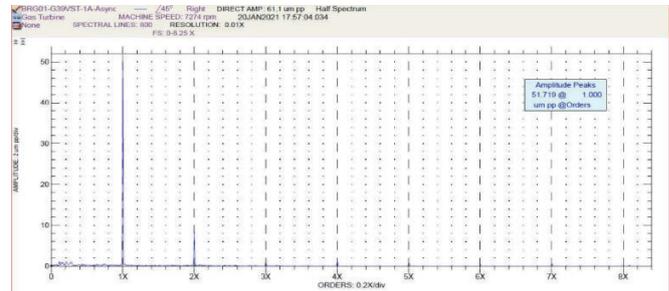


Fig. 13 (a) Bearing 1X spectrum plot

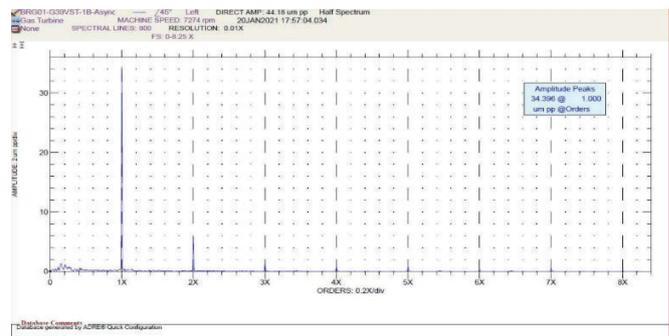


Fig. 13 (b) Bearing 1Y spectrum plot

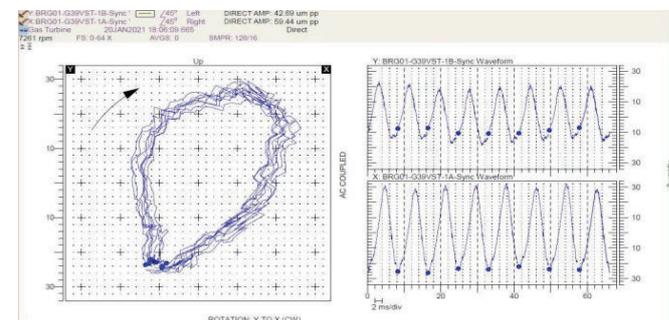


Fig. 14 (a) Bearing 1 orbit plot @ 4 MW

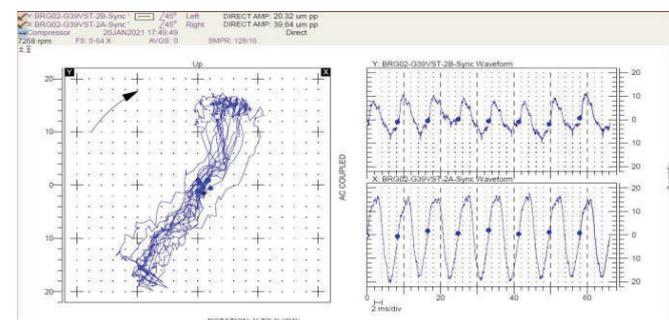


Fig. 14 (b) Bearing 2 orbit plot @ 4 MW

V.LEARNINGS

Recurrence of this issue within two years was a major setback as bearing 1 inspection requires at least five days wrench time costing plant reduced production and difficulty in load management. Significant learnings were captured from

this episode which highlighted gaps in design, and operation & maintenance practices necessitating immediate rectification to resolve this issue for good. Following were the key learnings:

- 1) In 2018 vibration events, compromised oil quality coupled with overdue bearing inspection was established as the root cause which assumed aging as the major contributor so cooling & sealing air path issue remains veiled.
- 2) Reliance on OEM technical literature can be dubious as cooling and sealing air diagram provided did not explicitly mention bearing 1 sealing resulting in this issue unaddressed in 2018.
- 3) Even though bearing cooling fan filters were replaced timely but its seating area was damaged and never repaired resulting in unfiltered air passing through bearing 1 air path. This issue was rectified with bearing inspection.
- 4) 3rd stage aft wheel space temperature (3AO) was controlled through cooling air circulating through bearing 1 housing. Due to clogging of this path, 3AO temperature were running 50-100 °C higher thus reducing the creep life of components.
- 5) Compared to GT-2101 A, the other H-25-unit GT-2101 B was doing well as it was not exposed to open surroundings with dust in vicinity so alongside bearing inspection activity, west side area of GT-2101 A was cemented to improve air quality as shown in figure 15 a & b.



Fig. 15 (a) West side area of H-25-unit GT-2101 A before root cause was established. Full of loose sand.



Fig. 15 (b) West side area of H-25-unit GT-2101 A after root cause was established. Cemented to improve air quality around turbine vicinity.

VI. RECOMMENDATIONS

Most of the gaps identified in this incident RCA were immediately addressed alongside bearing inspection however following recommendations were generated to ensure its

closure proper closure.

- 1) Considering low world air quality index (WAQI) of Karachi, Pakistan which remains above 100, it is recommended to perform bearings inspection at reduced frequency to clean cooling and sealing air path. It should be preferably integrated with HGPI at 32,000 hours.
- 2) Pressure- drop monitoring to be configured across bearing cooling fan filters so they are timely replaced rather than visual condition.
- 3) Bearings cooling fan filter original make was Viledon PS600 however they were localized few years ago as part of side wide supply chain exercise to reduce cost and lead time. This change to be restudied for drop in filter efficiency and revert if deviation is found.
- 4) An enclosed canopy to be built around bearing cooling fans to pre filter air and reduce direct exposure to atmosphere.
- 5) OEM to be notified for updating its cooling and sealing air diagram so other H-25 users can understand its significance and timely address it.
- 6) 3rd Stage wheel space temperatures (3AO1-2) to be closely monitored as its higher value is a clear indicator of clogging of sealing and cooling air path.

ACKNOWLEDGMENT

The authors of this paper are grateful to management of EPCL for providing this opportunity and their assistance in transmuting this vital experience into publication so other turbine users can incorporate its learnings to improve their operations and maintenance practices.

REFERENCES

- [1] Rashid, Abubakar, Saad, Muhammad, and Ahmed, Faheem. "Diagnosis of Intermittent High Vibration Peaks in Industrial Gas Turbine Using Advanced Vibrations Analysis." *WASET International Journal of Mechanical and Mechatronics Engineering*, vol:13, no:9, June 2019.
- [2] ISO – 10816 – Mechanical Vibration
- [3] Desimone, G., "Fundamentals of Signal Processing applied to rotating machinery diagnostics," Proceedings of the 43rd Turbomachinery and 30th Pump Users Symposia, 2014.
- [4] Eisenmann, Robert C Sr and Eisenmann, Robert C Jr." Common Malfunctions." *Machinery Malfunction Diagnosis and Correction*. PTR Prentice Hall, New Jersey (1997): pp. 440-443
- [5] Hatch, C. & Donald E. Bently, Fundamentals of Rotating Machinery Diagnostics. Bently Nevada Press, 2002.
- [6] ISO – 7919 – Mechanical Vibration

Transmission Characteristics and Anisotropy of Epsilon Near Zero Behaviour in Photonic Hypercrystal

Munazza Zulfiqar Ali

Abstract— Photonic Hypercrystal (PHC) is a recently studied novel phenomenon that incorporates the characteristics of hyperbolic metamaterial (HMM) and photonic crystals (PC). We investigate two types of one-dimensional photonic hypercrystals, one that exhibits Bragg gaps and the other that does not. Three different types of gaps emerging from three different mechanisms are investigated theoretically. For transverse magnetic (TM) polarized incident light, there is a frequency region where the HMM is of type II prohibiting propagation in the direction of growth of the crystal giving rise to a transmission gap. The second type is the conventional Bragg gap and the third type is the plasmon-polariton gap arising from the coupling of the longitudinal bulk plasmon mode and the perpendicular component of the electric field. We also report two frequency regions where parallel or perpendicular components of permittivity tensor become nearly zero and anisotropic epsilon near zero (AENZ) phenomenon takes place. The dependence of these gaps on the layer widths, incident angle, the polarization of incident radiation, and the filling factor of the crystal is investigated theoretically. Propagation of a plane wave at some frequencies of interest is studied by the intensity profile of transmitted wave using finite difference time domain (FDTD) simulations.

Keywords— photonic crystals, hyperbolic metamaterials, epsilon near-zero materials, transmission characteristics.

Fibrations Property of Embedding Maps of Orbifold Charts

Hakeem A. Othman * and Santanu Acharjee **

*Department of Mathematics, AL-Qunfudhah University college, Umm Al-Qura University, KSA.

*Department of Mathematics, Rada'a College of Education and Science, Albaydha University, Albaydha, Yemen.

**Department of Mathematics, Gauhati University, Guwahati-781014, Assam, India.

* hakim_albdoie@yahoo.com **sacharjee326@gmail.com

Abstract

In this paper we introduce the notion of Hurewicz fibrations in the class of embedding maps of orbifold charts by giving the concept of E-fibration embedding. We study the fundamental properties of this concept such as the restriction, product and relationship between it and Hurewicz fibration. Furthermore, we introduce the notion of lifting functions of E-fibration embedding and study preserving projection property of these lifting functions.

Keywords: Orbifold, Embedding, Fibration, Homotopy.

AMS classification: 32C05, 14E25, 55P05.

1 Introduction

The concept of orbifold has been described using various mathematical constructions and contexts. It was first introduced by Satake [7], [8]. He called them V-manifolds. He defined orbifolds as topological spaces with an atlas of charts. He viewed orbifolds as a generalization of manifolds. Orbifolds and Manifolds are described by charts. An orbifold chart under any space X is defined as a triple (X_U, G_U, Γ_U) , where X_U is open set in R^n , G_U is a finite group of homeomorphisms of X_U and Γ_U is a map. One of the issues with the atlas definition is that there is no canonical notion of map between orbifolds. Satake introduced maps of orbifolds, called embedding maps which considered a generalization of smooth maps of manifolds.

In this paper, Section 2 introduces the concept of E-fibration embedding and studies the fundamental properties of E-fibration embedding such as the restriction property, product property and relationship between E-fibration embedding maps and Hurewicz fibrations. In Section 3, we introduce the notion of lifting functions of E-fibration embedding by giving the concept of E-lifting function and regular E-lifting function. In Section 4, we show preserving projection property for E-lifting functions.

Throughout this paper all spaces will be assumed Hausdorff spaces. For any space X , X^I denotes the set of all continuous functions (paths) from $I = [0, 1]$ into X , we take this set with the compact-open topology. For all $x \in X$, by \tilde{x} we mean the constant path at a point x . For two paths $\alpha, \beta \in X^I$ with $\alpha(1) = \beta(0)$, by $\alpha \star \beta$ we mean the path in X defined by

$$(\alpha \star \beta)(t) = \begin{cases} \alpha(2t), & 0 \leq t \leq \frac{1}{2}; \\ \beta(2t - 1), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

Definition 1.1. [9] For two continuous functions (or maps) $f, g : X \rightarrow Y$, the function f is called a *homotopic* to g and write $f \simeq g$ if there exists a continuous function (homotopy) $H : X \times I \rightarrow Y$ satisfies $H_0 = f$ and $H_1 = g$.

Generalized Bivariate Blending Type Bernstein Operators by a New Class of Coefficients

Seyed Yashar Zaheriani

Seyed Yashar Zaheriani is with the Eastern Mediterranean University, Cyprus (e-mail: yashar.isfahan@gmail.com)

Abstract

In the current paper we modify the operators, that are defined by Kajla and Acar. we show that our modification gives a better approximation, which is remarkably well. Also we study approximation property of our operators using some unknown sequences as a coefficient on our operators. More over we prove the Korovkin and a Voronoskaja type theorems for these operators. The last part is concerned fulfilling the GBS operators and the degree of approximation . Finally we prove our claim using some numerical examples to show that how our modification works.

AMS Mathematics Subject Classification(2021): 41A25,41A36,40A35

Keywords: Bernstein Polynomial, GBS operators, Korovkin type approximation, degree of approximation.

1 Introduction and Preliminaries

Let f be a function on $C(D^2)$, then the two dimensional Bernstein operators ([1], [7]) are defined by,

$$B_{n,m}(f, x, y) = \sum_{k=0}^n \sum_{t=0}^m f\left(\frac{k}{n}, \frac{t}{m}\right) \binom{n}{k} \binom{m}{t} x^k (1-x)^{n-k} y^t (1-y)^{m-t}, \quad (1)$$

or

$$B_{n,m}(f, x, y) = \sum_{k=0}^n \sum_{t=0}^m f\left(\frac{k}{n}, \frac{t}{m}\right) B_{k,n}(x) B_{t,m}(y).$$

($D^2 = [0, 1] \times [0, 1]$) as a special case. Where $C(D^2)$ be the space of all ordinary continuous functions of two variables on D^2 , which is equipped by the norm $\|f\|_{C(D^2)} = \sup_{(x,y) \in D^2} |f(x, y)|$.

Theorem 1 Let $h \in C(D^2)$ be a bivariate positive function and suppose that $L_{n,m}$ for $n, m \geq 1$, are the sequence of positive linear operators such that,

$$\begin{aligned} \lim_{n,m} L_{n,m}(1; x, y) &= g \\ \lim_{n,m} L_{n,m}(t; x, y) &= tg \\ \lim_{n,m} L_{n,m}(s; x, y) &= sg \\ \lim_{n,m} L_{n,m}(t^2; x, y) &= t^2g \\ \lim_{n,m} L_{n,m}(s^2; x, y) &= s^2g \\ \lim_{n,m} L_{n,m}(s^2 + t^2; x, y) &= g(s^2 + t^2). \end{aligned}$$

Uniformly on D^2 .

Proof. Lets consider that,

$$\lim_{n,m} L_{n,m}(f; x, y) = fg.$$

By Korovkin Theorem and taking the linear positive sequences $\frac{1}{g}L_{n,m}$ for any $n, m \geq 1$, the proof is complete. ■

Generalized bivariate blending type Bernstein operators by a new class of coefficients

For any $f \in C(D^2)$ and $0 \leq \alpha_1, \alpha_2 \leq 1$, bivariate modified extension of the operator (1) is defined by

$$U_{n,m}^{M,\alpha_1,\alpha_2}(f; x, y) = \sum_{k=0}^n \sum_{j=0}^m J_{n,m,k,j}^{M,\alpha_1,\alpha_2}(x, y) f\left(\frac{k}{n}, \frac{j}{m}\right). \tag{2}$$

Where

$$J_{n,m,k,j}^{M,\alpha_1,\alpha_2}(x, y) = [(1 - \alpha_1)a(x, n)B_{k,n-2}(x) + (1 - \alpha_1)a(1 - x, n)B_{k-2,n-2}(x) + \alpha_1 B_{k,n}(x)] \\ \times [(1 - \alpha_2)b(x, n)B_{j,m-2}(x) + (1 - \alpha_2)b(1 - y, m)B_{j-2,m-2}(y) + \alpha_2 B_{j,m}(y)], \tag{3}$$

and $a(x, n) = a_1(n)x + a_0(n)$, $b(x, n) = b_1(m)y + b_0(m)$, where $a_i(n), b_j(m), i, j = 0, 1, n, m = 0, 1, 2, \dots$ are unknown sequences which will be determined later. In this part we want to study the unknown sequences under some conditions on the bivariate functions $f \in C(D^2)$. Asymptotically we have,

$$\lim_{n,m \rightarrow \infty} nm(U_{n,m}^{M,\alpha_1,\alpha_2}(f; x, y) - f(x, y)) = C(f; x, y). \tag{4}$$

Where $C(f; x, y)$ is a form depends on f_x, f_y, f_{xx} or f_{yy} . If relation (4) holds true, so clearly $a_1(n) + 2a_0(n) = 1$ and $b_1(m) + 2b_0(m) = 1$, using the given fact the following are considered,

- (condition1) If $a_1(n) = b_1(m) = 0$, then $a_0(n) = b_0(m) = \frac{1}{2}$.
- (condition2) If $0 < a_1(n), b_1(m) < 1$, then $a_0(n), b_0(m) > 0$.
- (condition3) If $a_1(n) = b_1(m) = 1$, then $a_0(n) = b_0(m) = 0$.
- (condition4) If $a_1(n), b_1(m) > 1$, then $a_0(n), b_0(m) < 0$.
- (condition5) If $-1 < a_1(n), b_1(m) < 0$, then $a_0(n), b_0(m) > 0$ and $a_1(n) + a_0(n) > 0, b_1(m) + b_0(m) > 0$.
- (condition6) If $a_1(n) = b_1(m) = -1$, then $a_0(n) = b_0(m) = 1$.
- (condition7) If $a_1(n), b_1(m) < -1$, then $a_0(n), b_0(m) > 0$ and $a_1(n) + a_0(n) < 0, b_1(m) + b_0(m) < 0$.

condition5 reduces $U_{n,m}^{M,\alpha_1,\alpha_2}$ to $U_{n,m}^{\alpha_1,\alpha_2}$. Also it can be easily verify that $U_{n,m}^{M,\alpha_1,\alpha_2}$ is positive for all parts except (condition4) and 7. The following Lemma 1, Corollary 1 and Theorem 2 are proved for condition1,2,3,5 and 6. Please note that through out of the paper all of the mentioned unknown sequences, $a_0(n), a_1(n), b_0(m)$ and $b_1(m)$ for all $n, m \geq 1$, are considered as bounded in all conditions1-7.

Lemma 1 For the operators $U_{n,m}^{M,\alpha_1,\alpha_2}(f; x, y)$ we have,

- (i) $U_{n,m}^{M,\alpha_1,\alpha_2}(1; x, y) = [(1 - \alpha_1)(a_1(n) + 2a_0(n)) + \alpha_1] \times [(1 - \alpha_2)(b_1(m) + 2b_0(m)) + \alpha_2]$.
- (ii) $U_{n,m}^{M,\alpha_1,\alpha_2}(t; x, y) = [(1 - \alpha_1)x(a_1(n) + 2a_0(n)) + \frac{2}{n}(a_0(n) + a_1(n))(1 - 2x)(1 - \alpha_1) + x\alpha_1] \times [(1 - \alpha_2)(b_1(m) + 2b_0(m)) + \alpha_2]$.
- (iii) $U_{n,m}^{M,\alpha_1,\alpha_2}(s; x, y) = [(1 - \alpha_2)y((b_1(m) + 2b_0(m)) + \frac{2}{m}(b_0(m) + b_1(m))(1 - 2y)(1 - \alpha_2) + y\alpha_2)] \times [(1 - \alpha_1)(a_1(n) + 2a_0(n)) + \alpha_1]$.

$$\begin{aligned}
 (iv) \quad U_{n,m}^{M,\alpha_1,\alpha_2}(t^2; x, y) &= [(1-\alpha_1)[x^2 \frac{(n-2)(n-3)}{n^2}(a_1(n)+2a_0(n)) + \frac{n-2}{n^2}a_1(n)x^2 + x \frac{(n-2)}{n^2}(a_0(n)-5a_1(n)) \\
 &\quad -xa_1(n)\frac{4}{n^2} + (\frac{4}{n^2} + \frac{5x(n-2)}{n^2})(a_0(n)+a_1(n))] + \frac{\alpha_1(n-1)x^2}{n} + \frac{\alpha_1x}{n} \times [(1-\alpha_2)(b_1(m)+2b_0(m)) + \alpha_2]. \\
 (v) \quad U_{n,m}^{M,\alpha_1,\alpha_2}(s^2; x, y) &= [(1-\alpha_2)[y^2 \frac{(m-2)(m-3)}{m^2}(b_1(m)+2b_0(m)) + \frac{m-2}{m^2}b_1(m)y^2 + y \frac{(m-2)}{m^2}(b_0(m)- \\
 &\quad 5b_1(m))-yb_1(m)\frac{4}{m^2} + (\frac{4}{m^2} + \frac{5y(m-2)}{m^2})(b_0(n)+b_1(m))] + \frac{\alpha_1(m-1)y^2}{m} + \frac{\alpha_2y}{m} \times [(1-\alpha_1)(a_1(n)+2a_0(n)) + \\
 &\quad \alpha_1]. \\
 (vi) \quad U_{n,m}^{M,\alpha_1,\alpha_2}(t^3; x, y) &= [(1-\alpha_1)[\frac{x^3(n-2)(n-3)(n-4)}{n^3}(a_1(n)+2a_0(n)) - \frac{5x^3a_1(n)(n-2)(n-3)}{n^3} + \frac{11x^2a_0(n)(n-2)(n-3)}{n^3} \\
 &\quad - \frac{11x^2a_1(n)(n-2)}{n^3} + \frac{12x(n-2)}{n^3}(a_1(n)+a_0(n)) + \frac{a_0(n)x(n-2)}{n^3} + \frac{8a_0(n)}{n^3} - \frac{a_1(n)8x}{n^3} + \frac{8x^2(n-2)(n-3)a_1(n)}{n^3}] + \frac{\alpha_1x}{n^2} \\
 &\quad + \frac{3x^2(n-1)}{n^2} + \frac{\alpha_1x^3(n-1)(n-2)}{n^2}] \times [(1-\alpha_2)(b_1(m)+2b_0(m)) + \alpha_2] \\
 (vii) \quad U_{n,m}^{M,\alpha_1,\alpha_2}(s^3; x, y) &= [(1-\alpha_2)[\frac{y^3(m-2)(m-3)(m-4)}{m^3}(b_1(m)+2b_0(m)) - \frac{5y^3b_1(m)(m-2)(m-3)}{m^3} + \frac{11y^2b_0(m)(m-2)(m-3)}{m^3} \\
 &\quad - \frac{11y^2b_1(m)(m-2)}{m^3} + \frac{12y(m-2)}{m^3}(b_1(m)+b_0(m)) + \frac{b_0(m)y(m-2)}{m^3} + \frac{8b_0(m)}{m^3} - \frac{b_1(m)8y}{m^3} + \frac{8y^2(m-2)(m-3)b_1(m)}{m^3}] + \\
 &\quad \frac{\alpha_1y}{m^2} + \frac{3y^2(m-1)}{m^2} + \frac{\alpha_1y^3(m-1)(m-2)}{m^2}] \times [(1-\alpha_1)(a_1(n)+2a_0(n)) + \alpha_1]. \\
 (viii) \quad U_{n,m}^{M,\alpha_1,\alpha_2}(t^4; x, y) &= [[(1-\alpha_1)[\frac{x^4(n-2)(n-3)(n-4)(n-5)}{n^4}(a_1(n)+2a_0) + \\
 &\quad +a_1(n)(\frac{7x^3(n-2)(n-3)+6x^4(n-2)(n-3)(n-4)-56x^2(n-2)+34x^2(n-2)(n-3)+11x^3(n-2)(n-3)(n-4)-11x^4(n-2)(n-3)(n-4)}{n^4}) \\
 &\quad +a_1(n)(\frac{-34x^4(n-2)(n-3)-16x}{n^4})+(a_1+a_0)(\frac{16+57x(n-2)}{n^4})+a_0(\frac{(n-2)x+17x^3(n-2)(n-3)(n-4)+41x^2(n-2)(n-3)}{n^4}) \\
 &\quad +\alpha_1(\frac{7x^2(n-1)+9x^3(n-1)(n-2)+x^4(n-1)(n-2)(n-3)}{n^3}]] \times [(1-\alpha_2)(b_1(m)+2b_0(m)) + \alpha_2]. \\
 (viiii) \quad U_{n,m}^{M,\alpha_1,\alpha_2}(s^4; x, y) &= [[(1-\alpha_2)[\frac{y^4(m-2)(m-3)(m-4)(m-5)}{m^4}(b_1(m)+2b_0) + \\
 &\quad +b_1(m)(\frac{7y^3(m-2)(m-3)+6y^4(m-2)(m-3)(m-4)-56y^2(m-2)+34y^2(m-2)(m-3)+11y^3(m-2)(m-3)(m-4)}{m^4}) \\
 &\quad +b_1(m)(\frac{-34y^4(m-2)(m-3)-16y-11y^4(m-2)(m-3)(m-4)}{m^4})+(b_1+b_0)(\frac{16+57y(m-2)}{m^4}) \\
 &\quad +b_0(\frac{(m-2)y+17y^3(m-2)(m-3)(m-4)+41y^2(m-2)(m-3)}{m^4}) \\
 &\quad +\alpha_2(\frac{7y^2(n-1)+9y^3(m-1)(m-2)+y^4(m-1)(m-2)(m-3)}{m^3}]] \times [(1-\alpha_1)(a_1(n)+2a_0(n)) + \alpha_1]
 \end{aligned}$$

Proof. Lets just prove part (i) and (iv) then the rest are similarly to find. As,

$$\sum_{k=0}^{n-2} \frac{k}{n} B_{k,n-2} = \frac{(n-2)}{n}x, \quad \sum_{k=0}^n \frac{k}{n} B_{k,n} = x, \quad \sum_{k=2}^{n-2} B_{k-2,n-2} = \frac{2}{n} + \frac{(n-2)x}{n}.$$

After replacing in (3) and some calculation it is clear to see that, the part (i) is held. To prove part (iv) we have,

$$\begin{aligned}
 \sum_{k=0}^{n-2} \frac{k^2}{n^2} B_{k,n-2} &= \frac{(n-2)(n-3)}{n^2}x^2 + \frac{(n-2)}{n^2}x, & \sum_{k=0}^n \frac{k^2}{n^2} B_{k,n} &= \frac{(n-1)}{n}x^2 + \frac{1}{n}x, \\
 \sum_{k=2}^{n-2} B_{k-2,n-2} &= \frac{4}{n^2} + \frac{5(n-2)x}{n^2} + \frac{(n-2)(n-3)x^2}{n^2}.
 \end{aligned}$$

After replacing in (3) and some easy calculations it is easy the proof is completed. ■

Theorem 2 Consider $a_1(n), b_1(m), a_0(n)$ and $b_0(m)$ be four sequences for which the operators $U_{n,m}^{M,\alpha_1,\alpha_2}$ is positive(conditions 1,2,3,5 and 6 are held), if $f \in C(D^2)$ then;

$$\lim_{n,m \rightarrow \infty} \|U_{n,m}^{M,\alpha_1,\alpha_2} - f\|_{C(D^2)} = 0$$

Proof. The conditions 1,2,3,5 and 6 are satisfied then $|a_i(n)| \leq M$ and $|b_j(m)| \leq M_1$ for $i, j = 0, 1$ and also by Lemma 2 we have,

$$\begin{aligned} \lim_{n,m \rightarrow \infty} U_{n,m}^{M,\alpha_1,\alpha_2}(1; x, y) &= 1, & \lim_{n,m \rightarrow \infty} U_{n,m}^{M,\alpha_1,\alpha_2}(t; x, y) &= x \\ \lim_{n,m \rightarrow \infty} U_{n,m}^{M,\alpha_1,\alpha_2}(s; x, y) &= y, & \lim_{n,m \rightarrow \infty} U_{n,m}^{M,\alpha_1,\alpha_2}(t^2; x, y) &= x^2 \\ \lim_{n,m \rightarrow \infty} U_{n,m}^{M,\alpha_1,\alpha_2}(s^2; x, y) &= y^2, & \lim_{n,m \rightarrow \infty} U_{n,m}^{M,\alpha_1,\alpha_2}(t^2 + s^2; x, y) &= x^2 + y^2. \end{aligned}$$

uniformly on D^2 , using Korovkin type approximation theorem for any function $f \in C(D^2)$, the proof is complete. ■

Corollary 1 *If $a_1(n) + 2a_0(n) = 1$ and $b_1(m) + 2b_0(m) = 1$, then*

$$\begin{aligned} U_{n,m}^{M,\alpha_1,\alpha_2}((t-x)^2; x, y) &\leq \frac{2(1+a_1(n))(5x^2-x+1)}{n} \\ U_{n,m}^{M,\alpha_1,\alpha_2}((s-y)^2; x, y) &\leq \frac{2(1+b_1(m))(5y^2-y+1)}{m} \end{aligned}$$

Proof.

$$\begin{aligned} (1-\alpha_1) \left[x^2 \frac{(n-2)(n-3)}{n^2} + \frac{n-2}{n^2} a_1(n) x^2 + x \frac{(n-2)}{n^2} (a_0(n) - 5a_1(n)) \right. \\ \left. - x a_1(n) \frac{4}{n^2} + \left(\frac{4}{n^2} + \frac{5x(n-2)}{n^2} \right) (a_0(n) + a_1(n)) \right] + \frac{\alpha_1(n-1)x^2}{n} + \frac{\alpha_1 x}{n} \\ - 2x^2 + \frac{-2x(1+a_1(n))(1-2x)(1-\alpha_1)}{n} \end{aligned}$$

After simplifying and taking $a_0(n) = \frac{1-a_1(n)}{2}$ easily it can be seen that

$$\begin{aligned} (1-\alpha_1) \left[\frac{6x^2}{n^2} + a_1(n) \left(\frac{x(x-3)}{n} + \frac{2(-x^2+x+1)}{n^2} \right) + \frac{2x(2x-1)}{n} + \frac{2a_1(n)x(2x-1)}{n} \right] \\ - \frac{x^2}{n} + \frac{x\alpha_1}{n} + \left(\frac{3x}{n} + \frac{2(1-3x)}{n^2} \right) (1-\alpha_1) \\ \leq \frac{6x^2}{n} + a_1(n) \left(\frac{x(x-3)}{n} + \frac{2(-x^2+x+1)}{n} + \frac{2x(2x-1)}{n} \right) + \frac{2x(2x-1)}{n} + \frac{4x}{n} + \frac{2(1-3x)}{n} \\ \leq \frac{a_1(n)(3x^2-3x+2)}{n} + \frac{10x^2-2x+2}{n} \leq \frac{2(1+a_1(n))(5x^2-x+1)}{n} \end{aligned}$$

■

Theorem 3 *Let $f \in C(D^2)$ and suppose that the operator $U_{n,m}^{M,\alpha_1,\alpha_2}$ is positive (conditions 1,2,3,5 and 6 are hold) so,*

(a) *If $a_1(n) + a_0(n) = 0$ and $b_1(n) + b_0(n) = 0$, ($n = 0, 1, 2, \dots$) and second partial derivatives f_{xx}, f_{yy}, f_{xy} exist then, we have*

$$\begin{aligned} \lim_{n \rightarrow \infty} n(U_{n,n}^{M,\alpha_1,\alpha_1}(f; x, y) - f(x, y)) &= \frac{1}{2} \left((1-\alpha_1) [a_1(n)(-4x^2-5x) + a_0(n)(-10x^2+x)] \right. \\ &\quad \left. + \alpha_1 x + x^2 \alpha_1 \right) f_{xx} \\ &\quad + \frac{1}{2} \left((1-\alpha_1) [b_1(n)(-4y^2-5y) + b_0(n)(-10y^2+y)] \right. \\ &\quad \left. + \alpha_1 y + y^2 \alpha_1 \right) f_{yy}. \end{aligned}$$

(b) If $a_1(n) + a_0(n) \neq 0$ and $b_1(n) + b_0(n) \neq 0$, ($n = 0, 1, 2, \dots$) and the first partial derivatives f_x, f_y just exist, in the way that there is no information about the second partial derivatives of f for any $(x, y) \in D^2$, we can get the following

$$\begin{aligned} \lim_{n \rightarrow \infty} n(U_{n,n}^{M,\alpha_1,\alpha_1}(f; x, y) - f(x, y)) &= \frac{2}{n}(a_0(n) + a_1(n))(1 - 2x)(1 - \alpha_1)f_x \\ &+ \frac{2}{n}(b_0(n) + b_1(n))(1 - 2y)(1 - \alpha_1)f_y \end{aligned}$$

Proof. Lets prove part (a) and the proof of second part is similar. For any $(x, y) \in D^2$, by the Taylor expansion formula we have

$$\begin{aligned} f(t, s) &= f(x, y) + f_x(x, y)(t - x) + f_y(x, y)(s - y) \\ &+ \frac{1}{2} \{ f_{xx}(x, y)(t - x)^2 + 2f_{xy}(x, y)(t - x)(s - y) + f_{yy}(x, y)(s - y)^2 \} \\ &+ \sigma(t, s; x, y) \sqrt{(t - x)^4 + (s - y)^4}. \end{aligned}$$

For any $(x, y) \in D^2$, where $\sigma(t, s; x, y) \in C(D^2)$ we have $\sigma(t, s; x, y) \rightarrow 0$, while $(t, s) \rightarrow (x, y)$. By applying $U_{n,n}^{M,\alpha_1,\alpha_1}$ on the given Taylor expansion we have,

$$\begin{aligned} U_{n,n}^{M,\alpha_1,\alpha_1}(f; x, y) &= f(x, y) + f_x(x, y)U_{n,n}^{M,\alpha_1,\alpha_1}((t - x); x, y) + f_y(x, y)U_{n,n}^{M,\alpha_1,\alpha_1}((s - y); x, y) \\ &+ \frac{1}{2} \{ f_{xx}(x, y)U_{n,n}^{M,\alpha_1,\alpha_1}((t - x)^2; x, y) + 2f_{xy}(x, y)U_{n,n}^{M,\alpha_1,\alpha_1}((t - x)(s - y); x, y) \} \\ &+ \frac{1}{2} \{ f_{yy}(x, y)U_{n,n}^{M,\alpha_1,\alpha_1}((s - y)^2; x, y) \} + U_{n,n}^{M,\alpha_1,\alpha_1}(\sigma(t, s; x, y) \sqrt{(t - x)^4 + (s - y)^4}; x, y). \end{aligned}$$

On the other hand by properties of part (a), Lemma 1, parts (i) and (ii) we have,

$$U_{n,n}^{M,\alpha_1,\alpha_1}((t - x); x, y) = 0,$$

also we should notice that similarly we get,

$$U_{n,n}^{M,\alpha_1,\alpha_1}((t - x)(s - y); x, y) = (U_n^{M,\alpha_1}((t - x); x), U_n^{M,\alpha_1}((s - y); y)) = 0.$$

From the Cauchy Schwarz inequality we obtain

$$\begin{aligned} &n U_{n,n}^{M,\alpha_1,\alpha_1}(\sigma(t, s; x, y) \sqrt{(t - x)^4 + (s - y)^4}; x, y) \\ &\leq \sqrt{U_{n,n}^{M,\alpha_1,\alpha_1}(\sigma(t, s; x, y); x, y)} \times \sqrt{n^2 U_n^{M,\alpha_1}((t - x)^4; x) + n^2 U_n^{M,\alpha_1}((s - y)^4; y)}. \end{aligned}$$

Using the given fact that $\sigma(t, s; x, y) \in C(D^2)$ we have $\sigma(t, s; x, y) \rightarrow 0$, while $(t, s) \rightarrow (x, y)$ and also by the Theorem 2 we deduce that,

$$\lim_{n \rightarrow \infty} U_{n,n}^{M,\alpha_1,\alpha_1}(\sigma(t, s; x, y); x, y) = 0,$$

for any $(x, y) \in D^2$ approaches to zero uniformly. By the corollary 4 and the fact that $|b_j(n)|, |a_i(n)| \leq M, (i, j = 0, 1)$, we may say that,

$$\begin{aligned} n^2 U_n^{M,\alpha_1}((t - x)^4; x) &\leq 4(1 + a_1(n))^2(5x^2 - x + 1)^2 \\ n^2 U_n^{M,\alpha_1}((s - y)^4; y) &\leq 4(1 + b_1(n))^2(5y^2 - y + 1)^2. \end{aligned}$$

Therefore

$$\lim_{n \rightarrow \infty} n \left(U_{n,n}^{M,\alpha_1,\alpha_1}(\sigma(t, s; x, y) \sqrt{(t - x)^4 + (s - y)^4}; x, y) \right) = 0.$$

So the proof is complete. ■

Remark 1 *The beneficial of using this method is that respect to the pervious theorem while $a_1(n) + a_0(n) \neq 0$ and $b_1(n) + b_0(n) \neq 0$ the existance of f_{xx}, f_{yy} is no longer important for any certain point $(x, y) \in D^2$.*

Consider the space H_{w_2} of real-valued functions, defined on D^2 and satisfying,

$$|f(u, v) - f(x, y)| \leq w_2(f; |\frac{u}{1+u} - \frac{x}{1+x}|, |\frac{v}{1+v} - \frac{y}{1+y}|).$$

Where w_2 is a non-negative function on D^2 , which is increasing for both variables and satisfying;

- (i) $w_2(f; \delta_1 + \delta_2, \delta) \leq w_2(f; \delta_1, \delta) + w_2(f; \delta_2, \delta)$.
- (ii) $w_2(f; \delta, \delta_1 + \delta_2) \leq w_2(f; \delta, \delta_1) + w_2(f; \delta, \delta_2)$.
- (iii) $\lim_{\delta_1 \rightarrow 0, \delta_2 \rightarrow 0} w_2(f; \delta_1, \delta_2) = 0$.

Let say that the $f(x, y) \in C(D^2)$ such that $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial^2 f}{\partial x^2}$ and $\frac{\partial^2 f}{\partial y^2}$ are in $C(D^2)$. The norm of $C(D^2)$ is defined as following,

$$\|f\|_{C^2(D^2)} = \|f\|_{C(D^2)} + \sum_{i=1}^2 (\|\frac{\partial^i f}{\partial x^i}\|_{C(D^2)} + \|\frac{\partial^i f}{\partial y^i}\|_{C(D^2)}), \tag{5}$$

The peeters $K - function$ of the function $f \in C(D^2)$ is defined by

$$K(f; \delta) = \inf \{ \|f - g\|_{C(D^2)} + \delta \|g\|_{C^2(D^2)}, \delta > 0 \},$$

where $g \in C^2(D^2)$. Also the following inequality is hold for any $\delta > 0$.

$$K(f; \delta) \leq L (\omega_2(f; \sqrt{\delta})). \tag{6}$$

The constant L is independent of δ and f . It is obvious that ω_2 is the second order of continuity.

Theorem 4 *If $f(x, y)$ is bound for any $(x, y) \in D^2$, then*

$$|U_{n,m}^{M, \alpha_1, \alpha_2}(f; x, y) - f(x, y)| \leq C \omega_2(f; \sqrt{\delta_{n,m}(x, y)}).$$

Where $\delta_{n,m}(x, y) = \delta_n^2(x) + \delta_m^2(y)$, $\delta_n^2 = \frac{10x^2 - 5x + 2}{n}(1 + a_1(n))$ and $\delta_m^2 = \frac{10y^2 - 5y + 2}{m}(1 + b_1(m))$

Suppose that $h \in C^2(D^2)$ and $t, s \in D^2$. Using Teylor's theorem we can write

$$\begin{aligned} h(t, s) - h(x, y) &= \frac{\partial h(x, y)}{\partial x}(t - x) + \int_x^t (t - u) \frac{\partial^2 h(u, y)}{\partial u^2} du \\ &+ \frac{\partial h(x, y)}{\partial x}(s - y) + \int_y^s (s - v) \frac{\partial^2 h(x, v)}{\partial v^2} dv. \end{aligned}$$

After applying $U_{n,m}^{M,\alpha_1,\alpha_2}$ on the given equation and by (5) we have,

$$\begin{aligned}
 & |U_{n,m}^{M,\alpha_1,\alpha_2}(f; x, y) - f(x, y)| = |h_x(\frac{2}{n}(a_0(n) + a_1(n))(1 - 2x)(1 - \alpha_1)) \\
 & + U_{n,m}^{M,\alpha_1,\alpha_2}(\int_x^t (t - u) \frac{\partial^2 h(u, y)}{\partial u^2} du) + h_y(\frac{2}{m}(b_0(m) + b_1(m))(1 - 2y)(1 - \alpha_2)) \\
 & + U_{n,m}^{M,\alpha_1,\alpha_2}(\int_y^s (s - y) \frac{\partial^2 h(x, v)}{\partial v^2} dv)| \leq |h_x(\frac{2}{n}(a_0(n) + a_1(n))(1 - 2x)(1 - \alpha_1))| \\
 & + |h_y(\frac{2}{m}(b_0(m) + b_1(m))(1 - 2y)(1 - \alpha_2))| + U_{n,m}^{M,\alpha_1,\alpha_2}(|\int_x^t |t - u| \frac{\partial^2 h(u, y)}{\partial u^2} dv|) \\
 & + U_{n,m}^{M,\alpha_1,\alpha_2}(|\int_y^s |s - y| \frac{\partial^2 h(x, v)}{\partial v^2} dv|) \leq |h_x(\frac{2}{n}(a_0(n) + a_1(n))(1 - 2x)(1 - \alpha_1))| \\
 & + |h_y(\frac{2}{m}(b_0(m) + b_1(m))(1 - 2y)(1 - \alpha_2))| + (U_{n,m}^{M,\alpha_1,\alpha_2}((t - x)^2; x, y)) \|h\|_{C^2(D^2)} \\
 & + (U_{n,m}^{M,\alpha_1,\alpha_2}((s - y)^2; x, y)) \|h\|_{C^2(D^2)}.
 \end{aligned}$$

Using $a_0(n) = \frac{1-a_1(n)}{2}$, $b_0(m) = \frac{1-b_1(m)}{2}$, (5) and corollary 4 clearly we have,

$$\begin{aligned}
 & |h_x(\frac{2}{n}(a_0(n) + a_1(n))(1 - 2x)(1 - \alpha_1))| + |h_y(\frac{2}{m}(b_0(m) + b_1(m))(1 - 2y)(1 - \alpha_2))| \\
 & + (U_{n,m}^{M,\alpha_1,\alpha_2}((t - x)^2; x, y)) \|h\|_{C^2(D^2)} + (U_{n,m}^{M,\alpha_1,\alpha_2}((s - y)^2; x, y)) \|h\|_{C^2(D^2)} \\
 & \leq \|h\|_{C^2(D^2)} (\frac{1 + a_1(n)}{2}) + \|h\|_{C^2(D^2)} (\frac{1 + b_1(m)}{2}) \leq |h_x(\frac{2}{n}(a_0(n) + a_1(n))(1 - 2x)(1 - \alpha_1))| \\
 & + |h_y(\frac{2}{m}(b_0(m) + b_1(m))(1 - 2y)(1 - \alpha_2))| + \|h\|_{C^2(D^2)} ((\frac{1 + a_1(n)}{n})(10x^2 - 2x + 2)) \\
 & + \|h\|_{C^2(D^2)} ((\frac{1 + b_1(m)}{m})(10y^2 - 2y + 2)) \\
 & \leq 3\|h\|_{C^2(D^2)} \left(\frac{1 + a_1(n)}{n}(10x^2 - 2x + 2) + \frac{1 + b_1(m)}{m}(10y^2 - 2y + 2) \right).
 \end{aligned}$$

It is while for all $f \in C(D^2)$, $h \in C^2(D^2)$ and by (6) we have,

$$\begin{aligned}
 |U_{n,m}^{M,\alpha_1,\alpha_2}(f; x, y) - f(x, y)| & \leq |U_{n,m}^{M,\alpha_1,\alpha_2}(f - h; x, y)| + |f(x, y) - h(x, y)| + |U_{n,m}^{M,\alpha_1,\alpha_2}(h; x, y) - h(x, y)| \\
 & \leq 3\|h\|_{C^2(D^2)}(1 + M)(\delta_n^2 + \delta_m^2) + 2\|f - g\|_{C^2(D^2)} \leq 3K(f; \delta_{n,m}(x, y)) \\
 & \leq C\omega_2(f; \sqrt{\delta_{n,m}}).
 \end{aligned}$$

In this part we consider that the conditions 4 and 7 are held, so using this two properties we continue to prove the following theorems.

Theorem 5 *If $f \in C(D^2)$, then for any sequences $a_0(n), a_1(n), b_1(m)$ and $b_0(m)$ satisfying either conditions 4 and 7 we have,*

$$\lim_{n,m} U_{n,m}^{M,\alpha_1,\alpha_2}(f; x, y) = f(x, y),$$

uniformly on D^2 .

Proof. Let $a_1(n), b_1(m) > 0$ and $a_0(n), b_0(m) < 0$ (condition 4). It is easy to verify that $D_1 - D_2 = U_{n,m}^{M,\alpha_1,\alpha_2}$, where

$$D_1(f; x, y) = \left(\sum_{j=0}^m \sum_{k=0}^n [(1 - \alpha_1)(xa_1(n)B_{k,n-2}(x) + a_1(n)B_{k-2,n-2}(x)) + \alpha_1 B_{k,n}(x)] \times \right. \\ \left. [(1 - \alpha_2)(yb_1(m)B_{j,m-2}(y) + b_1(m)B_{j-2,m-2}(y)) + \alpha_2 B_{j,m}(y)] \right. \\ \left. + \sum_{j=0}^m \sum_{k=0}^n [(1 - \alpha_1)(xa_1(n)B_{k,n-2}(x) + a_1(n)B_{k-2,n-2}(x)) + \alpha_1 B_{k,n}(x)] \times \right. \\ \left. [(1 - \alpha_2)(-b_0(m)B_{j,m-2}(y) + (b_1(m)y - b_0(m))B_{j-2,m-2}(y))] \right) f\left(\frac{k}{n}, \frac{j}{m}\right)$$

and

$$D_2(f; x, y) = \left(\sum_{j=0}^m \sum_{k=0}^n [(1 - \alpha_1)(-a_0(n)B_{k,n-2}(x) + (a_1(n)x - a_0(n))B_{k-2,n-2}(x))] \times \right. \\ \left. [(1 - \alpha_2)(yb_1(m)B_{j,m-2}(y) + b_1(m)B_{j-2,m-2}(y)) + \alpha_2 B_{j,m}(y)] \right. \\ \left. + \sum_{j=0}^m \sum_{k=0}^n [(1 - \alpha_1)(-a_0(n)B_{k,n-2}(x) + (a_1(n)x - a_0(n))B_{k-2,n-2}(x))] \times \right. \\ \left. [(1 - \alpha_2)(-b_0(m)B_{j,m-2}(y) + (b_1(m)y - b_0(m))B_{j-2,m-2}(y))] \right) f\left(\frac{k}{n}, \frac{j}{m}\right).$$

So clearly they are positive operators. Now we have the following,

$$\begin{aligned}
 D_1(1; x, y) &= [(1 - \alpha_1)(1 + x)a_1(n) + \alpha_1] [(1 - \alpha_2)(1 + y)b_1(m) + \alpha_2] + [(1 - \alpha_2)(-2b_0(m) + yb_1(m))] \\
 D_1(t; x, y) &= [(1 - \alpha_1)(x^2 a_1(n)(\frac{n-2}{n}) + a_1(n)(\frac{2}{n} + \frac{(n-2)x}{n})) + x\alpha_1] \times \\
 &\quad ([(1 - \alpha_2)(1 + y)b_1(m) + \alpha_2] + [(1 - \alpha_2)(-2b_0(m) + yb_1(m))]) \\
 D_1(t^2; x, y) &= [(1 - \alpha_1)(a_1(n)\frac{(n-2)(n-3)x^3}{n^2} + \frac{(n-2)x^2}{n^2}) + a_1(n)(\frac{4}{n^2} + \frac{5(n-2)x}{n^2} + \frac{(n-2)(n-3)x^3}{n^2})] \times \\
 &\quad ([(1 - \alpha_2)(1 + y)b_1(m) + \alpha_2] + [(1 - \alpha_2)(-2b_0(m) + yb_1(m))]) \\
 D_1(s; x, y) &= [(1 - \alpha_1)(1 + x)a_1(n) + \alpha_1] ([(1 - \alpha_2)(y^2 b_1(m)\frac{(m-2)}{m} + b_1(m)(\frac{2}{m} + \frac{(m-2)y}{m}) + y\alpha_2] + \\
 &\quad [(1 - \alpha_2)(-b_0(m)\frac{(m-2)y}{m} + (b_1(m)y - b_0(m))(\frac{2}{m} + \frac{(m-2)y}{m}))]) \\
 D_1(s^2; x, y) &= [(1 - \alpha_1)(1 + x)a_1(n) + \alpha_1] \times ([(1 - \alpha_2)(b_1(m)\frac{(m-2)(m-3)y^3}{m^2} + \\
 &\quad b_1(m)(\frac{4}{m^2} + \frac{5(m-2)y}{m^2} + \frac{(m-2)(m-3)y^2}{m^2})) + \alpha_2(\frac{(m-1)y^2}{m} + \frac{y}{m})] + \\
 &\quad [(1 - \alpha_2)(-b_0(m)(\frac{(m-2)(m-3)y^2}{m^2} + \frac{(m-2)y}{m^2}) + (b_1(m)y + b_0(m))(\frac{4}{m^2} + \frac{5(m-2)y}{m^2}) + \\
 &\quad \frac{(m-2)(m-3)y^2}{m^2})]) \\
 D_2(1; x, y) &= [(1 - \alpha_1)(-2a_0(n) + a_1(n)x)] ([(1 - \alpha_2)(1 + y)b_1(m) + \alpha_2] + [(1 - \alpha_2)(-2b_0(m) + b_1(m)y)]) \\
 D_2(1; x, y) &= ((1 - \alpha_1)(-a_0(n)(\frac{(n-2)(n-3)x^2}{n^2} + \frac{(n-2)x}{n^2})) + \alpha_1) ([(1 - \alpha_2)(1 + y)b_1(m) + \alpha_2] + \\
 &\quad [(1 - \alpha_2)(-2b_0(m) + b_1(m)y)]) \\
 D_2(t^2; x, y) &= ((1 - \alpha_1)(-a_0(n)(\frac{(n-2)(n-3)x^3}{n^2} + \frac{(n-2)x^2}{n^2}) + (a_1(n)x - a_0(n))(\frac{4}{n^2} + \frac{5(n-2)}{n^2} \\
 &\quad + \frac{x^2(n-2)(n-3)}{n^2})) + \alpha_1) ([(1 - \alpha_2)(1 + y)b_1(m) + \alpha_2] + [(1 - \alpha_2)(-2b_0(m) + b_1(m)y)]) \\
 D_2(s; x, y) &= ((1 - \alpha_1)(-2a_0(n) + a_1(n)x)) ([(1 - \alpha_2)(y^2 b_1(m)\frac{(m-2)}{m} + b_1(m)(\frac{2}{m} + \frac{(m-2)y}{m}) + y\alpha_2] + \\
 &\quad [(1 - \alpha_2)(-b_0(m)\frac{(m-2)y}{m} + (b_1(m)y - b_0(m))(\frac{2}{m} + \frac{(m-2)y}{m}))]) \\
 D_2(s^2; x, y) &= [(1 - \alpha_1)(1 + x)a_1(n) + \alpha_1] ([(1 - \alpha_1)(b_1(m)(\frac{(m-2)(m-3)y^3}{m^2} + \frac{(m-2)y^2}{m^2})) + \\
 &\quad b_1(m)(\frac{4}{m^2} + \frac{5y(m-2)}{m^2} + \frac{y^2(m-2)(m-3)}{m^2})] + \\
 &\quad [(1 - \alpha_2)(-b_0(m)(\frac{(m-2)(m-3)y^2}{m^2} + \frac{(m-2)y}{m^2}) + (b_1(m)y - b_0(m)) \times \\
 &\quad (\frac{4}{m^2} + \frac{5(m-2)y}{m^2} + \frac{(m-2)(m-3)y^2}{m^2}))].
 \end{aligned}$$

Now lets,

$$\lim_{n,m} U_{n,m}^{M,\alpha_1,\alpha_2}(f; x, y) = \lim_{n,m} D_1(f; x, y) - \lim_{n,m} D_2(f; x, y).$$

Lets take $\lim_n a_1(n) = L_1$, $\lim_m b_1(m) = L_2$ and by $a_0(n) = \frac{1-a_1(n)}{2}$ and $b_0(m) = \frac{1-b_1(m)}{2}$ we have,

$$\begin{aligned}
 \lim_{n,m} D_1(f; x, y) &= f(x, y) [(1 - \alpha_1)(1 + x)L_1 + \alpha_1] ([(1 - \alpha_2)(1 + y)L_2 + \alpha_2] + [(1 - \alpha_2)(1 + y)L_1 - 1]) \\
 \lim_{n,m} D_2(f; x, y) &= f(x, y) [(1 - \alpha_1)(1 + x)L_1 - 1] ([-(1 - \alpha_2)(1 + y)L_2 + \alpha_2] + [(1 - \alpha_2)(1 + y)L_1 - 1]).
 \end{aligned}$$

So by theorem 7 the proos is completed. For the condition7 the proof is in the same way. ■

Theorem 6 Let $f \in C(D^2)$ and suppose that the sequences $a_1(n), a_0(n), b_1(n)$ and $b_0(n)$ are bounded and they satisfy in conditions 4 and 7, so we have the following parts,

(a) If $a_1(n) + a_0(n) = 0$ and $b_1(n) + b_0(n) = 0$, ($n = 0, 1, 2, \dots$) and second partial derivatives f_{xx}, f_{yy}, f_{xy} exist then, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} n(U_{n,n}^{M,\alpha_1,\alpha_1}(f; x, y) - f(x, y)) &= \frac{1}{2}((1 - \alpha_1)[a_1(n)(-4x^2 - 5x) + a_0(n)(-10x^2 + x)] \\ &\quad + \alpha_1 x + x^2 \alpha_1) f_{xx} \\ &\quad + \frac{1}{2}((1 - \alpha_1)[b_1(n)(-4y^2 - 5y) + b_0(n)(-10y^2 + y)] \\ &\quad + \alpha_1 y + y^2 \alpha_1) f_{yy}. \end{aligned}$$

(b) If $a_1(n) + a_0(n) \neq 0$ and $b_1(n) + b_0(n) \neq 0$, ($n = 0, 1, 2, \dots$) and the first partial derivatives f_x, f_y just exist, in the way that there is no information about the second partial derivatives of f for any $(x, y) \in D^2$, then we have the following,

$$\begin{aligned} \lim_{n \rightarrow \infty} n(U_{n,n}^{M,\alpha_1,\alpha_1}(f; x, y) - f(x, y)) &= \frac{2}{n}(a_0(n) + a_1(n))(1 - 2x)(1 - \alpha_1)f_x \\ &\quad + \frac{2}{n}(b_0(n) + b_1(n))(1 - 2y)(1 - \alpha_1)f_y \end{aligned}$$

Proof. It is enough to prove the first part, then part b can be concluded similarly. The proof is the same with theorem 4, but it remains to show that,

$$\lim_n n |U_{n,m}^{M,\alpha_1,\alpha_2}(\sigma(t, s; x, y) \sqrt{(t-x)^4 + (s-y)^4})| = 0. \quad (7)$$

By (2) and boundedness of sequences $a_1(n), a_0(n), b_1(n)$ and $b_0(n)$ clearly we have

$$\begin{aligned} |U_{n,n}^{M,\alpha_1,\alpha_2}(\sigma(t, s; x, y) \sqrt{(t-x)^4 + (s-y)^4})| &\leq [(1 - \alpha_1)|a_1(n)x + a_0(n)|B_{k,n-2}(x) \\ &\quad + (1 - \alpha_1)|a_1(n) - xa_1(n) + a_0(n)|B_{k-2,n-2}(x) + \alpha_1 B_{k,n}(x)] \\ &\quad \times [(1 - \alpha_2)|b_1(n)y + b_0(n)|B_{j,n-2}(x) \\ &\quad + (1 - \alpha_2)|b_1(n) - b_1(n)y + b_0(n)|B_{j-2,n-2}(y) + \alpha_2 B_{j,n}(y)] \\ &\quad \times \sigma(t, s; x, y) \sqrt{(t-x)^4 + (s-y)^4}. \end{aligned}$$

Lets take

$$\begin{aligned} &B_{n,n}^{M,\alpha_1,\alpha_2} \\ &= [(1 - \alpha_1)|a_1(n)x + a_0(n)|B_{k,n-2}(x) + (1 - \alpha_1)|a_1(n) - xa_1(n) + a_0(n)|B_{k-2,n-2}(x) + \alpha_1 B_{k,n}(x)] \\ &\quad \times [(1 - \alpha_2)|b_1(n)y + b_0(n)|B_{j,n-2}(x) + (1 - \alpha_2)|b_1(n) - b_1(n)y + b_0(n)|B_{j-2,n-2}(y) + \alpha_2 B_{j,n}(y)]. \quad (8) \end{aligned}$$

It is clear that the operators $B_{n,n}^{M,\alpha_1,\alpha_2}$ is positive, so by Cauchy Schwarz inequality we obtain

$$\begin{aligned} &B_{n,n}^{M,\alpha_1,\alpha_2}(\sigma(t, s; x, y) \sqrt{(t-x)^4 + (s-y)^4}) \\ &\leq (B_{n,n}^{M,\alpha_1,\alpha_2}(\sigma^2(t, s; x, y)))^{1/2} \times (n^2 B_{n,n}^{M,\alpha_1,\alpha_2}((t-x)^4 + (s-y)^4; x, y))^{1/2} \\ &\leq (B_{n,n}^{M,\alpha_1,\alpha_2}(\sigma^2(t, s; x, y)))^{1/2} \times (n^2 B_{n,n}^{M,\alpha_1,\alpha_2}(t-x)^4; x, y) + n^2 B_{n,m}^{M,\alpha_1,\alpha_2}((s-y)^4; x, y))^{1/2} \end{aligned}$$

By easy calculation and using the same method of Corollary 4 we have,

$$\begin{aligned} \lim_{n \rightarrow \infty} B_{n,n}^{M,\alpha_1,\alpha_2}((t-x)^4; x, y) &= ((1 - \alpha_1) \left| \frac{a_1(n)(2x+1) + 1}{2} \right| \left(\frac{10x^2 + 16x}{n} \right) + \frac{\alpha_1 x}{n})^2 \\ &\quad \times ((1 - \alpha_2)(2y|b_1(n)| + 2|b_0(n)| + |b_1(n)| + \alpha_2)^2 \quad (9) \end{aligned}$$

and

$$\begin{aligned} \lim_{n \rightarrow \infty} B_{n,n}^{M,\alpha_1,\alpha_2}((s-y)^4; x, y) &= ((1-\alpha_2) \left| \frac{b_1(n)(2y+1)+1}{2} \right| \left(\frac{10y^2+16y}{n} \right) + \frac{\alpha_2 y}{n})^2 \\ &\times ((1-\alpha_1)(2y|a_1(n)| + 2|a_0(n)| + |a_1(n)| + \alpha_1)^2). \end{aligned}$$

On the other hand

$$\lim_{n \rightarrow \infty} B_{n,n}^{M,\alpha_1,\alpha_2}(\sigma^2(t, s; x, y); x, y) = 0$$

which deduced from the definition of $B_{n,n}^{M,\alpha_1,\alpha_2}$, Korovkin theorem and the fact that

$$\lim_{n \rightarrow \infty} B_{n,n}^{M,\alpha_1,\alpha_2}(\sigma^2(t, s; x, y); x, y) \leq N\sigma^2(t, s; x, y); x, y),$$

where N is a positive integer, that completes the proof. ■

2 Construction of GBS operators of generalized blending type

The concepts of B-continuity and B-differentiability has been initiated by Bögel in [18] and [20]. Later, Dobrescu and Matei [25] obtained the approximating properties of GBS of bivariate Bernstein polynomials by using B-continuity and B-differentiability. Recently, Bögel space gained popularity among researchers and investigated by researchers in different directions and modification on the Bernstein operators have got a special place as a hot topic between the researchers which can be seen in ([8]-[32]), also see [2]. Let $f(x, y)$ be a real valued function which is defined on the compact set $[0, 1] \times [0, 1]$, the mixed difference of f is denoted by $\Delta_{(t,s)}f(x, y)$ and defined as

$$\Delta_{(t,s)}f(x, y) = f(x, y) - f(x, s) - f(t, y) + f(t, s).$$

The function $f(x, y)$ is called B-continuous at (x, y) if

$$\lim_{(t,s) \rightarrow (x,y)} \Delta_{(t,s)}f(x, y) = 0$$

for all $(x, y) \in [0, 1] \times [0, 1]$. Similarly, The real valued function $f(x, y)$ is called B-Bounded on $[0, 1] \times [0, 1]$ if there exists $M > 0$ such that $|\Delta_{(t,s)}f(x, y)| \leq M$, for every $(x, y), (t, s) \in [0, 1] \times [0, 1]$. The set of all B-continuous and B-bounded functions will be denoted by $C_b([0, 1] \times [0, 1])$ and $B_b([0, 1] \times [0, 1])$ respectively. In the present section we introduced following GBS case of the operators $U_{n,m}^{M,\alpha_1,\alpha_2,s_1,s_2}(f; x, y)$,

$$P_{n,m}^{\alpha_1,\alpha_2}(f(t, s); x, y) = U_{n,m}^{M,\alpha_1,\alpha_2,s_1,s_2}(f(x, s) + f(t, y) - f(t, s); x, y).$$

if $n, m \geq 2$.

That is

$$\begin{aligned} &P_{n,m}^{\alpha_1,\alpha_2}(f(t, s); x, y) \\ &= \sum_{k=0}^n \sum_{j=0}^m (1-\alpha_1) \left[a(x, n) \binom{n}{k-2} x^{k-2} (1-x)^{n-k-2} + a(1-x, n) \binom{n-2}{k-2} x^{k-2} (1-x)^{n-k} \right. \\ &\quad \left. + \alpha_1 \binom{n}{k} x^k (1-x)^{n-k} \right] \\ &\times (1-\alpha_2) \left[b(y, m) \binom{m}{j-2} y^{j-2} (1-y)^{m-j-2} + b(1-y, m) \binom{m-2}{j-2} y^{j-2} (1-y)^{m-j} \right. \\ &\quad \left. + \alpha_2 \binom{m}{j} y^j (1-y)^{m-j} \right] \left(f\left(\frac{k}{n}, y\right) + f\left(x, \frac{j}{m}\right) - f\left(\frac{k}{n}, \frac{j}{m}\right) \right). \end{aligned}$$

It should be mention that the operators $P_{n,m}^{\alpha_1,\alpha_2}$ is positive if conditions 1,2,3,5 and 6 are held.

3 Degree of Approximation by $P_{n,m}^{\alpha_1,\alpha_2}(f(t,s);x,y)$

The mixed modulus of smoothness of a real valued function $f(x,y)$ is denoted by $\omega_{mixed}(f, \delta_1, \delta_2)$ and defined as,

$$\omega_{mixed}(f, \delta_1, \delta_2) := \sup\{|\Delta_{(t,s)} f(x,y)| : |t-x| < \delta_1, |s-y| < \delta_2\}$$

where $(x,y), (t,s) \in [0,1] \times [0,1]$ and $\delta_1, \delta_2 > 0$. It is well known that $\omega_{mixed}(f, \delta_1, \delta_2)$ has similar properties with usual modulus of continuity.

Theorem 7 Let (x,y) be any point in $[0,1] \times [0,1]$, and let $f \in C_b([0,1] \times [0,1])$, for real numbers $0 \leq \alpha_1, \alpha_2 \leq 1$, $n, m \geq 2$ and for conditions 1,2,3,5 and 6, we have

$$|P_{n,m}^{\alpha_1,\alpha_2}(f(t,s);x,y) - f(x,y)| \leq (1 + \sqrt{Z})^2 \omega_{mixed}(f, \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{m}})$$

where Z is a positive constant, and for conditions 4, 7 we have

$$|P_{n,m}^{\alpha_1,\alpha_2}(f(t,s);x,y) - f(x,y)| \leq (1 + \sqrt{Z_1})^2 \omega_{mixed}(f, \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{m}})$$

where Z_1 is a positive constant.

Proof. Recall that $\omega_{mixed}(f, \delta_1, \delta_2)$ has similar properties with usual modulus of continuity, therefore for any positive real numbers λ_1 and λ_2 we have,

$$\omega_{mixed}(f, \lambda_1 \delta_1, \lambda_2 \delta_2) \leq (1 + \lambda_1)(1 + \lambda_2) \omega_{mixed}(f, \delta_1, \delta_2).$$

For any $\delta_1 > 0, \delta_2 > 0$ and for every $(x,y), (t,s) \in [0,1] \times [0,1]$ we can write that,

$$\begin{aligned} \Delta_{(t,s)} f(x,y) &\leq \omega_{mixed}(f, |t-x|, |s-y|) \\ &\leq (1 + \frac{|t-x|}{\delta_1})(1 + \frac{|s-y|}{\delta_2}) \omega_{mixed}(f, \delta_1, \delta_2) \end{aligned}$$

On the other hand by using the definition of $\Delta_{(t,s)} f(x,y)$ we have,

$$f(t,y) + f(x,s) - f(t,s) = f(x,y) - \Delta_{(t,s)} f(x,y)$$

and if we consider conditions 1,2,3,5 and 6,

$$\begin{aligned} &|P_{n,m}^{\alpha_1,\alpha_2}(f(t,s);x,y) - f(x,y)| \\ &\leq T_{n,m}^{M,\alpha_1,\alpha_2}(|\Delta_{(t,s)} f(x,y)|;x,y) \\ &\leq \omega_{mixed}(f, \delta_1, \delta_2) T_{n,m}^{M,\alpha_1,\alpha_2}((1 + \frac{|t-x|}{\delta_1})(1 + \frac{|s-y|}{\delta_2});x,y) \\ &\leq \omega_{mixed}(f, \delta_1, \delta_2) \left(1 + \frac{1}{\delta_1} \sqrt{T_{n,m}^{M,\alpha_1,\alpha_2}((t-x)^2;x,y)} \right. \\ &\quad \left. + \frac{1}{\delta_2} \sqrt{T_{n,m}^{M,\alpha_1,\alpha_2}((s-y)^2;x,y)} \right. \\ &\quad \left. + \frac{1}{\delta_1 \delta_2} \sqrt{T_{n,m}^{M,\alpha_1,\alpha_2}((t-x)^2;x,y) T_{n,m}^{M,\alpha_1,\alpha_2}((s-y)^2;x,y)} \right) \end{aligned}$$

Using Lemma 1, we get,

$$\begin{aligned} &|P_{n,m}^{\alpha_1,\alpha_2}(f(t,s);x,y) - f(x,y)| \\ &\leq \omega_{mixed}(f, \delta_1, \delta_2) \left(1 + \frac{1}{\delta_1} \sqrt{\frac{2(1+a_1(n))(5x^2-x+1)}{n}} + \frac{1}{\delta_2} \sqrt{\frac{2(1+b_1(m))(5y^2-y+1)}{m}} \right. \\ &\quad \left. + \frac{1}{\delta_1 \delta_2} \sqrt{\left(\frac{2(1+a_1(n))(5x^2-x+1)}{n} \right) \left(\frac{2(1+b_1(m))(5y^2-y+1)}{m} \right)} \right) \end{aligned}$$

Now, if we take $\delta_1 = n^{-\frac{1}{2}}$ and $\delta_2 = m^{-\frac{1}{2}}$ we get

$$\begin{aligned} & |P_{n,m}^{\alpha_1, \alpha_2, s_1, s_2}(f(t, s); x, y) - f(x, y)| \\ & \leq \omega_{mixed}(f, n^{-\frac{1}{2}}, m^{-\frac{1}{2}}) \left(1 + \sqrt{2(1 + a_1(n))(5x^2 - x + 1)} + \sqrt{2(1 + b_1(m))(5y^2 - y + 1)} \right. \\ & \quad \left. + 2\sqrt{(1 + a_1(n))(5x^2 - x + 1)(1 + b_1(m))(5y^2 - y + 1)} \right) \\ & \leq (1 + \sqrt{Z})^2 \omega_{mixed}(f, n^{-\frac{1}{2}}, m^{-\frac{1}{2}}) \end{aligned}$$

where $Z := \max\left\{2(1 + a_1(n))(5x^2 - x + 1), 2(1 + b_1(m))(5y^2 - y + 1)\right\}$. Now let the conditions 4,7 are held therefore, using (8) and (9) and the same method, which is used in the first part we can write,

$$|P_{n,m}^{\alpha_1, \alpha_2}(f(t, s); x, y) - f(x, y)| \leq (1 + \sqrt{Z_1})^2 \omega_{mixed}(f, n^{-\frac{1}{2}}, m^{-\frac{1}{2}}),$$

where

$$\begin{aligned} Z_1 := \max \left\{ \right. & ((1 - \alpha_1) \left| \frac{a_1(n)(2x + 1) + 1}{2} \right| (10x^2 + 16x) + \alpha_1 x)^2 \\ & \times ((1 - \alpha_2)(2y|b_1(n)| + 2|b_0(n)| + |b_1(n)| + \alpha_2)^2, \\ & ((1 - \alpha_2) \left| \frac{b_1(n)(2y + 1) + 1}{2} \right| (10y^2 + 16y) + \alpha_2 y)^2 \\ & \left. \times ((1 - \alpha_1)(2y|a_1(n)| + 2|a_0(n)| + |a_1(n)| + \alpha_1)^2 \right\}. \end{aligned}$$

■

4 Numerical examples and related graph

In this section we illustrate a graph to show approximation of the our operators $U_{n,m}^{\alpha_1, \alpha_2}(f; x, y)$ and GBS operators $P_{n,m}^{\alpha_1, \alpha_2}(f; x, y)$. Also we compare it with the operators $T_{n,m}^{\alpha_1, \alpha_2}(f; x, y)$ defined in [26] to a function $f(x, y)$ where $f(x, y) = x \sin(\pi x)y$. In this examples we can observe that the approximation of our modified operators is remarkably better than the operators $T_{n,m}^{\alpha_1, \alpha_2}(f; x, y)$. It should be mention that for unknown sequences two cases have (ii) and (iiv) have been considered as follows,

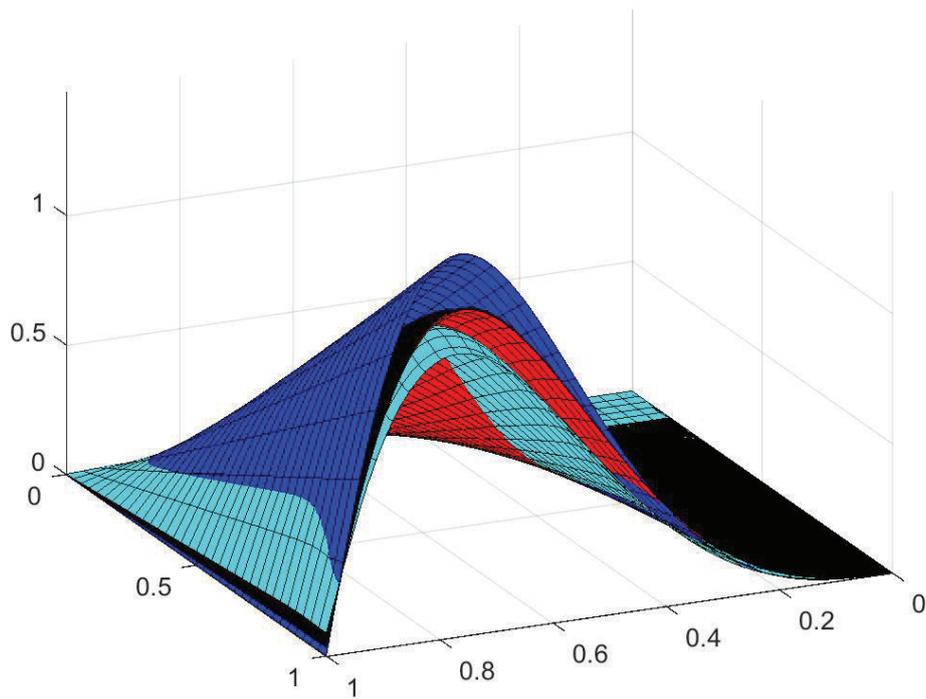


Figure 1: , $n = m = 40, \alpha_1 = \alpha_2 = 0.9$, operators $U_{m,n}^{(\alpha_1, \alpha_2)}$ (red color) For the case (ii) where $a_1 = \frac{1}{n}, b_1 = \frac{1}{m}, a_0 = \frac{1}{2} - \frac{1}{2n}$ and $b_1 = \frac{1}{2} - \frac{1}{2m}$ and for case (vii) applied on GBS operators where $a_0 = \frac{1}{2n} + 1, b_0 = \frac{1}{2m} + 1, a_1 = \frac{-1}{n} - 1$ and $b_1 = \frac{-1}{m} - 1$ (black color) respectively, $T_{m,n}^{(\alpha_1, \alpha_2)}$ (light blue color) the (blue color) comes to the given function f .

References

- [1] Cheney, E. W. and Sharma, A., Bernstein power series. *Can. J. Math.*, 16 (1964), 241-252.
- [2] Jakimowski, A., A generalization of the Lototsky Method of summability. *Michigan Math. J.* 6 (1959), 277-290.
- [3] Jayasri, C., On generalized Lototsky Summability. *Indian J. pure appl. Math.* 13(7), pp. 795-805, 1982.
- [4] King, J. P., Jayasri Summability. *Indian J. pure appl. Math.* 33(6), pp. 797-806, 2002.
- [5] King, J. P., The Lototsky transform and Bernstein Polynomilas. *Canad. J. Math.* 18 (1966), 89-91.
- [6] Korovkin, P., *Linear operatora and approximation theory* (translated from Russian edition of 1959, Delhi, 1960).
- [7] Meyer-König, W. and Zeller, K., Bernsteinsche Potenzreihen, *Studia Math.*, 19, (1960), 89,94.
- [8] Acar, T., Kajla, A., Blending Type Approximation by GBS Operators of Generalized Bernstein-Durrmeyer Type, *Miskolc Mathematical Notes* , 19(1), 319–336 (2018).
- [9] Acar, T., Kajla, A., Degree of Approximation for Bivariate Generalized Bernstein Type Operators, *Results Math.*, 73(2) (2018), doi.org/10.1007/s00025-018-0838-1.
- [10] Agrawal,P.N., Ispir, N., Kajla, A.: GBS operators for Lupa -Durrmeyer, type based on Polyadistributions. *Results Math.* 69, 397-418, (2016).
- [11] Aktuğlu, H., Atamert, M.S., Baytuñç E., Gezer H., Approximation properties of generalized blending type Lototsky-Bernstein operators, submitted.
- [12] Altomere, F., Campiti, M., Korovkin type approximation theory and its applictions, de Gruyter *Stud. Math.* (Berlin: de Gruyter) (1994) vol. 17.
- [13] Aral, A., Erbay, H., Parametric generalization of Baskakov operators, *Mathematical Communications*, 24(1), 119-131, (2019).
- [14] Baxhaku, B., Kajla A.,. Blending type approximation by bivariate generalized Bernstein type operators. *Questiones Math.*, <https://doi.org/10.2989/16073606.2019.1639843>.
- [15] Badea, C., Badea, I., Cottin, C., Gonska, H.H.:Notes on the degree of approximation of B-continuous and B-differentiable functions. *Approx. Theory Appl.* 4, 95-108 (1988)
- [16] Badea, C., Badea, I., Cottin, C.:Korovkin Type theorems for Generalized Boolean Sum Type operators, *Approximation Theory* (Kecskemet, 1900). In :*Colloquia Mathematica Societatis Janos Bolyai*, Amsterdam, vol. 58, pp. 51-68 (1990).
- [17] Bărbosu, D., Acu, A.M., Muraru, C.V.: On certain GBS Durrmeyer operators based on q-integers. *Turk. J. Math.* <https://doi.org/10.3906/mat-1601-34>
- [18] Bögel, K., Über die mehrdimensionale Differentiation, *Integration und beschränkte Variation*, (German)*J. Reine Angew. Math.*, 173 (1935), 5-29.
- [19] Bögel, K., Über die mehrdimensionale Differentiation, *Jahresber. Deutsch Math.-Ver.* 65 (1962), 45-71.
- [20] Bögel, K., Mehrdimensionale Differentiation von Funtionen mehrerer Veränderlicher, *J. Reine Angew. Math.* 170 (1934), 197-217.
- [21] Cai, Q. B., Xu, X. W., Shape-preserving properties of a new family of generalized Bernstein operators, *Journal of Ineq. and Appl.* 2018:241 <https://doi.org/10.1186/s13660-018-1821-9>.
- [22] Cetin, N., Approximation and Geometric Properties of Complex -Bernstein Operator, *Results Math* (2019) 74: 40. <https://doi.org/10.1007/s00025-018-0953-z>.
- [23] Cetin, N, Radu, V, A., Approximation by generalized Bernstein-Stancu operators, *Turkish Journal of Math.*, 43(4), 2032-2048, (2019).
- [24] Chen, X., Tan, J., Liu, Z., Xie, J. Approximation of functions by a new family of generalized Bernstein operators, *J. Math. Anal. Appl.* 450 (2017) 244–261.
- [25] Dobrescu, E., Matei, I., The approximation by Bernstein type polynomials of bidimensionally, continuous functions,(Romanian) *An. Univ. Timișoara Ser Ști. Mat.-Fiz.*, 4 (1996), 85–90.
- [26] Kajla, A., Acar, T., Modified alpha-Bernstein operators with better approximation properties, *Annals of Func. Anal.*, 10(4), 570-582, 2019.
- [27] Kajla, A, Goyal, M.,, Generalized Bernstein-Durrmeyer operators of blending type, *Afrika Matem-*

- atika, 30 (7-8), 1103-1118, (2019).
- [28] Kajla, A., Miclăuș, D., Blending Type Approximation by GBS Operators of Generalized Bernstein-Durrmeyer Type, *Results Math.*, 73(1) (2018), doi.org/10.1007/s00025-018-0773-1.
 - [29] Miclăuș, D.: On the GBS Bernstein-Stancu's type operators. *Creat. Math. Inform.* 22, 73-80 (2013)
 - [30] P.P. Korovkin, On convergence of linear positive operators in the space of continuous functions, *Dokl. Akad. Nauk. SSSR* 90(1953), 961-964.
 - [31] Pop, O.T.: Approximation of B-differentiable functions by GBS operators. *Anal. Univ. Oradea Fasc. Mat* 14,15-31(2007)
 - [32] Pop, O., Bărbosu, D.: GBS operators of Durrmeyer-Stancu type. *Miskolc Math. Notes.* 9,53-60(2008)
 - [33] Davis, P.J., *Interpolation and approximation.* Courier Corporation (1975).

Contribution of the SidePlate Beam-Column Connections to the Seismic Responses of Special Moment Frames

Gökhan Yüksel, Serdar Akça, İlker Kalkan

Abstract— The present study is an attempt to demonstrate the significant levels of contribution of the moment-resisting beam-column connections with side plates to the earthquake behavior of special steel moment frames. To this end, the moment-curvature relationships of a regular beam-column connection and its SidePlate counterpart were determined with the help of finite element analyses. The connection stiffness and deformability values from these finite element analyses were used in the linear time-history analyses of an example structural steel frame under three different seismic excitations. The top-story lateral drift, base shear, and overturning moment values in two orthogonal directions were obtained from these time-history analyses and compared to each other. The results revealed the improvements in the system response with the use of SidePlate connections. The paper ends with crucial recommendations for the plan and design of further studies on this very topic.

Keywords—Seismic detailing, special moment frame, steel structures, beam-column connection, earthquake-resistant design.

I. INTRODUCTION

THE earthquake resistant design of regular residential and commercial structures cannot fully rely on the elastic design philosophy. Resisting the most severe seismic excitations in the elastic range of response is a rather challenging task, since it necessitates the use of large member sizes, particularly for the vertical load-bearing members of the structure. This conservative design is atypical for structures with low degrees of significance due to the high construction costs. In this respect, regular structural systems are expected to dissipate a significant portion of the earthquake-induced free energy within the inelastic range of deformations. Structures susceptible to inelastic material response during the design earthquake need to be provided with sufficient deformation capacity at a particular lateral strength. The adequate deformability of a structure without compensating from its strength to a major level is also denoted as the ductility, which constitutes the leading aspect in the inelastic seismic design of structures.

Unlike lateral strength, providing a structure with sufficient ductility and deformation capacity is an exhausting task due to need for various major and minor details in the system. The ductile behavior of the entire structural frame strongly depends on the resistance and ductility of moment-resisting beam-

column connections. The ductile behavior of the connection regions can only be achieved by surpassing the triaxial state of stress at the surface of contact between the column flange and the beam bottom flange. In this way, the shear stresses, which cater for the plastic deformation capacity of the connection region, can govern the beam-column connection behavior. The devastating effects of this triaxial state of stress were commonly observed in the 1994 Northridge earthquake. These undesired effects were exacerbated in the presence welded flange-bolted web (WFBW) traditional connections [1]. To overcome the brittle response of the connection regions, the ANSI/AISC 358-16 [2] standard presents the so-called prequalified beam-column connections with specific dimensions and details that can easily tolerate significant ground excitations.

The present study pertains to the efficiency of one of the prequalified types of connections in ANSI/AISC 358-16 [2], which is the SidePlate (SP) moment-resisting connection (Fig. 1). This type of connection offers the following advantages as compared to the conventional pre-Northridge WFBW connections:

1. The physical separation of the beam and column prevents the formation of a triaxial stress state at the connection. Many cases of failure in the Northridge earthquake were associated with the triaxial stress concentrations at the joint regions, which is unavoidable in the case of direct contact between the beam and column.
2. The side plates, which sandwich the column, avert the distortion of the panel zone by supporting the column web in resisting the unbalanced shear forces.
3. The extension of the side plates into the beam provides the plastic hinges to be more distant to the column face. In this way, the amount of dissipated energy increases without the need for the enlargement of the beam section.
4. The generality of the welds in this connection type are in parallel direction to the load, and therefore, the ductility of these welds is improved when compared to traditional connections, in which the flange welds are in perpendicular direction to internal forces.
5. Cover plates above and below the beam compensates for the differences between the beam and column widths.
6. Solely using fillet welds in this type of connection prevents the involvement of any concern for the notch effects in the

Gökhan Yüksel is the founder and general director of the Sayhan Engineering, Consulting, Construction and Machinery Limited Company, Ankara, Turkey (e-mail: gokhan@sayhan.com.tr)

Serdar Akça is a civil engineer at the Directorate of Investment Monitoring, General Directorate of TEDAŞ (Turkish Power Distribution Incorporate),

Ministry of Energy and Natural Resources, 06520 Ankara, Turkey. (e-mail: akca_serdar@hotmail.com).

Prof. Dr. İlker Kalkan is with the Department of Civil Engineering, Faculty of Engineering and Architecture, Kırıkkale University, 71450 Kırıkkale, Turkey (e-mail: ilkerkalkan@kku.edu.tr).

connection region.

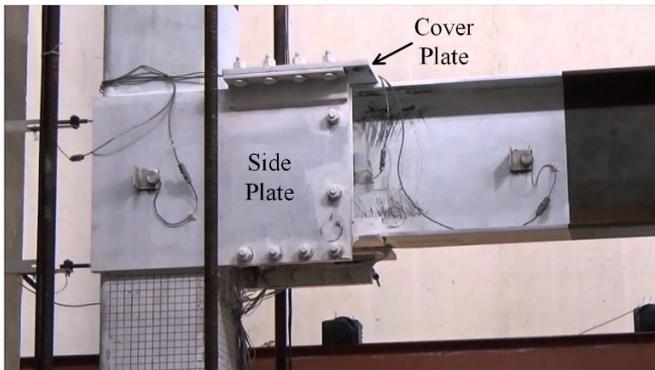


Fig. 1 SidePlate connection [3]

Significant effort has been put forward in the literature to unfold the effects of the SP connection type on the frame behavior. Engelhardt and Sabol [4] underscored the significance of beam cover plates in the joint regions to achieve ductile connection region behavior based on a total of 12 tests on full-scale specimens. Deylami and Ashraf [1] adopted the thickness values of the cover, side and shear plates as the analysis parameters and conducted detailed finite element analyses (FEA) on SP connections. These analyses indicated that SP connections have remarkable energy dissipation capacities and these capacities are not affected by the side plate thickness to a major extent. Chou et al. [5] used side plates to rehabilitate the existing beam-column connections. The tests on unrehabilitated and rehabilitated WFBW specimens depicted the efficiency of side plates to reduce the tensile strains in the bottom flange of the beam in the vicinity of column face. Shiravand and Deylami [6] investigated the applicability of the SidePlate connection details to I-beam to double-I built-up column steel connections. Nonlinear FEA on traditional connections, where the beam is directly welded to the cover plate, and the proposed SidePlate connections indicated that the SidePlate connections can be classified as full strength. The ductility of the connection was shown to exhibit a substantial increase with the use of this new connection type. Jalali et al. [7] conducted nonlinear seismic analyses on steel beams with different height and a side-plate connection subassembly. The connection region was modeled with an elastic panel zone and two degrading rotational springs that represent the critical connection section. Seismic demand hazard curves were established based on the application of the probabilistic seismic demand analysis approach to the analysis results. Nejad et al. [8] established the effectiveness of this connection type in allowing the connected beam to develop full catenary action and inelastic capacity in the case of a sudden column loss.

The studies on SidePlate connections in the last couple years concentrated on the applicability of this connection type to the concrete filled steel tubular (CFST) column-steel beam connections and the improvement to be achieved thanks to this connection detail. Within this scope, Huang et al. [9] tested full-scale walled column-H-beam specimens. The column flange and web dimensions in the panel zone and the extension of the

side plates were adopted as test parameters. This study also presents theoretical bending capacity equations for three different failure modes of the connections with side plates, namely the plastic hinging of the beam away from the connection, failure of the weld between beam flange and side plate and the flexural failure of the side plate. Among the three tested specimens, the one with plastic hinging in the beam reached the highest ductility value due to the complete elastic behavior of the panel zone and plastic behavior at the desired location on the beam. In the search for effective connection details for CFST column-steel beam connections, Zhang et al. [10] designed and applied a side plate connection reinforced with an arc expanded cover plate (ACPSP connection) and a grooved side plate connection reinforced with a cover plate (GSPCP connection). This study mostly focused on the ability of the beam-column connections to withstand the catenary action in the case of the sudden loss of a column. These two new connection types improved the resistances of the beams to catenary action when compared to the traditional SidePlate connection. In a similar study, Liu et al. [11] examined the effectiveness of the SidePlate connection in wall-type CFST to I-beam connections. The pseudostatic cyclic loading tests and FEA indicated that the axial compression load on the column and the height of the side plates play an important role on the ductility of the connection region. Insufficient side plate height was found to trigger the failure of the gap between the beam and column, which in turn was responsible for the reduced strength of the connection. The tests of Liu et. [12] on walled CFST column-steel beam joint specimens indicated that the flexural moments are resisted by the internal force couple along the flange of the side plate initially. Later, part of the bending moments are transmitted to the column web, while the remaining portion of these bending moments are resisted by the lower flange of the beam on opposite side of the connection region.

The detailed literature survey, summarized above, clearly indicates that the effects of the SidePlate beam-column connections on the frame seismic behavior has not been studied previously to a major extent. The present study pertains to the contribution of the SidePlate beam-column connection to the resistances of the structural steel frames to ground excitations and the changes in the seismic response parameters (top-story lateral drift, base shear force and base overturning moment) with the use of this prequalified joint type. In this respect, the moment-curvature relations of a traditional pre-Northridge connection and a SidePlate one were determined numerically using a commercial software [13]. Later, linear time-history analyses of a regular structural frame with the two different connection types were conducted under three different ground motion records. The analyses were realized with the help of another commercial matrix structural analysis software [14]. The comparison of the response parameters from the analyses on the two connection types helped the authors to clearly demonstrate the great contribution of the SidePlate connection to frame behavior.

II. ANALYSIS PROCEDURE

A. Connection Details

Two connections are adopted in the present study. The first connection is the SidePlate connection, whose dimensional details are illustrated in Fig. 2. The extension of side plates in this connection are connected to the beam cover plates with the help of bolts to increase the resistance of the connection regions to dynamic loading. The parts of side plates, sandwiching the column, are welded to the stiffeners of the column. In this way, the ductility of the connection regions were increased since the fillet welds have the highest ductility if applied in parallel direction to the applied loading (lateral earthquake loading in this case). The column and beam have HEB500 and IPE450 cross-sections. The angles connecting the cover and side plates have L 110.25 cross-section. The grades of profiles, welds and M36 bolts of the connection are given in Table 1.

The second connection is a regular haunched beam-column connection for special moment frames (SMF) and all dimensional details of this connection are illustrated in Fig. 3. The profile and bolt sizes and grades of this connection are identical to the first connection. In the second connection, the beam is welded to an auxiliary plate, which is bolted to the column. The stress concentrations in the connection region were also relieved with the help of a haunch. The welds, connecting the flanges to the auxiliary plate, are expected to have low ductility, due to the orthogonality of these welds to lateral loading.

B. Stiffness Analysis of the Connections

The moment-curvature relations of the two connections were determined with the help of analyses on subassembly models, complying with the requirements of the section K2 of AISC 341-16 [15]. Accordingly, the connection models were composed of a single beam, attached to one side of the column, representing the edge and corner beam-column connections. As suggested by AISC 341-16 [15], the member lengths in the connection models were chosen based on the inflection points in the steel structural frame of the present study under earthquake loading. The inflection points on the beams are expected to develop at mid-length of the span in the case of lateral earthquake loading. Therefore, a beam length of 4000 mm was assigned to the models, considering the beam length in the analyzed frames was 8000 mm. Similarly, the upper and lower member lengths in the connection subassemblies had total lengths of 1000 mm and 1450 mm, respectively, based on the assumption that these inflection points develop at about 0.25 and 0.35 times the column length (4300 mm) from the center of the connection under earthquake loading.

The rotational stiffness analyses of both connections were conducted with the help of the IDEA StatiCa software [13] and the key parameters from these analyses are tabulated in Table 2. The structural analyses on the steel frame models adopted the values in this table for the connection region. The tabulated values clearly indicate that the SidePlate connection is superior to the haunched connection in terms of both strength and ductility for both reverse and forward directions of loading.

TABLE I
SECTIONS AND GRADES OF MEMBERS IN THE CONNECTIONS

Member	Section or Size	Grade
Beam	IPE 450	S275JR
Column	HEB 500	S275JR
Angles	L110.110.25	S355JR
Bolts	M36	10.9
Stiffener Plates	20 mm thick	S355JR
Side and Cover Plates	25 mm thick	S355JR
Fillet Welds	throat thickness in the range of 4.0-10.6 mm	E70xx ¹

¹a specified minimum tensile strength of 70 ksi (482 MPa)

TABLE II
RESULTS OF THE STIFFNESS ANALYSES

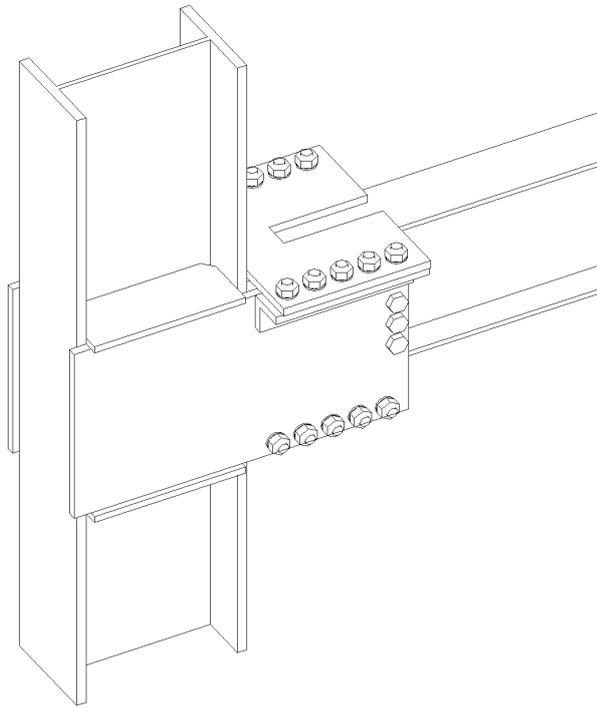
Loading Direction	Parameter	Haunched	SidePlate	% Increase
Reverse	Bending Resistance (kN.m)	541.2	585.1	8.1
	Initial rotational stiffness ¹ (MNm/rad)	470.6	636.5	35.2
	Rotational Capacity (mrad)	24.8	25.8	4.0
	Class	Rigid	Rigid	-
Forward	Bending Resistance (kN.m)	541.2	585.1	8.1
	Initial rotational stiffness ¹ (MNm/rad)	1072.2	574.1	46.4
	Rotational Capacity (mrad)	24.2	25.9	7.0
	Class	Rigid	Rigid	-

¹ limit stiffness for the rigid joint 177.1 MNm/rad and pinned joint 17.7 MNm/rad

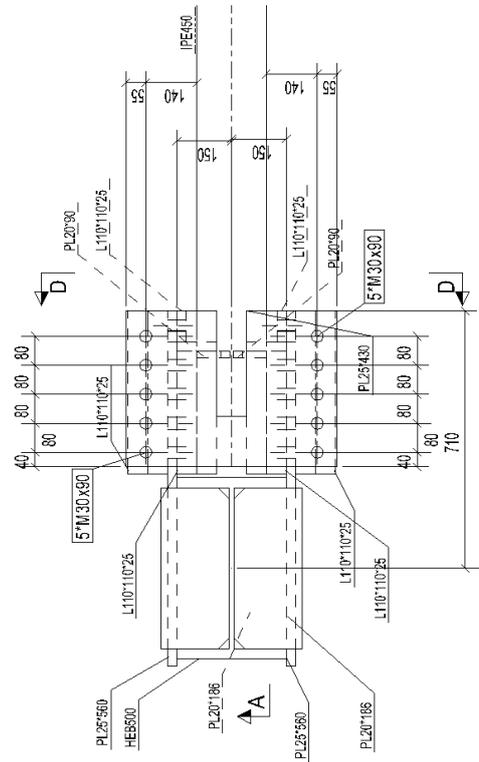
Only the rotational stiffness of the haunched connection in forward direction exceeds the respective value of the SidePlate one due to the contribution of haunch in this direction. One should remember that the ductility constitutes the most crucial parameter for the inelastic response of steel structural frames during an earthquake. The deformation capacities of the SidePlate connection were obtained to be 4% and 7% above the respective values of the haunched connection in the reverse and forward directions of loading, respectively.

C. Structural Model

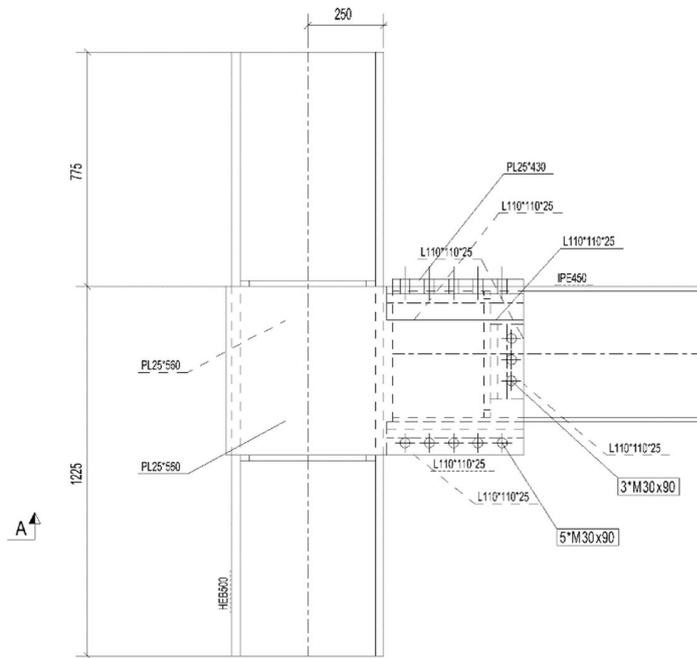
A four-bay four-story steel structural frame (Fig. 4) was modeled in the SAP2000 matrix structural analysis software [14]. This benchmark frame was regular in plan and had no structural irregularities whatsoever. In this way, the contribution of this new connection detail to the seismic responses of regular structures was aimed to be unfolded. All columns were fixed to the ground. All beams and columns had cross-sections of HEB500 and IPE450 of grade S275JR.



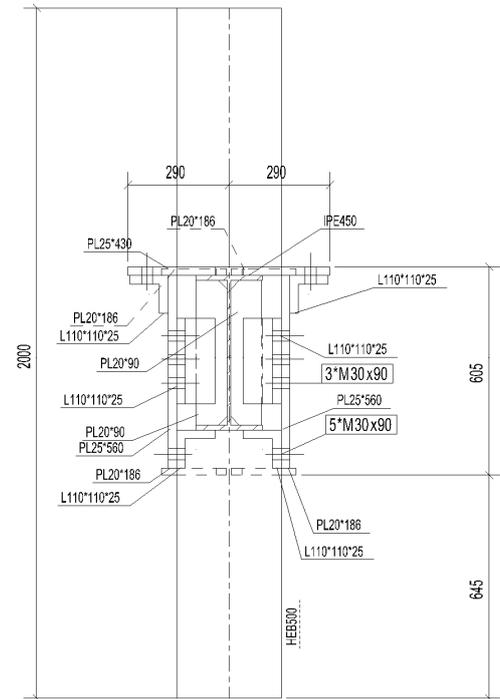
a-) Isometric View



b-) Top View

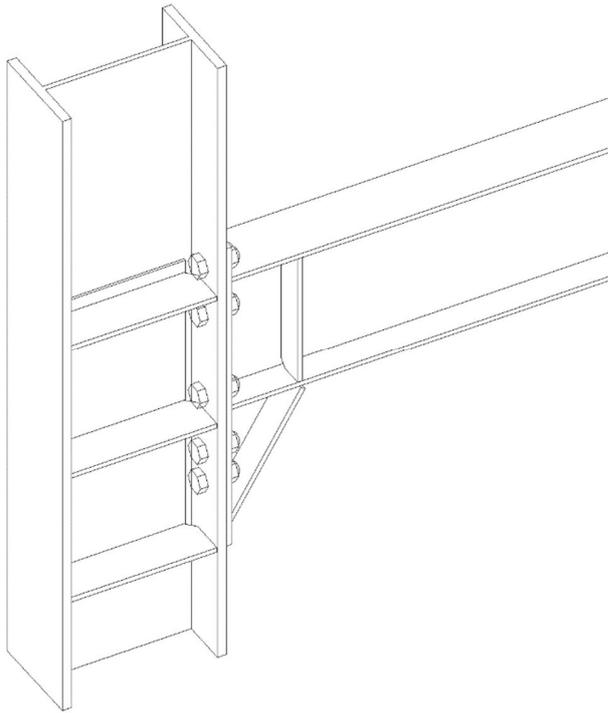


c-) Elevation View

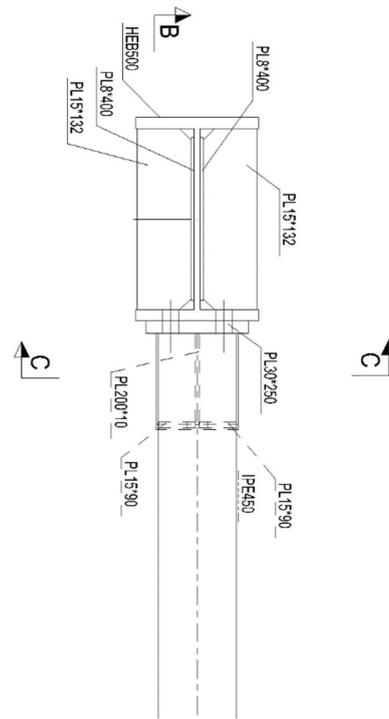


d-) Cross-Sectional View

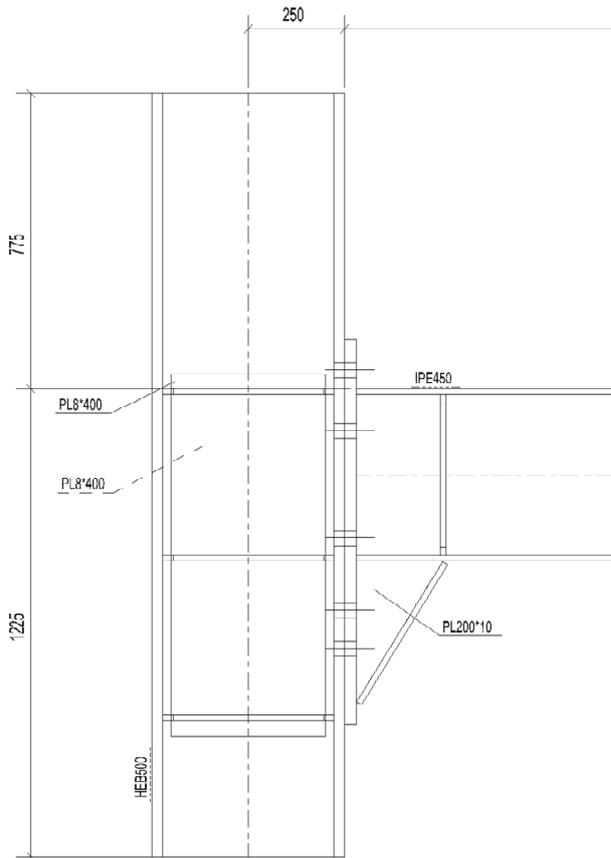
Fig. 2 SidePlate connection details



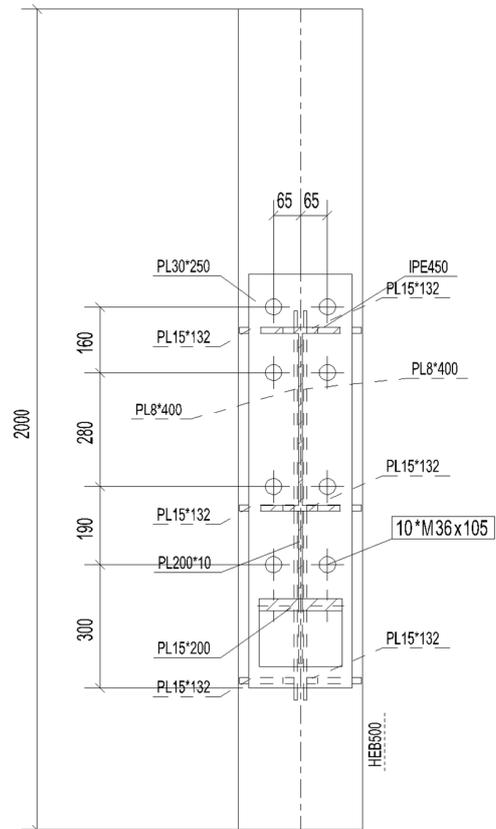
a-) Isometric View



b-) Top View



c-) Elevation View



d-) Cross-Sectional View

Fig. 3 Conventional haunched beam-column connection details

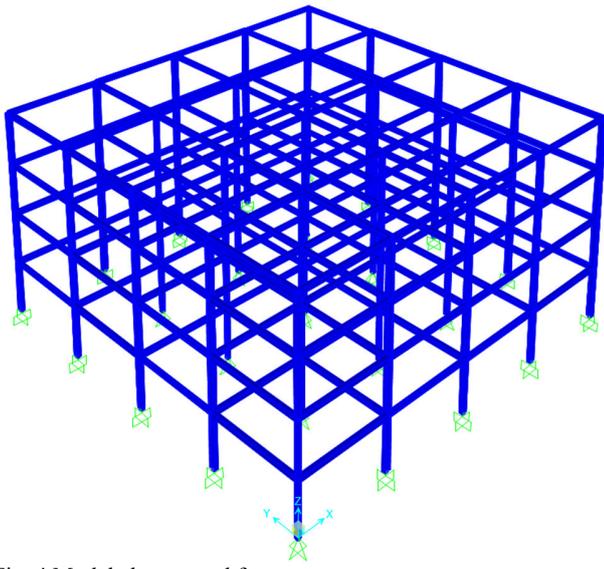


Fig. 4 Modeled structural frame

D. Earthquake Analyses

Linear time-history analysis method was used in the present study to examine the improvement in the overall earthquake performance of the benchmark structure with the use of SidePlate connection. Time-history analysis is the most realistic earthquake analysis method since it utilizes real ground motion records, scaled to the design response spectrum of the related structure. Furthermore, ground excitations, modified according to the local ground conditions of the location, can also be utilized in this method to obtain more realistic and reliable estimates [16].

The American [17], European [18], Turkish [19] and Australian-New Zealand [20] practice allows the use three ground motion records in the analyses as long as the maximum values of the structural response parameters (story drift, base shear, base overturning moment and story twist) are taken into account in the evaluation and comparison of analysis results. All these codes of practice [17]-[20] necessitate the simultaneous application of two horizontal and one vertical components of the earthquake record and choosing the records according to the seismic characteristics (peak ground acceleration, distance to fault and source mechanism of the seismic action) of the region. Based on the previous studies of the authors [21], [22], three records from the PEER website [23], namely the Bolu, Imperial Valley and Kocaeli records, were used in the present analyses. These records imitate the worst earthquake scenario of the City of Kirikkale, where the model structure was assumed to be constructed. The ground characteristics of the region were completely ignored in the present analyses by assuming full fixity at foundation level.

III. DISCUSSION OF RESULTS

Table 3 illustrates the structural response parameters for the two structural models in the case of three different ground motions.

TABLE III
RESULTS OF STRUCTURAL ANALYSES

Ground Motion	Dir.	Response Parameter	Haunc.	Side Plate	% Change
Imperial Valley	+ x	Base Shear (kN)	1729.6	1754.4	+1.4
		Overturning Moment (kN.m)	31986.6	31785.6	-0.7
		Lateral Drift (mm)	161.8	161.9	+0.1
	- x	Base Shear (kN)	1969.5	1978.2	+0.4
		Overturning Moment (kN.m)	50488.9	50856.0	+0.7
		Lateral Drift (mm)	196.3	194.9	-0.7
	+ y	Base Shear (kN)	1964.9	2017.4	+2.7
		Overturning Moment (kN.m)	50620.4	51022.2	+0.8
		Lateral Drift (mm)	195.2	199.0	+1.9
	- y	Base Shear (kN)	2469.4	2471.8	+0.1
		Overturning Moment (kN.m)	32944.3	32944.4	0
		Lateral Drift (mm)	204.8	203.5	-0.6
Bolu	+ x	Base Shear (kN)	1650.3	1661.5	+0.7
		Overturning Moment (kN.m)	49613.1	49605.2	-0.1
		Lateral Drift (mm)	203.2	201.9	-0.6
	- x	Base Shear (kN)	1943.5	1960.4	+0.9
		Overturning Moment (kN.m)	54848.4	55121.1	+0.5
		Lateral Drift (mm)	172.9	172.3	-0.3
	+ y	Base Shear (kN)	1950.5	1977.6	+1.4
		Overturning Moment (kN.m)	56906.8	56967.7	+0.1
		Lateral Drift (mm)	141.0	143.9	+2.1
	- y	Base Shear (kN)	1934.2	1966.4	+1.7
		Overturning Moment (kN.m)	50070.9	50278.1	+0.4
		Lateral Drift (mm)	164.2	166.7	+2.0
Kocaeli	+ x	Base Shear (kN)	1738.2	1736.4	-0.1
		Overturning Moment (kN.m)	36740.0	36834.2	+0.3
		Lateral Drift (mm)	190.0	190.1	-0.1
	- x	Base Shear (kN)	1529.4	1553.0	+1.5
		Overturning Moment (kN.m)	38419.0	38420.6	0
		Lateral Drift (mm)	190.4	190.2	-0.1
	+ y	Base Shear (kN)	1830.1	1906.9	+4.2
		Overturning Moment (kN.m)	40169.3	41603.4	+3.5
		Lateral Drift (mm)	228.3	234.2	+2.5
	- y	Base Shear (kN)	2312.7	2326.4	+0.6
		Overturning Moment (kN.m)	41195.9	40675.4	-1.3
		Lateral Drift (mm)	223.2	229.6	+2.7

The values in the table indicate that the use of SidePlate connection in replacement for the haunched connection yielded to a general trend of increase up to 4.2 % in the base shear and up to 3.5 % in the base overturning moment. This replacement resulted in the reduction of the top-story lateral drift in some analyses, while increasing the drift in the remaining analyses. In general, the effect of improving the connection region on the story drift is much less than its

effects on base shear and overturning moment. The increase in the lateral forces and bending moments for almost fixed values of lateral drift stems from the larger stiffness of the SidePlate connection as compared to the haunched connection. As the connections become more rigid, the general stiffness and strength of a structure increases. Similar lateral deflections result in greater internal forces and moments in the more rigid structural members, which in turn increases the forces and moments at the base of structure.

The almost unchanged values of the top-story lateral deflections, on the other hand, originate from the limited increase in the rotational stiffness of the connection with the use of side plates. The main superiority of the SidePlate connection is the tolerance to greater deflections in the inelastic range rather than its stiffness and strength. Hence, nonlinear time-history analyses on the same model structures will definitely demonstrate much greater differences between the seismic response parameters of these two models with different connection details. The linear analyses of the present study could only show the partial advantage of this connection.

IV. CONCLUSION

The present study is an attempt to unfold the advantages of the use of SidePlate connections in Special Moment Frames under seismic actions over conventional beam-column connections. For this purpose, two connections, including a conventional haunched one and a side-plated one, were designed. The moment-curvature relationships of the two connections were determined. Later, two structural models with no structural irregularities were developed for each connection type. Linear time-history analyses were conducted on both models by using three different ground motion records. These records were chosen based on the seismic features of the assumed location of the model structures. The two horizontal and one vertical components of each ground motion were applied to the structure simultaneously. The effect of the improved connection regions on the structural response was evaluated through the use of three response parameters, namely the top-story lateral drift, the base shear and the base overturning moment in both orthogonal plan dimensions.

The analyses indicated that the base shear and overturning moment values increased to a considerable extent, while the drift remained almost unchanged between the two models. The increases in the shear and moment values with the use of SidePlate connection were attributed to the increase in the structural stiffness thanks to the improved rotational stiffness of the connections. The linear analyses could not fully grasp the effects of the SidePlate connections on structural deflections. With the use of the full moment-curvature plots of the connection regions, the nonlinear analyses are expected to show that the structures with SidePlate connections undergo smaller lateral deflections thanks to additional reserve stiffness in the inelastic portion of the moment-curvature plot.

V. FUTURE STUDIES

The present study is the first stage of an analytical program. In this stage, only regular structural systems were analyzed and linear time-history analyses were conducted to identify the seismic performances of these regular frames. In the further stages of the study, frames with different structural irregularities (regular in plan, weak story, etc.) will be analyzed. What is more, nonlinear time-history analyses will be conducted to reach the damage states of the frames and their nonlinear deformations to uncover the actual performances.

ACKNOWLEDGMENT

The present paper is a condensation of the M.Sc. thesis of Mr. Gökhan Yüksel under the supervision of Prof. Dr. İlker Kalkan.

REFERENCES

- [1] A. Deylami and R. Ashraf, "Moment resisting connection with SidePlate (geometric aspects)," Paper No. 194, in *Proc. 13th World Conference on Earthquake Engineering*, Vancouver, Canada, August 1-6, 2004.
- [2] American Institute of Steel Construction (AISC), *ANSI/AISC 358-16: Prequalified Connections for Special and Intermediate Steel Moment Frames for Seismic Applications*. Chicago, U.S.A, 2016.
- [3] MiTek, "Sideplate plus connection design SMF test H2 close up," <https://www.sideplate.com/r-d/>
- [4] M. D. Engelhardt and T. A. Sabol, "Reinforcing of steel moment connections with cover plates: Benefits and limitations," *Eng. Struct.*, vol. 20, issues 4-6, pp. 510-520, 1998.
- [5] C. C. Chou, K. C. Tsai, Y. Y. Wang, and C. K. Jao, "Seismic rehabilitation performance of steel side plate moment connections," *Earthquake Eng. Struct. Dyn.*, vol. 39, pp. 23-44, 2010. DOI: <http://doi.org/10.1002/eqe.931>.
- [6] M. R. Shiravand and A. Deylami, "Application of full depth side plate to moment connection of I-beam to double-I column," *Adv. Struct. Eng.*, vol. 13, issue 6, pp. 1047-1062, 2010.
- [7] S. A. Jalali, M. Banazadeh, A. Abolmaali, and E. Tafakori, "Probabilistic seismic demand assessment of steel moment frames with side-plate connections," *Scientia Iranica, Trans. A: Civ. Eng.* vol. 19, issue 1, pp. 27-40, 2012.
- [8] I. Faridmehr, M. H. Osman, M. B. M. Tahir, A. F. Nejad, and R. Hodjati, "Seismic and progressive collapse assessment of SidePlate moment connection system," *Struct. Eng. Mech.*, vol. 54, issue 1, pp. 35-54, January 2015. DOI: <http://dx.doi.org/10.12989/sem.2015.54.1.035>
- [9] Y. Q. Huang, J. P. Hao, R. Bai, C. L. Fan, and Q. Xue, "Mechanical behaviors of side-plate joint between walled concrete-filled steel tubular column and H-shaped steel beam," *Adv. Steel Constr.*, vol. 16, issue 4, pp. 346-353, 2020. DOI: <https://doi.org/10.18057/IJASC.2020.16.4.7>.
- [10] W. Zhang, S. Jia, X. Xiong, Z. Chen, H. Liu, T. Su, and Q. Du, "Investigation of side plate connections in an S-CFST column frame under a column-loss scenario," *Struct.*, vol. 32, pp. 1302-1319, 2021. DOI: <https://doi.org/10.1016/j.istruc.2021.03.039>
- [11] H. Liu, J. Hao, Q. Xue, and X. Sun, "Seismic performance of a wall-type concrete-filled steel tubular column with a double side-plate I-beam connection," *Thin Walled Struct.*, vol. 159, pp. 107175, 2021. DOI: <https://doi.org/10.1016/j.tws.2020.107175>.
- [12] H. C. Liu, J. P. Hao, Q. Xue, and Y. Q. Huang, "Stress response and initial stiffness of side plate connections to WCFT columns," *Adv. Steel Constr.*, vol. 17, issue 3, pp. 306-317, 2021. DOI: <https://doi.org/10.18057/IJASC.2021.17.3.9>.
- [13] IDEA StatiCa S.R.O., *IDEA StatiCa- Structural Design Software for Steel and Concrete*, Brno, Czech Republic.
- [14] Computers & Structures, Inc., *SAP2000 Structural Analysis and Design*, Walnut Creek, CA, U.S.A.

- [15] American Institute of Steel Construction (AISC), ANSI/AISC 341-16: Seismic Provisions for Structural Steel Buildings. Chicago, U.S.A, 2016.
- [16] Y. B. Sonmezer, I. Kalkan, S. Bas, and S. O. Akbas, "Effects of the use of the surface spectrum of a specific region on seismic performances of R/C structures," *Nat. Hazard.*, vol. 93, pp. 1203–1229, 2018. <https://doi.org/10.1007/s11069-018-3347-3>
- [17] Federal Emergency Management Agency (FEMA), NEHRP recommended provisions for seismic provisions for new buildings and other structures (FEMA 450). Part 1: Provisions, Washington D.C., U.S.A., 2004.
- [18] European Committee for Standardization (CEN), EN 1998-1: Design of Structures for Earthquake Resistance- Part 1: General Rules, Seismic actions and Rules for Buildings. Brussels, Belgium, 2004.
- [19] Ministry of Public Works and Settlement, The Turkish Seismic Code (TSC) - Specification for Buildings to be Built in Seismic Zones, Ankara, Republic of Turkey, 2007.
- [20] Standards New Zealand, NZS1170.5:2004: Structural Design Actions, Parts 5 Earthquake Actions – New Zealand, Wellington, New Zealand, 2004.
- [21] B. Toprak, S. Bas, and I. Kalkan, "Effects of fly ash column treatment of HP clayey soils on seismic behavior of R/C structures," *Geomech. Eng.*, vol. 25, issue 6, pp. 473-480, 2021. <https://doi.org/10.12989/gae.2021.25.6.473>
- [22] S. Bas, J. H. Lee, M. Sevinc, and I. Kalkan, "Seismic performance of R/C structures under vertical ground motion," *Comput. Concr.*, vol. 20, issue 4, pp. 369-380, 2017. <https://doi.org/10.12989/gae.2021.25.6.473>
- [23] Pacific Earthquake Engineering Research Center, PEER Ground Motion Database. Retrieved from <https://ngawest2.berkeley.edu>.

Prof. Dr. Ilker Kalkan received his B.Sc. from the Department of Civil Engineering, Middle East Technical University, Ankara, Turkey in 2004; his M.Sc. and Ph.D. from Georgia Institute of Technology, Atlanta, Georgia, U.S.A in 2005 and 2009, respectively.

He is a professor in the Department of Civil Engineering, Faculty of Engineering and Architecture, Kırıkkale University, Kırıkkale, Turkey. He has authored or co-authored more than 40 SCI papers, published in several journals, including ACI Structural Journal, ASCE Journal of Structural Engineering, Engineering Structures, Composite Structures, Structural Engineering and Mechanics and Journal of Constructional Steel Research. His research interests include stability of steel and RC structural members, rehabilitation of reinforced concrete structures, the use of FRP reinforcing bars in concrete members and seismic behavior of steel and concrete structures.

Serdar Akca was born in Ankara, Turkey, 1987. He is M.Sc. student in Kırıkkale University, Institute of Science. He received his B.Sc. from the Department of Civil Engineering, Selçuk University, Konya, Turkey. His main research interests are steel structures and FEM modelling of structures.

Verifying Environmental Performance through Inventory and Assessment: Case Study of the Los Alamos National Laboratory Waste Compliance and Tracking System

Oral S. Saulters^{1*}, Shanon D. Goldberg*, Wendy A. Staples**, Ellena I. Martinez*, Lorie M. Sanchez*, Diego E. Archuleta*, Deborah L. Williams*, Scot D. Johnson*, Alex M. Baker*

Abstract--To address an important set of unverified field conditions, the Los Alamos National Laboratory Waste Compliance and Tracking System (WCATS) Wall-to-Wall Team performed an unprecedented and advanced inventory. This reconciliation involved confirmation analysis for approximately 5850 hazardous, low-level, mixed low-level, and transuranic waste containers located in more than 200 staging and storage areas across 33 Technical Areas. The interdisciplinary team scoped, planned, and developed the multidimensional assessments. Through coordination with cross-functional site hosts, they were able to verify and validate data while resolving discrepancies identified in WCATS. The results were extraordinary with an updated inventory, tailored outreach, more cohesive communications, and timely closed-loop feedbacks.

Keywords--Circular economy, environmental performance data, social-ecological-technological systems, waste management

I. INTRODUCTION

Along with the recent international attention surrounding the 26th United Nations Climate Change Conference of the Parties (COP26) in Glasgow, Scotland, the majority of Americans (64%) recognize the salience of environmental protection as an urgent and top priority [1]. From the global to the local, in this time of the Anthropocene, communities are galvanizing. For example, many local governments such as Los Alamos County, New Mexico have formed Resiliency, Energy, and Sustainability-related community-led initiatives to take action on these issues with new urgency [2].

This is also the case at Los Alamos National Laboratory (LANL or Laboratory); a unique and complex institution with a storied history (see Fig. 1 for LANL Technical Areas and surrounding lands). Building on that heritage as part of what US Secretary of Energy Granholm has described as America's "solutions department" [3], the current strategic LANL agenda is focused on: nuclear security; mission focused scientific, technical, and engineering; mission operations; and community relations [4]. Environmental, safety, and health (ESH) connections are essential to all of the noted agenda objectives. And accurate, reliable, and timely data are needed to effectively evaluate environmental performance goals [5]. As part of managing the dynamic ESH activities at LANL, the state-of-the art Waste Compliance and Tracking System (WCATS) was implemented in 2010 to successfully integrate waste management activities across functional areas (see Table 1 for the variety of LANL waste types, methods of disposal, and amounts). In particular, WCATS is a software application that incorporated the migration of electronic data from multiple pre-existing databases [6]. It has been specifically designed to manage LANL's wastes from cradle-to-grave. The interactive tool provides the proper support needed for the waste cycle: planning, forecasting, generation, characterization, processing, storage, treatment, shipment, and final disposal of wastes [7].

LANL has a tripartite governing policy for the environment which includes commitments to stewardship; continuous improvement; and risk-reduction/sustainability [8]. Management of wastes generated by Laboratory operations is a crucial component of compliance with environmental laws. Key federal and state rules such as DOE Order 435.1, Radioactive Waste Management; Resources Conservation and Recovery Act; 2016 Compliance Order on Consent; Federal Facility Compliance Act; and New Mexico Hazardous Waste Act provide regulatory drivers. Along with meeting federal, state,

*Los Alamos National Laboratory, Environmental Protection and Compliance, Los Alamos, NM USA

**Los Alamos National Laboratory, Quality Engineering and Inspection, Los Alamos, NM USA

¹Corresponding Author, osaulters@lanl.gov, 505.412.1810

and local regulations, Laboratory-specific ESH policies, directives, and procedures are designed to guide a diverse array of highly decentralized activities within a multilayered requirements structure and hierarchy [9]. Specifically, primary institutional requirements for environmental and waste management processes include but are not limited to: SD400-Environmental Management System, P409-LANL Waste Management, and the Hazardous Waste Facility Permit [10]-[12].

Due to its clandestine origin story as a “Secret City” of the Manhattan Project requiring the highest levels of security including its isolated location, LANL research facilities have developed somewhat independently. The corollary is a web of waste management processes which historically have been relatively disparate and semi-autonomous [13]. As early as 1948, the Atomic Energy Commission recognized the importance of institutionalized waste management [14]. Fast forward seven plus decades and in addition to the core LANL Environment and Waste Programs (EWP) functions, a distributed cadre of deployed ESH professionals, waste management coordinators (WMCs), and packaging specialists provide facility-level rooted support for generators at each of the Technical Areas on a day-to-day basis. However, with the velocity and flux of ever evolving R&D projects, construction, operations, demolition, and legacy cleanup initiatives, a distinct lag-time puzzle of uncertainty emerged. Verifying feedbacks across the laboratory campus requires broad collaboration and on the ground mobilization vis-à-vis WCATS auditing. Therefore, a central question crystallized, *how can the robust assortment of waste streams within WCATS and varied waste practices across the vast network of Laboratory facilities be effectively inventoried, measured, and validated?*

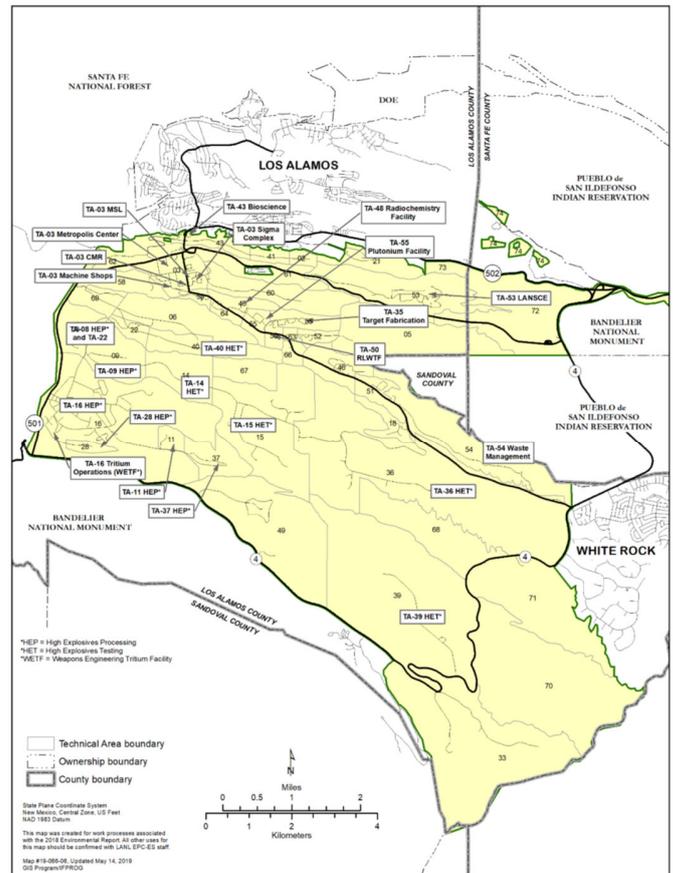


Fig. 1 Los Alamos National Laboratory Technical Areas and surrounding lands [8]

TABLE 1 LANL Waste Types and Disposal Methods [8]

Waste Type	Method for Disposal	Disposal Amount (2019)
Solid Transuranic Waste and Solid Mixed Transuranic Waste	The Laboratory sends solid transuranic and mixed transuranic wastes offsite to the Waste Isolation Pilot Plant in Carlsbad, NM, when the transuranic or mixed transuranic waste meets the plant’s waste acceptance criteria. Some transuranic and mixed transuranic waste is stored at the Laboratory while waiting for an acceptable disposal pathway to be identified. In 2019, LANL waste was also shipped from	8,270 cubic yards (6,322.5 cubic meters) from LANL; 4 cubic yards (3.1 cubic meters) from Waste Control Specialists in Andrews County, Texas

	long-term storage at Waste Control Specialists (Andrews County, TX) to WIPP.	
Solid Low-level Radioactive Waste	The Laboratory sends solid low-level radioactive waste offsite to licensed treatment, storage, and disposal facilities. These sites include the Nevada Nuclear Security Site, operated by the DOE, and commercial facilities operated by Energy Solutions (Clive, UT); Perma-Fix Northwest, Inc. (Richland, WA), and Waste Control Specialists (Andrews County, TX).	48,957 cubic yards (37,431 cubic meters)
Liquid Radioactive Waste	The Laboratory treats liquid radioactive waste onsite at the Radioactive Liquid Waste Treatment Facility in TA-50. The treated water is either evaporated or released at permitted Outfall 051.	694,982 gallons
Solid Hazardous Waste	The Laboratory sends solid hazardous waste offsite for treatment and disposal at licensed treatment, storage, and disposal facilities. In 2019, these facilities included Veolia North America (Henderson, CO) and Clean Harbors (Clive, UT).	150 tons
Solid Mixed Low-Level Waste	The Laboratory sends solid mixed low-level waste offsite to licensed TSDFs. In 2019, these facilities included Energy Solutions (Clive, UT), Perma-Fix of FL, Inc. (Gainesville, FL), and Waste Control Specialists (Andrews County, TX). Some mixed	6,108 cubic yards (4,670 cubic meters)

	low-level waste is treated at one of the licensed treatment, storage, and disposal facilities to meet land-disposal restrictions and is then disposed of at the Nevada Nuclear Security Site.	
Solid Nonhazardous Waste	The Laboratory sends sanitary solid waste, construction debris, and demolition debris to the Los Alamos County Eco Station for transfer to municipal landfills such as the landfill in Rio Rancho, NM. LA County operates this transfer station and is responsible to the State of NM for obtaining all related permits for these activities. The Laboratory also sends solid nonhazardous waste to regional facilities in AZ and CO.	2,727 tons
Liquid Sanitary Waste	The Laboratory treats liquid sanitary waste onsite at the Sanitary WWTP. Treated water is reused in Laboratory cooling towers and is released at permitted Outfall 001.	1,249,214 gallons
PCB Wastes	Waste containing PCBs, including fluorescent light ballasts and contaminated soils, was sent to U.S. EPA-authorized TSDFs, including Clean Harbors (Clive, UT) and Veolia North America (Henderson, CO).	358 tons
Asbestos Waste	Asbestos-containing waste is deposited at any of several waste disposal sites operated in accordance w/ 40 CFR 61, 154	88 tons

II. DESCRIPTION

By internalizing ESH values, principles, and tools into organizational systems and culture, better outcomes can be achieved [15]. Per LANL System Description, SD400, Environmental Management System, the first important step for formal LANL approval of potential waste management begins with the Integrated Review Tool (IRT), a decision hub that interfaces with the Permits and Requirements Identification (PRID) Tool [10]. Subject Matter Experts (SMEs) from key disciplines and programs through use of the PRID Tool conduct extensive reviews of all proposed actions submitted by proponents to determine which multimedia approvals and permits may be required. For example, if a hazardous waste permit is required, the generator, WMC, SMEs, leaders, and intergovernmental partners can work together through formal coordination (e.g., Integrated Project Team [IPT]) and use of policies and procedures, such as P409 LANL waste management, which incorporate decision hierarchies [16] to ensure that the IPT have thoroughly scoped all lifecycle stages; and have planned for the safe, efficient, sustainable, and just path forward. Concurrently and symbiotically, efforts are initiated through WCATS to setup the necessary architecture configurations for waste requirements (see Fig. 2 for WCATS main screen format).

This opens up options for creatively analyzing waste streams as potential feedstock for materials, energy, and information flows; through industrial ecology and circular economy inspired frameworks [17]-[19]. Accordingly, it is recognized that pollution prevention and waste minimization are preferred; and ultimately **Waste Remediation (WR)** can be an important ecosystem service for protection of human health and the environment [20].

Therefore, a conceptual framework for understanding and advancing lessons from this interdependent **Social-Ecological-Technological system (SET)** case study can be used [21], [22]. Consequently, the (S) are the human dimensions involving the LANL workforce and broader communities; the (E) are the natural settings including northern New Mexico and associated ecosystems involved; and the (T) are the infrastructure and products as represented by WCATS.

To advance the effort of establishing the capacity of WCATS to sustain WR in the face of acute and chronic stresses with ongoing evolution and improvement, a tiered process took place. First, the properties of WCATS were thoroughly analyzed. Second, implications for the key attributes of the broader LANL institutional governance arrangements continue to be put forward.

The first phase of the LANL Wall-to-Wall Inventory (W2W) Team was comprised of multidisciplinary members with distinct academic and professional backgrounds including: various ESH programs, areas of regulatory specialty, waste expertise, software engineering, database experience, technology training and transfer, and metric development skills; all of which contributed to the success of the project.

They met daily to communicate the plan of the day, assign and provide status on actions, determine facility schedule, ensure access was granted, and discuss outstanding issues or concerns before reporting to sites. Some had the ability to query WCATS to generate ad-hoc reports to track process and define the scope of the project, others provided WCATS training for users and other members of the team, some had backgrounds in quality assurance/quality control (QA/QC), data presentation, graph and metric data, and document preparation. The team collaborated with each other, partner organizations and facility subject matter experts that managed the staging/storage areas to be able to accomplish the inventory with efficiency and accuracy while minimizing impacts to routine operations.

The W2W Team leveraged existing protocols and data while preparing and developing a pragmatic proposal for performing this unprecedented strategic initiative by conducting research, analysis, and project scoping [23]. Empowered with authorization by LANL leadership as illustrated by laboratory-wide notifications, their creative and innovative approach in establishing the methods of the project, querying the WCATS baseline inventory report for designated areas to identify the number of sealed/closed containers and items in the database, and building the platform to evaluate approximately 5850 waste containers and items on site proved to be effective. Following pilot scale tests, they geospatially verified the location and status of over 3000 containers and items (~52%) in onsite staging/storage areas in nearly 200 locations across 33 Technical Areas; while also closing the data loop on those dispositioned off site. For instance, given the numerous types of waste streams generated at LANL, many commercial off-site Treatment, Storage, Disposal Facilities (TSDFs) are partners for ultimate disposal in the WR interstate network (e.g., UT, NV, TN, WA, CO, NM, and TX). The evaluation process included review of available container data such as histories, shipping manifests, supporting reports, field logs, and employee interviews to identify and ground truth any relics. Clarification of these records provides a more accurate representation of active waste storage by LANL [6].

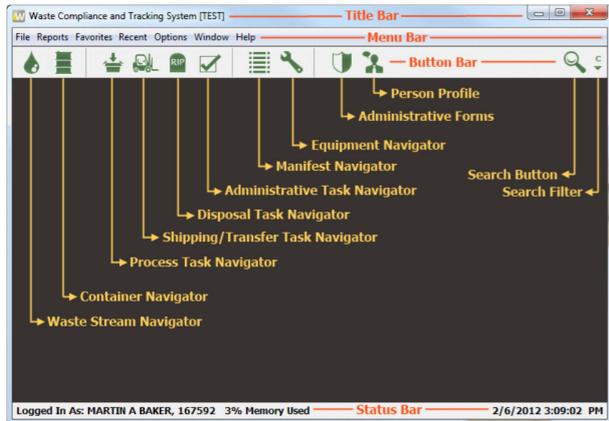


Fig. 2 Waste Compliance Assessment and Tracking System (WCATS) main screen [7]

III. DISCUSSION

The field component of the initiative took place from June 2019 through January 2020. As part of the SET, the cross-functional W2W Team developed new procedures, checklists, and processes used to: execute the work in the field, decommission items in WCATS, manage, track and develop reporting metrics and graphs. The team used held-hand mobile devices to scan and record 3600 containers on-site and evaluated how wide spread the use of these mobile devices was by embedded facility personnel [24]; see Fig. 3 for examples of mobile device and waste containers in storage. The W2W team developed a two-person (buddy strategy) approach to independently scan containers and then compare results to ensure that no containers were missed or double counted. The procedures and checklists were revised based on experience to improve the efficiency and accuracy of the inventory as they conducted the verification.

The multidimensional assessment process was flexible enough to accommodate special situations. An example was inventorying waste containers at TA-55, where arduous security, potential radiation exposure, and the significant number of waste containers made the traditional approach impractical. Instead, the WCATS inventory was compared to Nuclear Material Control and Accountability (NMCA) information contained in the Local Area Nuclear Material Accountability Software (LANMAS) database for transuranic wastes. Due to rigorous DOE and LANL standards, the quality of the LANMAS data is very high and was deemed adequate to authenticate the data in WCATS. This allowed the inventory of those items to be cross-walked and completed with minimal safety and security risk and no violations.

Overall, the W2W Team synchronized with all levels of management, waste generators, deployed ESH professionals, waste management coordinators, packaging/transportation services leaders, and facility representatives to organize access approvals, obtain escorts as required, ensure site access training, walk-down the facility sites, report issues and concerns, develop and convey WCATS training and changes, and to provide monitoring. Diligent, rigorous, and iterative steps were taken to account for containers throughout the waste management continuum and their operational and administrative status. From being under the control of generators to being processed, packaged, and transported on-site or shipped off-site during the noted interval a more accurate empirical distribution was ascertained. In summary, more than half of the baseline containers and items were physically confirmed. Containers present in the WCATS database but not located on-site were potential database migration relics (e.g., containers or items that were a product of the transfer of historic database records into the WCATS database).

Due to the information migration, container histories were sometimes incomplete, re-packaging, and shipping histories on paper-based systems were not effectively captured electronically. Of this artifact portion, reconciliation efforts directly reduced the active waste container inventory by decommissioning and other clarifying mechanisms. The synergistic initiative resulted in improved data quality, streamlined processes, better management operations, and capacity building. Lessons learned from this visionary strategy can be evaluated for customization and potential replication at other DOE laboratories, production facilities, and beyond.

The SET is consistent with the cradle-to-cradle guidance by McDonough and Braungart [25] for design innovators and organizational leaders, as the W2W initiative continues to: signal intention (LANL distinguished performance award for the effort); restore (building capacity); innovate further (re-engineering and upgrade of WCATS and next generation mobile devices); understand and prepare for the learning curve (new procedures and outreach); and exert intergenerational responsibility (succession planning through mentoring, cross-training, and partnerships).



Fig. 3 WCATS Mobile Device and Waste Containers [7]

IV. CONCLUSIONS

Several macro-level institutional compliance feedback mechanisms were strengthened and reinforced such as: the Issues Management actions to provide revisions and updates to the Laboratory's Waste Management Policy, P409; crosswalks with the Hazardous Waste Facility Permit; linkages with the Environmental Management System; and many internal procedures in multiple organizational units to ensure principles for building resilience and sustaining ecosystem services. The second phase of the W2W Team continues to exhibit the unusual commitment and dedication by ensuring that WCATS remains accurate and current as new containers and items are added to the software. This effort is being accomplished by hiring new data stewards to track and monitor data, ongoing mentoring and training for users of the network, developing over 45 training videos during the pandemic, and ensuring system updates are completed in a timely manner. The original estimate was that the Wall-to-Wall inventory would take approximately one year to complete. The team continued to demonstrate going above and beyond normal expectations by refining the process and coordinating efforts with host facility site personnel, resulting in the inaugural field work portion of the inventory being completed in approximately five months. The W2W Team demonstrated schedule flexibility to prevent impact to facility operations and avoided curtailing or pausing site activities.

With respect to next steps for the project, it is recommended that the resilience of WR, a key ecosystem service as managed by WCATS, continue to be evaluated. This represents an impactful area of ESH knowledge and practice for supporting current and future LANL needs. By using the principles of SET: 1) diversity and redundancy; 2) connectivity; 3) slow variables and feedbacks; 4) complex systems thinking; 5) learning; 6) participation; and 7) polycentricity; advances can be developed toward best

practices, gap analysis, and the path forward to sustainable, regenerative, and long-term stewardship.

ACKNOWLEDGMENT

The Authors thank the many colleagues, partners, and collaborators that assisted with and supported these initiatives in various ways. This work is funded by Los Alamos National Laboratory, Triad National Security. The Authors declare that there is no conflict of interest.

REFERENCES

- [1] Pew Research Center, "As Economic Concerns Recede, Environmental Protection Rises on the Public's Policy Agenda" (2020).
- [2] LARES, "Los Alamos Resiliency, Energy and Sustainability Task Force: Interim Report to County Council", August 2021 (2021).
- [3] Discover LANL, "From Climate and Cancer Solutions to Commitment to the Mission: Secretary of Energy Granholm 'Visits' Los Alamos for Packed Virtual Tour June 17, 2021" (2021). <https://discover.lanl.gov/news/stories/granholm-visit>
- [4] Los Alamos National Laboratory, "Laboratory Agenda", LA-UR-21-20301 (2021).
- [5] M. Song et al., "Environmental Performance Evaluation with Big Data: Theories and Methods", *Ann Oper Res*, 270: 459-472 (2018).
- [6] S. Goldberg and E. Martinez, "2019 Wall-to-Wall Waste Compliance and Tracking System Inventory Report, EPC-DO-20-179" LA-CP-21-20529 (2020).
- [7] M. Baker, "Waste Compliance and Tracking System (WCATS) Users Guide" LA-UR-20-23289 (2020).
- [8] L. Hansen, et al., "Los Alamos National Laboratory 2019 Annual Site Environmental Report", LA-UR-20-26673 (2020).
- [9] Los Alamos National Laboratory, "Requirements Systems and Hierarchy", *PD311 Revision 4* (2021).
- [10] Los Alamos National Laboratory, "Environmental Management System", *SD400 Revision 5* (2021).
- [11] Los Alamos National Laboratory, "LANL Waste Management", *P409 Revision 8* (2021).
- [12] New Mexico Environment Department, "Los Alamos National Laboratory Hazardous Waste Facility Permit", EPA ID NM0890010515, (2010), <https://www.env.nm.gov/hazardous-waste/lanl-permit/>.
- [13] S. Jones, et al., "Preparing Los Alamos National Laboratory's Waste Management Program for the Future", *WM2012 Conference Proceedings*, February 26 - March 1, 2012, Phoenix, Arizona USA (2012).
- [14] L.A. Emelity, *Waste Management at Los Alamos: Protecting Our Environment*, pp. 3, Los Alamos National Laboratory, LALP-90-30 (1991).
- [15] S. Stricoff, "Safety Leadership: Avoid the False Dichotomy of Systems vs. Culture", *Safety and Health Magazine*, 190, 4 (2014).
- [16] U.S. EPA, "Municipal Solid Waste in the United States: 2012 Facts and Figures", EPA-530-R-13-001. Office of Solid Waste, Washington, DC (2014).
- [17] S. Erkman, "Industrial Ecology: An Historical View", *J. Cleaner Prod*, 5, 1-2: 1-10 (1997).
- [18] R.L. Smith et al., "An Industrial Ecology Approach to Municipal Solid Waste Management", *Resources, Conservation and Recycling*, 104: 317-326 (2015).
- [19] A. Pires and G. Martinho, "Waste Hierarchy for Circular Economy in Waste Management", *Waste Management*, 95: 298-305 (2019).

- [20] S.C.L. Watson et al., "A Conceptual Framework for Assessing the Ecosystem Service of Waste Remediation: In the Marine Environment", *Ecosystem Services*, 20: 69-81 (2016).
- [21] B. Cosens et al., "Governing Complexity: Integrating Science, Governance, and Law to Manage Accelerating Change in the Globalized Commons" *PNAS*, 118, 36: 1-9 (2021).
- [22] R. Biggs et al., "Toward Principles for Enhancing the Resilience of Ecosystem Services" *Annual Review of Environment and Resources*, 37: 421-48 (2012).
- [23] Los Alamos National Laboratory (LANL), "Waste Compliance and Tracking System (WCATS): Managing Containers Wall-to-Wall" (2017).
- [24] E. Rogers, *Diffusion of Innovations*, Fourth Edition. The Free Press: New York (1995).
- [25] W. McDonough, and M. Braungart, *Cradle-to-Cradle: Remaking the Way We Make Things*, pp. 182-186, North Point Press: New York (2002).

Programmed Speech to Text Summarization Using Graph Based Algorithm

Ms. Hamsini Pulugurtha
Pursing B.Tech 2nd Year
Department of Computer Science and Engineering
Gayatri Vidya Parishad College Of Engineering (Autonomous)
Visakhapatnam ,India.

Dr. P.V.S.L Jagadamba
Professor,
Department of Computer Science and Engineering,
Gayatri Vidya Parishad College of Engineering for Women, Visakhapatnam, India.
drpvsljagadamba@gvpcew.ac.in

Abstract— Programmed Speech to Text and Text Summarization Using Graph-based Algorithms can be utilized in gatherings to get the short depiction of the gathering for future reference. This gives a signature check utilizing Siamese neural organization to confirm the personality of the client and convert the client gave sound record which is in English into English text utilizing the discourse acknowledgment bundle given in Python. At times just the outline of the gathering is required, the answer for this text rundown. Thus, the record is then summed up utilizing the regular language preparing approaches, for example, solo extractive text outline calculations.

Keywords—Siamese Neural Network, English Speech to English Text, Natural Language Processing, Unsupervised extractive text Summarization.

I. INTRODUCTION

Signature Verification authenticates the identity of individuals by measuring their signatures. The signature is treated as a series of movements that contain unique biometric data. The growing need for a full proof signature verification scheme ensure that the proposed scheme can provide comparable and if possible better performance than already established signature verification schemes.

Speech is the most important part of communication between human beings. Speech recognition is the process of making a machine recognize the speech of different people based on certain words or phrases it can be converted into text.

Text Summarization solves this problem by providing a shortened summary of it with consumes lot of time. In the proposed work a combination of speech to text conversion and text summarization is implemented.

II. LITERATURE SURVEY

In a paper performing Signature verification using Siamese Neural Network [1], the study presented an algorithm based on an artificial neural network, called “Siamese” time delay neural network, which consists of two identical networks

joined at their output. During training the network learns to measure the similarity between pairs of signatures. When used for verification, only one half of the Siamese network is evaluated. The output of this half network is the feature vector for the input signature. Verification consists of comparing this feature vector with a stored feature vector for the signer. The signature closer than a chosen threshold to this stored representation are accepted, all other signatures are rejected as forgeries.

The paper, Evaluating Google Speech-to-Text API's Performance for Romanian e-Learning Resources [2], presents the history of ASR systems together with main approaches used by the algorithms behind these systems.

In a paper, Automatic Text Summarization using text rank algorithm [3], had taken the data from various sources for a particular topic and summarized it for the convenience of the people, so that they do not have to go through multiple sites for relevant data.

The paper, a Survey on Extractive Text Summarization, explains about the different types of summarization methods and describes briefly about the extractive text summarization techniques.

III. PROPOSED SYSTEM

The orientation provided by this work can be very competent in finding the authentication through digital signature verification, audio files for speech to text conversion and then text summarization. This leads to improve authentication and summary making. Already, the users signature dataset will be stored, and user audio will also be recorded. The work is divided into three modules Signature verification, Speech to text conversion, Text summarization.

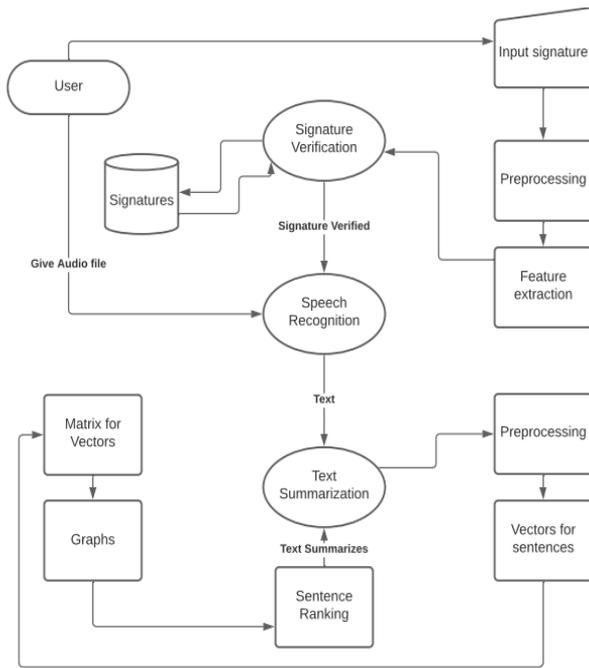


Fig. 1 Content Diagram

IV. SIAMESE NEURAL NETWORK

A. Data

The initial data collection was made by all our team members and each genuine signature is scanned and placed in an input1 folder with 250 total signatures. Later when producing genuine or forgery signatures, the signer was shown an example of the genuine signature and is placed in a folder named: input2. Finally, the information about the signature whether it is a genuine or forgery is stored in a .csv file with 0 or 1 representation for forgery or genuine kind signatures.

B. Pre-Processing

The provided 500 signatures are initially converted into black and white image. To make sure that each and every image is of the same size, resizing the signatures is required to 200 x 500 dimensions. Finally, after converting the images into same size, normalization is performed to make the values range between specific values. Now the data is ready to split into train, test, validation set. We have considered 180 images to be in the training set and 30 signatures for validation and 40 signatures for test set.

C. Architecture

The module needs to learn to discriminate forgeries from genuine signatures, led us to consider the use of artificial neural networks techniques. Neural networks can learn from examples and they can also be used to compress data. This

work is based on a neural network architecture, called a Siamese network, which first extracts features from two signatures and then evaluates the distance between these two sets of features.

The network has two input fields to compare the two patterns and one output whose state value corresponds to a measure of the distance between the two patterns. Two separate subnetworks based on Time Delay Neural Network (TDNN) act, one on each input pattern, to extract features, then the cosine of the angle between the two feature vectors is calculated and this represents the distance value. The cosine distance measure was rather than the Euclidean distance. Using Euclidean distance and requiring two genuine signatures to have a small distance could lead to the trivial solution of zero-sized feature vectors.

The subnetworks are constrained to be identical; they are multi-layered and feed-forward with several layers of feature extraction before distance measure. TDNN's use local processing units having inputs that are delayed in time. In other words, a unit receives inputs from just a portion of the signature trajectory. In a multi-layer TDNN, successive layers of units can extract features spread over a wide range of time.

To train the network, two patterns are applied, one to each input, and a desired value for the output is used to backpropagate the error. The desired output is for a small angle between the outputs of the two subnetworks (f_1 and f_2) when two signatures are presented, and a large angle if one of the signatures is forgery. For the cosine distance used here:

$$(f_1 \cdot f_2) / (|f_1| |f_2|)$$

The desired output is in between 0 and 1 and the threshold value decides whether the respective signature is a genuine one or not.

D. Training the Neural Networks

During training, part of the data set is used for learning the weights of the units, another part, called the validation set, is used after a training iteration to test the test the network's performance. A measure of the performance of the network at accepting genuine pairs and rejecting forgeries is found by calculating the percentage of genuine signature pairs for which the output was greater than threshold value.

V. SPEECH TO TEXT CONVERSION

Speech recognition is the ability of a computer software to identify words and phrases in spoken language and convert them to human readable text. We do not need to build any machine learning model from scratch, this library provides with convenient wrappers for various well known public speech recognition APIs. We have used Google Speech Recognition in this application, as it is straightforward and does not require any API key.

Make sure the audio file in the .wav format and contains English speech. Then, load the audio file, and create a speech recognition object and then split the audio file into

chunks and apply speech recognition on each of these chunks. Finally, returns the text for all chunks detected. We can also change the parameters suitable for your audio file.

VI. TEXT SUMMARIZATION

A. Reading the Data

We read the data from output generated by the Google Speech Recognition API and now we split the text into individual sentences.

B. Pre Processing

One of the major forms of preprocessing is to filter out useless data. In natural language processing, useless words, are referred to as stop words. We apply preprocessing on the text which enables us to remove all the extra data from it which are of no use, example: "the", "a", "an", "in" etc.

C. Word Embeddings, Similarity Matrix and Graphs

Word embeddings are vector representation of words. These word embeddings will be used to create vectors for our sentences. The Words from vocabulary is mapped to vectors or numbers. Word embeddings are in fact a class of techniques where individual words are represented as real-valued vectors in a predefined vector space. This is achieved by mapping words into a meaningful space where the distance between words is related to semantic similarity. Based on vector of sentences we prepare a similarity matrix, i.e., find similarities between the sentences, and we will use the cosine similarity approach. Let us first define a zero matrix and populate it with cosine similarities of the sentences. Cosine similarity to compute the similarity between a pair of sentences and initialize the matrix with cosine similarity scores. Now convert the matrix into graph. The nodes of this graph will represent the sentences and the edges will represent the similarity scores between the sentences.

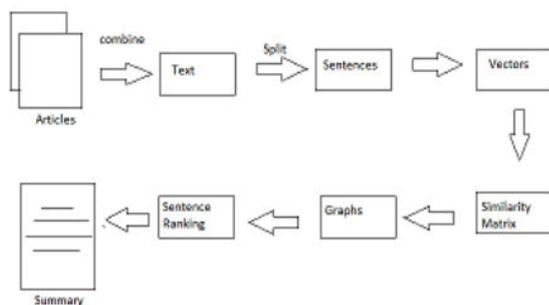


Fig. 2 Text Summarization Process

D. Page Rank Algorithm

Page rank is used primarily for web pages, we would have to compute a score called page rank score. The undirected graph is then passed its time to extract the top N sentences based on their rankings for summary generation. For our, based on similarity we take out top 2 or 3 sentences and summarize them. The basic idea of Text Rank is to provide

a score for each sentence in a text, then you can take the top-n sentences and sort them as they appear in the text to build an automatic summary.

VII. CONCLUSION

The Web application is designed with sign in and signs up, Initially the user authentication with the signature verification using Siamese Neural Network model is the one module of the paper. Once the user is verified means, the audio file is chosen from the files and then, on clicking submit button, audio file goes to speech recognition process, where python code is used and then the audio file is converted to text file. The text file goes through page rank algorithm and summarized text is produced. Finally, the converted text and summarized text are produced as the output of the paper.

The real idea of the paper is to prepare a summary of meetings, using meeting recorded audio files. This will save time and also authentication using signature verification helps in not accessing or entering of any other persons into a meeting. Text summarization refers to the technique of shortening long pieces of text.

The proposed method is useful to preparing notices and a summary and also the safety of user's data using authentication. In the future, the accuracy of the models used in this paper can be increased, For speech recognition, new functions and new models can be developed. For authentication, biometric models can be used, and a recording option can be added which will automatically recorded and covert to text. This paper can be used in many areas such as meetings, summarizing large data, making notes, conferences, hospitals. This paper's intention is to create a coherent and fluent summary having only the main points outlined in the document.

REFERENCES

- [1] Jane Bromley, James W. Bentz, Leon Bottou, Isabelle Guyon, Yann Lecun, Cliff Moore, Eduard Sackinger and Roopak Shah "Signature Verification Using A "Siamese" Time Delay Neural Network" By Carnegie Mellon University on 10/14/14, AT&T Bell Laboratories, Crawfords Corner Road Holmdel, NJ 07733, USA and NCR Corporation, E&M Atlanta 2651 Satellite Boulevard, Duluth, GA 30135, USA.
- [2] Iancu, Bogdan. (2019). Evaluating Google Speech-to-Text API's Performance for Romanian e-Learning Resources. *Informatica Economica*. 23. 17-25. 10.12948/issn14531305/23.1.2019.02.
- [3] Tanwi, Satanik Ghosh, Viplav Kumar, Yashika S Jain, Mr. Avinash B "Automatic Text Summarization using Text Rank" *International Research Journal of Engineering and Technology (IRJET)* Volume: 06 Issue: 04 | Apr 2019
- [4] N.Moratanch ,S.Chitrakala "A Survey on Extractive Text Summarization" *IEEE International Conference on Computer, Communication, and Signal Processing (ICCCSP-2017)*.

Review on Petg Material Parts Made Using Fused Deposition Modeling

Dhaval Chauhan,
Post Graduate Student,

Mahesh Chudasama
Assistant Professor

Mechanical Engineering Department,
Government Engineering College, Dahod, India

Abstract: This study has been undertaken to give a review of Polyethylene Terephthalate Glycol (PETG) material used in Fused Deposition Modelling (FDM). This paper offers a review of the existing literature on polyethylene terephthalate glycol (PETG) material, the objective of the paper is to providing guidance on different process parameters that can be used to improve strength of the part by performing various testing like tensile, compressive, flexural etc. This work is target to find new paths that can be used for further development of use of fiber reinforcement in PETG material.

Keywords: PETG, FDM, Tensile strength, Flexural strength, Fiber reinforcement

Knowledge of Artificial Insemination and Agribusiness Management for Social Innovation in Rural Populations

Yasser Y Lenis^{1,3}, Daniela Garcia Gonzalez¹, Cristian Solarte Bacca¹, Diego F Carrillo González², Amy Jo Montgomery⁴, Dursun Barrios⁵.

Abstract— Introduction: Artificial insemination in bovines helps to promote genetic improvement and can positively impact the rural economy. The Colombian armed conflict has forced a large portion of the rural population to abandon their territory, affecting their education, family integration and economics. **Justification:** The achievement of education in rural populations was one of the Millennium Development Goals (MDGs), made by the United Nations. During the last World Summit on Sustainable Development (WSSD), it was concluded that most of the world's poor, illiterate and undernourished population lives in rural areas; therefore, access to education is considered one of the greatest challenges for governments in countries with developing economies. **Objectives:** To study the effects of training in artificial insemination and rural management on the perception of knowledge and on the level of knowledge in rural residents affected by the armed conflict in Nariño, Colombia. **Methods:** The perception of knowledge and the theoretical-practical knowledge of 63 rural residents was evaluated on the topics of bovine agribusiness management, artificial insemination, and genetic improvement through the application of three surveys. 1) evaluated the perceived level of knowledge each rural resident had in relation to each topic, using the Likert scale, 2) evaluated the theoretical knowledge prior to training, and 3) evaluated the theoretical knowledge upon completion of training. **Results/discussion:** Of the surveyed rural residents, 54% stated that they knew how business management improved the performance of their bovine agribusiness, 54% answered the pre-training knowledge test correctly, while 83% correctly answered the post-training knowledge test. Only 6% of surveyed residents perceived that they had prior knowledge in topics related to artificial insemination and reproductive anatomy. Prior to training 35% of surveyed residents answered correctly in these topics, while upon completion of training 65% answered correctly. Regarding genetic improvement, 11% of participating rural residents stated that they knew this subject. The percentage of correct answers in this topic went from 57% to 89% prior to and post training, respectively. **Conclusion:** Rural extension programs contribute to the closing of knowledge gaps in relation to the use of reproductive biotechnologies and bovine management in rural areas affected by armed conflict.

Keywords— Agribusiness, insemination, knowledge, reproduction.

¹ Universidad Nacional de Colombia Sede-Palmira. Facultad de Ciencias Agropecuarias.

² Universidad de Sucre, Facultad de Ciencias Agropecuarias, Grupo de Investigación en Reproducción y Mejoramiento Genético Animal, ORCID 0000-0002-0326-0815. ³ Universidad de Antioquia, Facultad de Ciencias

Agrarias, Escuela de Medicina veterinaria, Grupo de Investigación OHVRI. ⁴ St. George's University, School of Medicine, St. George's, Grenada, West Indies. ⁵ Universidad Nacional de Colombia Sede-Bogotá. Facultad de Ciencias Agrarias, Departamento de Desarrollo Rural y Agroalimentario.

Evaluation of Interspecific Pollination of *Elaeis Guineensis* and *Elaeis oleifera* carried out in the Ucayali Region-Peru

Victor Sotero¹, Cindy Castro¹, Ena Velazco¹, Ursula Monteiro¹, Dora Garcia²

1. Universidad Nacional Intercultural de la Amazonia- Pucallpa-Perú, (phone: 51-998744387), e-mail: vsoteros@unia.edu.pe
2. Universidad Nacional de la Amazonia Peruana- Iquitos-Perú (pone: 51-998741535), e-mail: dora.garcia@unapiquitos.edu.pe

Abstract— The aim of this study is to carry out the evaluation of the artificial pollination of the female flowers of *E. oleifera* with pollen of *E. guineensis*, to obtain the hybrid Palma OXG, which presents two characteristics of interest, such as high resistance to the disease of spear rot and high concentration of oleic acid. The works were carried out with matrices from the experimental fields and INIA in the Province of Colonel Portillo in the Ucayali Region-Peru. From the pollination of five species of *E. oleifera*, fruits were obtained in two of them, called O7 and O68, with a percentage of 23.6% and 18.6% of fertile fruits. When germination was carried out in a controlled environment of temperature, air and humidity, only the O17 species was germinated with a yield of 68.7%.

Keywords— *Elaeis oleifera*, *Elaeis guineensis*, Palm OXG, pollination

Identification of Gm15441, a Txnip Antisense Lncrna, as a Critical Regulator in Liver Metabolic Homeostasis

Mingyang Xin¹, Qian Guo¹, Qingchun Lu¹, Juan Lu^{1,2}, Po-shun Wang¹, Yun Dong^{1,3}, Tao Li^{1,4}, Ye Chen⁵, Glenn S. Gerhard¹, Xiao-feng Yang⁶, Michael Autieri⁷, Ling Yang^{1*}

¹ Department of Medical Genetics and Molecular Biochemistry, Lewis Katz School of Medicine at Temple University, Philadelphia, PA19140, USA

² Present Address: Department of Intensive Care Unit, The First Hospital of Jilin University, Changchun 130021, China

³ Present Address: Department of Endocrinology, Nanxishan Hospital of Guangxi Zhuang Autonomous Region, Guilin 541001, China

⁴ Present Address: Department of Infectious diseases, Shandong Provincial Hospital Affiliated to Shandong First Medical University

⁵ Department of Mathematics and Statistics, Northern Arizona University, Flagsta, AZ, 86011, USA

⁶ Center for Metabolic Disease Research, Lewis Katz School of Medicine at Temple University, Philadelphia, PA19140, USA

⁷ Cardiovascular Research Center, Lewis Katz School of Medicine at Temple University, Philadelphia, PA19140, USA

* Corresponding Author:

Ling Yang, PhD

3440 N Broad St, Kresge Hall, Rm. 624

Philadelphia, PA19140, USA.

Email: ling.yang@temple.edu

Phone: +1-301-538-3931

Abstract

Background: The majority of mammalian genome is composed of non-coding regions, where numerous long non-coding RNAs (lncRNAs) are transcribed. Although lncRNAs have been identified to regulate fundamental biological processes, most of their functions remain unknown, especially in metabolic homeostasis. Analysis of our recent genome-wide screen reveals that Gm15441, a thioredoxin-interacting protein (Txnip) antisense lncRNA, is the most robustly induced lncRNA in the fasting mouse liver. Antisense lncRNAs are known to regulate their sense gene expression. Given that Txnip is a critical metabolic regulator of the liver, we aimed to investigate the role of Gm15441 in the regulation of Txnip and liver metabolism.

Methods: We examined the response of Gm15441 and Txnip under in vivo metabolic signals such as fasting and refeeding, and in vitro signals such as insulin and key metabolic transcription factors. We investigated the regulation of Txnip expression by Gm15441 and the underlying mechanism in mouse hepatocytes. Using adenovirus-mediated liver-specific overexpression, we determined whether Gm15441 regulates Txnip in the mouse liver and modulates key aspects of liver metabolism.

Results: We found that the expression levels of Gm15441 and Txnip showed a similar response pattern to metabolic signals in vivo and in vitro, but that their functions were predicted to be opposite. Furthermore, we found that Gm15441 robustly reduced Txnip protein expression in vitro through sequence-specific regulation and translational inhibition. Lastly, we confirmed the Txnip inhibition by Gm15441 in vivo (mice) and found that Gm15441 liver-specific overexpression lowered plasma triglyceride and blood glucose levels and elevated plasma ketone body levels.

Conclusions: Our data demonstrate that Gm15441 is a potent Txnip inhibitor and a critical metabolic regulator in the liver. This study reveals the therapeutic potential of Gm15441 in treating metabolic diseases.

Keywords: Gm15441, antisense lncRNA, Txnip, liver metabolism

Background

The central premise of gene expression is that DNA is transcribed into messenger RNAs (mRNAs), which then serve as templates for protein synthesis (1). However, in the past two decades, research has uncovered a large number of RNAs that are transcribed but do not encode proteins (2-6). Among the various types of non-coding RNA transcripts, long non-coding RNAs (lncRNAs), defined as RNA transcripts longer than 200 nucleotides (nt) without protein-coding potential, have attracted increasing attention (7, 8). lncRNAs can perform many functions through diverse mechanisms, including transcriptional regulation in cis or trans, organization of nuclear domains, and regulation of proteins and other RNA molecules (9). lncRNAs have been found to play crucial roles in fundamental cellular processes and various human diseases, such as cancers, cardiovascular diseases, and neurodegenerative diseases (10, 11). The newly recognized roles of lncRNAs across all taxa provide a novel perspective on the centrality of RNA in gene regulation (7). Nevertheless, the function and biological relevance of lncRNAs in metabolic homeostasis remain enigmatic (12).

Txnip is a protein-coding gene whose genetic or epigenetic variations are associated with chronic metabolic disorders such as diabetes and dyslipidemia (13-18). Overexpression of Txnip in animal models has been shown to promote hepatic glucose production, reduce insulin sensitivity, induce apoptosis of pancreatic β -cells, and decrease energy expenditure. Consistently, Txnip-deficient animals have reduced hepatic glucose production, enhanced ketogenesis, and are protected from type 1 and type 2 diabetes as well as diet-induced non-alcoholic liver disease (19, 20). Txnip has become a leading player in regulating glucose and lipid metabolism, and inhibition of Txnip is a prospective therapeutic strategy for diabetes and other metabolic disorders.

LncRNAs have been demonstrated to regulate the transcription, post-transcription, and translation of protein-coding genes (21). However, little is known about the lncRNA regulation of Txnip expression. In the current study, we sought to understand the role of a natural antisense lncRNA, Gm15441, which is transcribed from the reverse strand of Txnip, in regulating Txnip expression and metabolic homeostasis. We assessed the expression pattern of Gm15441 and Txnip in response to a variety of metabolic signals *in vivo* and *in vitro*, the regulatory role of Gm15441 on Txnip expression in mouse hepatocytes, and the physiological role of Gm15441 in the mouse liver. Altogether, our study indicates that Gm15441 is a potent translational inhibitor of Txnip and a critical metabolic regulator in the mouse liver. This study provides a better understanding of liver metabolic regulatory networks, laying the foundation for novel lncRNA-based gene therapy for metabolic diseases.

Results

Regulation of Gm15441 and Txnip expression in the liver

The liver is a crucial metabolic organ involved in energy metabolism. To identify pivotal lncRNA regulators in liver metabolism, we analyzed our recent genome-wide transcriptional profiling of mouse livers under 24-h fasting (GEO dataset GSE85439) (22). Fasting is a widely used extreme nutrient depletion condition, which can create acute metabolic stress and is suitable for identifying potential lncRNA regulators in liver metabolism. We identified Gm15441 as the most robust fasting-induced lncRNA in the mouse liver, suggesting that Gm15441 may have the potential to regulate liver metabolism (Fig. 1A). Gm15441 is a validated lncRNA in the NCBI database and is characterized as capped, spliced, and polyadenylated. Interestingly, Gm15441 is a natural antisense lncRNA to a protein-coding gene, Txnip (Fig. 1B). The expression of Gm15441 was dramatically increased in mouse livers after 24 h of fasting, but decreased after refeeding (Fig. 1C). We also found that the mRNA and protein expression levels of Txnip were significantly increased under fasting conditions and decreased after refeeding (Fig. 1, C-D). These data suggest that the expression of both genes responds similarly to changes in the nutritional state of mouse livers.

Because Txnip is associated with type 2 diabetes [19], we also analyzed the expression of Gm15441 and Txnip in the livers of *ob/ob* mice, a widely used type 2 diabetes model. Compared to lean controls, RNA expression levels of both Gm15441 and Txnip were increased in *ob/ob* mouse livers (Fig. 1E). The protein expression levels of Txnip also increased (Fig. 1F). These data reveal a potential disease association between Gm15441 and type 2 diabetes. Altogether, these results showed a similar response pattern of Gm15441 and Txnip expression to *in vivo* metabolic signals.

Regulation of Gm15441 and Txnip expression in diverse tissues

To elucidate whether this response pattern is exclusive to the liver or shared by other tissues, we analyzed the expression of Gm15441 and Txnip in six different tissues from wild-type mice under ad libitum feeding, 24-h fast, and 24-h fast followed by 4 h of refeeding, including inguinal white adipose tissue (iWAT), epididymal white adipose tissue (eWAT), kidney, intestine, gastrocnemius muscle, and liver. In all of these tissues, both Gm15441 and Txnip were increased under fasting and decreased after refeeding (Fig. 2, A-F). Nevertheless, Gm15441 in the liver increased most dramatically in response to fasting conditions among all tissues (Fig. 2F), which further supports that Gm15441 might be actively involved in regulating liver metabolism.

Regulation of Gm15441 and Txnip expression in vitro

Next, to investigate the response of Gm15441 and Txnip to metabolic signals in vitro, we first examined their regulation by essential metabolic hormones. Although Gm15441 and Txnip were induced in fasted livers, no response was observed upon fast-dominated hormone glucagon stimulation in vitro (Fig. 3A). However, we found that the fed-dominated hormone insulin inhibited both Gm15441 and Txnip expression (Fig. 3A), which is consistent with their reduction in refed livers.

Furthermore, we investigated whether Gm15441 and Txnip are regulated by master transcription factors that respond to nutrient changes. Forkhead box O1 (FOXO1) is a crucial fasting-induced transcription factor that activates gluconeogenesis and suppresses lipogenesis (23, 24). Hepatocyte nuclear Factor-4 α (HNF4 α) is a master transcription factor for liver gene expression and regulates gluconeogenesis during fasting (25, 26). We found that both Gm15441 and Txnip were significantly upregulated by FOXO1 but not HNF4 α (Fig. 3B). Intriguingly, FOXO1 is a well-known transcription factor that is inactivated by insulin. Therefore, the induction of Gm15441 and Txnip by FOXO1 is consistent with their inhibition by insulin, suggesting the potential involvement of Gm15441 and Txnip in liver glucose metabolism.

In addition, we investigated whether Gm15441 and Txnip are responsive to key nutrient-sensing nuclear receptors in the liver. Peroxisome proliferator-activated receptor- α (PPAR α) is a fasting-induced nuclear receptor, and the farnesoid X receptor (FXR) is a feeding-induced nuclear receptor. PPAR α promotes while FXR suppresses fatty acid oxidation and gluconeogenesis in the liver (27-30). In vitro treatment with PPAR α agonist WY14643 (WY) and GW7647 in mouse primary hepatocytes significantly increased the RNA expression levels of Gm15441 and Txnip, whereas the FXR agonist GW4064 robustly suppressed their expression (Fig. 3C). Likewise, Txnip protein levels were upregulated by PPAR α agonist and downregulated by FXR agonist (Fig. 3D). The regulation of Gm15441 and Txnip expression by PPAR α and FXR is consistent with their induction in fasted livers and suppression in refed livers. Consistently, the ChIP-sequencing data retrieved from the Gene Transcription Regulation Database (GTRD) revealed that multiple Foxo1 and PPAR α binding sites are located in the promoters and/or gene bodies of Gm15441 and Txnip (Supplementary Fig. 1). These data suggest that PPAR α contributes to the induction of both gene expression in the fasting condition, whereas insulin-FOXO1 and FXR contribute to the inhibition in the refed state. Collectively, our data showed that Gm15441 shares a similar expression pattern with Txnip in response to in vivo and in vitro metabolic signals, suggesting a potential close functional relationship.

Gm15441 is an inhibitor of Txnip protein expression

Antisense lncRNAs have been found to regulate the expression of their corresponding sense genes (31). The antisense lncRNA Gm15441 responds to metabolic signals similar to its sense gene Txnip, suggesting

that Gm15441 may function as either a booster or inhibitor of Txnip expression. To test this, we first used a bioinformatics approach to predict the functions of Txnip and Gm15441. It is a widely accepted approach to predict gene function by performing gene ontology analysis of its highly correlated genes. Thus, to predict the metabolic function of Txnip and Gm15441 in the liver, we performed gene ontology analysis of Txnip and Gm15441 correlated genes using gene expression profiles from the GEO dataset GSE85439 (22). The identified mRNAs that are highly correlated with Txnip and Gm15441 were then subject to Gene Ontology (GO) analysis. GO analysis of mRNAs highly correlated with Txnip indicated that these mRNAs are involved in gluconeogenesis and triglyceride synthesis, consistent with the known functions of Txnip in the liver (Fig. S2A). Therefore, we reasoned that the GO analysis of mRNAs highly correlated with Gm15441 could be used to predict the function of Gm15441 in the liver. Intriguingly, the GO analysis results suggest that Gm15441 may play opposite roles to Txnip in the liver (Fig. S2A and S2B). For example, Txnip was positively correlated with triglyceride biosynthesis; however, Gm15441 was negatively correlated with triglyceride biosynthesis (Fig. S2B). The complete list of mRNAs highly correlated with Gm15441 and Txnip is provided in the Supplementary Material. These results suggest that Gm15441 might be an inhibitor of Txnip expression.

To further investigate whether Gm15441 is an inhibitor of Txnip expression, we cloned Gm15441 using mouse liver cDNA. We found two Gm15441 isoforms, a relatively long (Gm15441-L) isoform and a short (Gm15441-S) isoform (Fig. 4A). The sequence alignment of Gm15441 to the Txnip mRNA indicates that the Gm15441-L isoform has an overlapping region with the Txnip 5' end (5'UTR and part of the first exon), whereas the Gm15441-S isoform does not (Fig. 4A). A recent study indicated that translational regulation of the human TXNIP gene could occur via an internal ribosome entry site (IRES) located in the Txnip 5'UTR (32). Thus, we reasoned that Gm15441-L isoform which contains an overlapping region with the Txnip 5' end would be able to regulate Txnip expression. Therefore, we cloned Txnip, which contains both coding sequence (CDS) and 5' UTR, from mouse liver cDNA (Fig. 4A). We then investigated whether Gm15441-L regulates Txnip expression. As shown in Fig. 4B, the ΔCt values of Gm15441 is about 1.5 Ct less than Txnip in Gm15441-L and Txnip co-overexpression mouse primary hepatocytes, suggesting that the expression level of Gm15441 RNA is around 3-fold ($2^{1.5}$ Ct) higher than Txnip mRNA. We also found that the Gm15441-L co-overexpression with Txnip dramatically reduced Txnip protein but not its RNA levels (Fig. 4, C-D). Moreover, we found that Gm15441 knockdown can counteract the Txnip protein reduction and restore levels, without interfering Txnip RNA expression (Fig. 4, E-F). Next, we investigated whether Gm15441 regulates endogenous Txnip. We found that Gm15441-L overexpression dramatically reduced endogenous Txnip protein but not its RNA levels in mouse hepatocytes AML12 (Fig. 5, A-B). We also found that knockdown of Gm15441 can also dramatically increase endogenous Txnip protein but not its RNA levels in mouse hepatocytes AML12 (Fig. 5, C-D). Altogether, these results suggest that Gm15441-L inhibits Txnip protein expression. Our data also imply that this inhibition is neither through transcriptional regulation nor affecting mRNA stability, but potentially through translational regulation.

Gm15441 inhibits Txnip protein through sequence specificity and translational regulation

Next, we aimed to further determine whether Gm15441-L regulates Txnip expression through its sequence overlapping with the Txnip 5' end. Compared to Gm15441-L, Gm15441-S lacks the sequence overlapping with the Txnip 5' end. Therefore, Gm15441-S could serve as a natural negative control for investigating the sequence specificity. We found that Gm15441-S could not reduce Txnip protein levels (Fig. 6A). This result suggests that the sequence in Gm15441-L overlapping with the Txnip 5' end is necessary for Gm15441-L to inhibit Txnip protein expression. To further confirm this, the 267-709 nt fragment of Gm15441-L, which overlaps with the Txnip 5' end, was subcloned. As shown in Fig. 6B, Txnip protein

expression was reduced to a similar level when Txnip was co-overexpressed with Gm15441-L or Gm15441-L 267-709 nt. Next, we divided Gm15441-L 267-709 nt into two pieces: Gm15441-L 267-430 nt, which is complementary to the entire 5'UTR of Txnip, and Gm15441-L 431-709 nt, which is complementary to part of the first exon of Txnip. We found that both Gm15441-L 267-430 nt and Gm15441-L 431-709 nt can suppress Txnip protein expression. These results support the sequence specificity of Gm15441-L 267-709 nt in the suppression of Txnip protein expression.

Next, we investigated whether Gm15441-L inhibits Txnip protein expression through translational regulation. We performed fluorescence in situ hybridization (FISH) to detect the localization of Gm15441-L and Txnip RNAs. As shown in Fig. 6C, Gm15441-L RNAs co-localized with Txnip RNAs. The colocalization between Gm15441 and Txnip RNAs suggests a potential interaction. We performed RNA-RNA pull down assays in mouse hepatocytes Hepa1-6 overexpressed with Txnip and Gm15441. We used Gm15441-specific oligo probes and non-specific oligo probe as control. We found that Gm15441 RNA was successfully pulled down by Gm15441-specific probes, so does Txnip mRNA (Fig. 6D). These results suggest a direct interaction between Gm15441 RNA and Txnip mRNA, which supports the idea that Gm15441 regulates Txnip mRNA translation. Furthermore, we used the translational inhibitor cycloheximide (CHX) to test this hypothesis. We found that the Txnip protein experienced a significant decline after CHX treatment, but the presence of Gm15441-L did not affect the rate of decline of Txnip protein (Fig. 6E). Because CHX treatment blocks Txnip translation, the rate of decline of Txnip protein is determined by protein degradation. Thus, the unchanged rate of decline of Txnip protein suggests that Gm15441-L does not affect Txnip protein degradation. Therefore, we concluded that Gm15441-L reduced Txnip protein through translational inhibition. Altogether, these results suggest that Gm15441-L is a potent translational inhibitor of Txnip, and the sequence 267-709 nt in Gm1544-L is necessary for Gm15441-L to inhibit Txnip protein expression.

Gm15441 regulates Txnip protein expression and energy metabolism in the mouse liver

To gain insight into how Gm15441 regulates Txnip in an in vivo setting, we used adenovirus to specifically overexpress Gm15441-L, Gm15441-S, or empty vector (control) in the mouse liver. Both isoforms of Gm15441 were successfully increased in the mouse liver after adenovirus administration (Fig. 7A). The Δ Ct values of Gm15441 is about 5 Ct less than Txnip in Gm15441-L or Gm15441-S overexpression livers, suggesting that the expression level of Gm15441 RNA is around 32-fold (2^5 Ct) higher than Txnip mRNA (Fig. 7B). Interestingly, significant changes were observed in glucose and lipid metabolic gene expression (Fig. 7A). Consistent with the in vitro results, Gm15441-L overexpression did not affect Txnip mRNA expression levels, but significantly decreased Txnip protein levels in vivo (Fig. 7, A and C). Moreover, Gm15441-S did not affect either mRNA or protein levels of Txnip (Fig. 7, A and C). These results strongly supporting that Gm15441-L downregulates Txnip protein through its sequence overlapping with the Txnip 5'end.

Previous studies have shown that Txnip regulates glucose and lipid metabolism (20, 33-41). Therefore, we reasoned that Gm15441-L inhibition of Txnip protein regulates glucose and lipid metabolism in mice. First, there is no significant difference in body weight after Gm15441-L or Gm15441-S overexpression (Fig. 7D). Next, we found that liver-specific overexpression of Gm15441-L, but not Gm15441-S, lowered blood glucose levels, increased plasma ketone bodies beta-hydroxybutyrate (BOH), and reduced plasma triglyceride (TG) levels (Fig. 7, E-G). These results are consistent with previous findings in Txnip knockout mice (40). Liver TG content was unaffected by either of the Gm15441 isoforms overexpressed in mice (Fig. 7H). These results suggest that Gm15441 is a key metabolic regulator in the liver.

Discussion

Recently, the novel roles of lncRNAs in both gene expression and energy metabolism have attracted increasing attention. This study investigated the regulatory relationship between the antisense lncRNA Gm15441 and its sense coding gene Txnip and revealed the role of Gm15441 in the liver. We showed that Txnip protein levels were suppressed by Gm15441 via its sequence overlapping with the Txnip 5' end. We also found that Gm15441 inhibits Txnip protein through translational inhibition. Furthermore, we demonstrated that liver-specific overexpression of Gm15441 in mice increases plasma ketone bodies while lowering circulating TG and blood glucose levels. Altogether, this study indicates that Gm15441 is a potent inhibitor of Txnip protein expression and a critical regulator of liver metabolic homeostasis.

Previous studies have demonstrated that translation of Txnip is mediated by the IRES located in the Txnip 5'UTR (32). Compared to Gm15441-L, Gm15441-S lacks the region overlapping with the Txnip 5'UTR. Here, we found that Gm15441-S did not affect either Txnip protein or RNA levels. In addition, liver-specific overexpression of Gm15441-S in mice did not affect blood glucose, plasma BOH, and TG levels. However, the exact function of Gm15441-S remains unclear and requires further investigation.

A recent study showed that knockdown of Gm15441 reduced BOH levels and increased TG content in cultured mouse hepatocytes (42). The results of this *in vitro* loss-of-function study strongly supports the phenotypes observed in our liver-specific Gm15441 overexpression mice with increased BOH and decreased plasma TG. In addition, a whole-body Gm15441 knockout mouse model showed increased serum TG levels and Txnip protein expression in the liver (43). These findings are consistent with our liver-specific Gm15441 overexpression mice, which showed decreased serum TG levels and Txnip protein expression in the liver. We investigated the function of Gm15441 under normal physiological conditions. Disease models, such as high fat diet induced obesity, could be used to further investigate the role of Gm15441 in metabolic diseases in the future.

Besides, expression of Gm15441 is not exclusive to the liver, which suggests that Gm15441 might play important roles in other metabolic organs. Thus, future studies on the functional role and mechanism of Gm15441 in other tissues are required to fully understand the importance of Gm15441 in the energy metabolism. Moreover, we found that both Txnip and Gm15441 can be regulated by various metabolic signals and transcription factors. Unraveling the complexity of these regulatory mechanisms may provide new directions for the roles of Txnip and Gm15441.

Conclusions

Our study demonstrated that lncRNA Gm15441 dynamically responds to *in vivo* and *in vitro* metabolic signals in a similar pattern to its sense gene Txnip. We also found that Gm15441 reduced Txnip protein expression through sequence specificity and translational inhibition. We further demonstrated that liver-specific overexpression of Gm15441 regulates glucose and lipid metabolism. Overall, our study indicates that lncRNA Gm15441 is a potent translational inhibitor of Txnip and a critical metabolic regulator in the liver. Inhibition of Txnip is considered a promising therapeutic strategy for metabolic disorders, such as diabetes (19). The approach in the present study, in which overexpression of Gm15441 was used to inhibit Txnip expression, sets a direction for future Gm15441-based gene therapy for metabolic diseases.

Methods

Animal experiments

All animal protocols were approved by Temple University Institutional Animal Care and Use Committee (IACUC). All the mice were purchased from the Jackson Laboratory and were acclimatized to the Temple animal housing unit for 10-14 days before any experiments. For experiments under fasting and refeeding conditions, 10-week old male C57BL/6 mice with free access to water and normal chow diet were used as the control (Ad Libitum) group, mice subjected to a 24-h fasting period before being euthanized for tissue harvest were used as the fasting group, and mice that fasted for 24-h followed by a 4-h normal chow diet feeding before tissue harvest were used as the refeeding group. For ob/ob mouse model, 10-week old ob/ob and lean control mice were subjected to a 4-h food withdrawal before tissue harvest. For Gm15441-L/S liver specific overexpression mouse model, 9-week old male C57BL/6 mice were injected with adenovirus expressing empty vector as control, or Gm15441-L, or Gm15441-S intravenously at 2×10^9 plaque forming units (pfu) per mouse (9 weeks old). Seven to eleven days post injection, mice were subjected to a 24-h food withdrawal before tissue harvest, then liver and blood samples were harvested for further analysis.

Isolation and culture of mouse primary hepatocytes

Mouse primary hepatocytes were isolated from 8-12 weeks old male C57BL/6 mice as described previously [22]. Briefly, immediately after anesthesia with Ketamine (100 mg/kg) and Xylazine (10 mg/kg), mice livers were perfused with Krebs Ringer buffer and digested using collagenase (Liberase TM Research Grade, Roche). Isolated hepatocytes were then purified with Percoll. Only cells with viability over 90% determined by trypan blue were seeded onto collagen-coated plates in DMEM supplemented with 5.5 mM glucose, 2 mM GlutaMAX™, and 10% Cosmic Calf Serum. 4 h after plating, cells were switched to maintenance medium DMEM supplemented with 5.5 mM glucose and 2 mM GlutaMAX™. Details of drug treatment are indicated in each experiment.

Cell culture and siRNA transfection

The mouse hepatocyte cell line AML12 cells (purchased from ATCC) were cultured in a 1:1 mixture of Dulbecco's modified Eagle's medium and Ham's F12 medium (Invitrogen) with 10% CCS, ITS (Invitrogen) and dexamethasone (40ng/ml). For siRNA transfection, 60% confluent AML12 cells were cultured with siRNA specifically targeting Gm15441 or siLacZ as control for 48 hours using Invitrogen RNAiMax transfection reagent. Then RNA and protein were harvested and quantified by real-time PCR analysis. The control siRNA targeting LacZ: sense CUACACAAAUCAGCGAUUU, antisense AAAUCGCUGAUUUGUGUAG; The siRNA sequence targeting Gm15441: sense GACGAGAACUUGUCAGUA, antisense UAUCUGACAAGUUCUCGUC.

Adenovirus production

Gm15441, Txnip containing both the 5' untranslated region (5'UTR) and coding sequence (CDS), and the constitutively active HNF4 α were cloned from mouse liver cDNA. The following primers were used for cloning: Gm15441 (Forward: 5'-GGAGCAAGCCGATAAGCAG, Reverse: 3'-ACATTTAAATTTTTATTTGGGTGTCTCTGGAGTG), Txnip (5'UTR-Forward: 5'-GACACTCTCCTCCTCTGGTCTC, Reverse: 3'-CTGCACGTTGTTGTTGTTGTT), HNF4 α (Forward: 5'-ATGCGACTCTCTAAAACCCTT, Reverse: 3'-CTAGATGGCTTCTTGCTTGGTGATC).

Gm15441, Txnip, HNF4 α , YFP (Addgene plasmid #15302), constitutively active mouse FOXO1 (Addgene plasmid #17547) were subcloned into adenoviral vector pAd/CMV/V5-DEST (Invitrogen). Adenoviruses were amplified in HEK293A cells and purified by CsCl density-gradient ultracentrifugation, then desalted using PD10 columns (GE Healthcare Life Sciences) and titered with an Adeno-X Rapid Titer Kit (Clontech). For in vitro adenovirus experiments, primary mouse hepatocytes were transduced with each adenovirus at a level of 50 multiplicity of infection (MOI) for 2 h, then samples were collected 24-h post transduction.

RNA extraction and quantitative real-time PCR analysis

Total RNA was isolated from mouse liver or primary hepatocytes using Trizol reagent (Invitrogen) followed by a Turbo DNA-free DNase treatment (Ambion). cDNA was generated using a reverse transcription system (SuperScript® III First-Strand Synthesis System, Invitrogen). Quantitative real-time PCR was performed using a real-time PCR system (Mastecycler; Eppendorf). The relative amount of mRNA in each sample was normalized to 18S transcript levels. The sequences for gene-specific RT-PCR primers are attached in Table S1.

Western blotting

Mouse liver samples and cells were lysed in RIPA buffer (Cell Signaling Technology) containing protease and phosphatase inhibitors (Thermos Scientific). The lysates were subjected to SDS-PAGE, transferred to polyvinylidene fluoride (PVDF) membranes, and incubated with the primary antibody followed by the fluorescence conjugated secondary antibody (LI-COR). The bound antibody was visualized using a quantitative fluorescence imaging system (LI-COR). The relative amount of protein in each sample was normalized to β -Actin protein levels. Primary antibodies include a mouse monoclonal anti-TXNIP antibody, clone JY2 (1:500; MABS1225, Sigma-Aldrich), and a rabbit monoclonal anti- β -Actin antibody (1:1000; 8457S, Cell Signaling Technology).

RNA in situ hybridization (ISH) Assays

RNA in situ hybridization (ISH) for Gm15441 and Txnip RNA was performed manually using a commercially available kit, RNAscope® Fluorescent Multiplex Reagent Kit (Advanced Cell Diagnostics, Inc., Hayward, CA), and RNAscope® Oligo Probes specific to Gm15441 and Txnip according to the manufacturer's instructions. Briefly, monolayer formalin-fixed Hepal-6 cell sections were pretreated with protease prior to hybridization with the Gm15441 and Txnip oligo probe mixture. Then, a cascade of signal amplification molecules was hybridized sequentially, followed by fluorescent dye-labeled probes hybridization. Target RNA signals were visualized by confocal microscopy.

RNA pull-down assay

Hepal-6 cells were seeded at 30~35% to a 10-cm cell culture dish; 24 hours later, cells were transduced with adenovirus expressing Txnip and Gm15441-L, MOI=50; 24h post transfection, cells were washed twice with 1X PBS, then 1ml lysis buffer (25 mM Tris_Cl pH7.4, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 5% glycerol), supplemented with protease inhibitors and Rnase inhibitors (200 U/mL) was added. Cell lysates were then on ice for 30 min, and centrifuge at 12000 rpm at 4 °C. 20 μ L of the supernatant

were saved as the input sample, 450ul μ L of the supernatant sample were then incubate with 100 pmol of non-specific oligo probe (GTTTGTGGTTTAACAGTGGGAAGGC/3BioTEG/) or Gm15441 probe mixture (TGAAGTCTTATGTAGCTGGGGCTGA/3BioTEG/, CACCAGAGCATTACCAGAAAGGAC/3BioTEG/) on a tube rotator under moderate agitation at RT for 6-8h. Then, cell lysates were incubated with 100 ul of Streptavidin Dynabeads M-270 (Invitrogen) under moderate agitation on the tube rotator at RT overnight. Beads were then washed briefly five times with lysis buffer and resuspended in 1 ml of TRI reagent. RNAs were isolated using Direct-zol RNA MiniPrep kit (Zymo) and analyzed by RT-PCR.

Blood and liver metabolite measurements

Blood glucose levels were assayed from a tail-clip and using an Ascensia Elite XL glucometer (Bayer Co.) after an overnight fast. Blood samples were collected after 24-h fasting by cardiac puncture using heparinized 25 G needles with 1 ml syringes during terminal anesthesia. Plasma samples were obtained by centrifuging blood samples at 5000 x g at 4 °C. Liver triglyceride (TG) content of mice and serum levels of TG were assayed by Triglyceride Determination kit (Sigma, TR0100). Plasma β -hydroxybutyrate (BOH) levels were assayed by β -HB Ketone Body Colorimetric Assay Kit (Cayman).

Bioinformatics analysis

The genome-wide transcriptional profiling (GSE85439) of mouse livers under 24-h fasting and Ad Libitum was used to identify potential lncRNA regulators in liver metabolism ref (22). Expression correlation analysis was performed to predict the function of Txnip and Gm15441 in the liver metabolism using liver samples (n=37) from the same datasets (GSE85439). Briefly, gene expression profiles of liver tissues from mice under Ad Libitum (n=4), 24-hour fasting (n=4), refeeding (n=4), normal diet (n=5), 48-hour high fat diet (n=5), 12-week high fat diet (n=5), and lean control mice (n=5) as well as ob/ob mice (n=5) were subjected to correlation analysis. We evaluated the co-expression of Gm15441 or Txnip with all detectable mRNAs in the 37 liver samples. The correlation coefficients of Gm15441 with mRNAs or Txnip with mRNAs were made by R packages based on the Pearson correlation method. The mRNAs that showed correlation coefficients > 0.7 (positive) or < 0.7 (negative) were used to perform GO analysis separately. GO analyses were performed by DAVID Bioinformatics Resources 6.7 (<https://david.ncifcrf.gov>) ($p < 0.05$).

Statistical Analyses

Comparisons between groups and the difference of the slopes were assessed by a two-tailed unpaired Student's t-test. Multiple comparisons were statistically evaluated by a one-way analysis of variance (ANOVA) followed by Bonferroni or Dunnett post-hoc test. Graphs are presented as means \pm SEM. Statistical significance was determined by $p < 0.05$.

Abbreviations

lncRNAs: long non-coding RNAs
 Txnip: Thioredoxin-interacting protein
 mRNAs: Messenger RNAs
 nt: Nucleotide

WAT: Inguinal white adipose tissue
eWAT : Epididymal white adipose tissue
FOXO1: Forkhead box O1
HNF4 α : Hepatocyte nuclear Factor-4 α
PPAR α : Peroxisome proliferator-activated receptor- α
FXR: Farnesoid X receptor (FXR)
WY: WY14643
GTRD: Gene Transcription Regulation Database
GO: Gene Ontology
IRES: Internal ribosome entry site
CDS: Coding sequence
FISH: Fluorescence in situ hybridization
CHX: Cycloheximide
BOH: beta-hydroxybutyrate
TG: triglyceride

Declarations

Ethics approval and consent to participate

All animal protocols were approved by Temple University Institutional Animal Care and Use Committee (IACUC).

Consent for publication

Not applicable

Availability of data and materials

Not applicable

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by NIH grant K22HL139921.

Author Contributions

MX, QL, QG, JL, PW, YD, and TL performed and analyzed the experiments. YC and LY performed the bioinformatics analysis. GG, XY, and MA provided important intellectual input to the manuscript. MX and LY wrote the manuscript. LY conceived and supervised the study. All authors read and approved the final manuscript.

References

1. Rinn JL, and Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem.* 2012;81:145-66.
2. Weinberg RA, and Penman S. Small molecular weight monodisperse nuclear RNA. *J Mol Biol.* 1968;38(3):289-304.
3. Paul J, and Duerksen JD. Chromatin-associated RNA content of heterochromatin and euchromatin. *Mol Cell Biochem.* 1975;9(1):9-16.
4. Salditt-Georgieff M, Harpold MM, Wilson MC, and Darnell JE. Large heterogeneous nuclear ribonucleic acid has three times as many 5' caps as polyadenylic acid segments, and most caps do not enter polyribosomes. *Mol Cell Biol.* 1981;1(2):179-87.
5. Salditt-Georgieff M, and Darnell JE. Further evidence that the majority of primary nuclear RNA transcripts in mammalian cells do not contribute to mRNA. *Mol Cell Biol.* 1982;2(6):701-7.
6. Nickerson JA, Krochmalnic G, Wan KM, and Penman S. Chromatin architecture and nuclear RNA. *Proc Natl Acad Sci U S A.* 1989;86(1):177-81.
7. Kopp F, and Mendell JT. Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell.* 2018;172(3):393-407.
8. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 2012;22(9):1775-89.
9. Ulitsky I, and Bartel DP. lincRNAs: genomics, evolution, and mechanisms. *Cell.* 2013;154(1):26-46.
10. Kornfeld JW, and Brüning JC. Regulation of metabolism by long, non-coding RNAs. *Front Genet.* 2014;5:57.
11. Li J, Xuan Z, and Liu C. Long non-coding RNAs and complex human diseases. *Int J Mol Sci.* 2013;14(9):18790-808.
12. Losko M, Kotlinowski J, and Jura J. Long Noncoding RNAs in Metabolic Syndrome Related Disorders. *Mediators Inflamm.* 2016;2016:5365209.
13. Kowluru RA, and Mishra M. Contribution of epigenetics in diabetic retinopathy. *Sci China Life Sci.* 2015;58(6):556-63.
14. De Marinis Y, Cai M, Bompada P, Atac D, Kotova O, Johansson ME, et al. Epigenetic regulation of the thioredoxin-interacting protein (TXNIP) gene by hyperglycemia in kidney. *Kidney Int.* 2016;89(2):342-53.
15. Ramus SM, Cilensek I, Petrovic MG, Soucek M, Kruzliak P, and Petrovic D. Single nucleotide polymorphisms in the Trx2/TXNIP and TrxR2 genes of the mitochondrial thioredoxin antioxidant system and the risk of diabetic retinopathy in patients with Type 2 diabetes mellitus. *J Diabetes Complications.* 2016;30(2):192-8.
16. van Greevenbroek MM, Vermeulen VM, Feskens EJ, Evelo CT, Kruijshoop M, Hoebee B, et al. Genetic variation in thioredoxin interacting protein (TXNIP) is associated with hypertriglyceridaemia and blood pressure in diabetes mellitus. *Diabet Med.* 2007;24(5):498-504.
17. Ferreira NE, Omae S, Pereira A, Rodrigues MV, Miyakawa AA, Campos LC, et al. Thioredoxin interacting protein genetic variation is associated with diabetes and hypertension in the Brazilian general population. *Atherosclerosis.* 2012;221(1):131-6.
18. Chambers JC, Loh M, Lehne B, Drong A, Kriebel J, Motta V, et al. Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: a nested case-control study. *Lancet Diabetes Endocrinol.* 2015;3(7):526-34.
19. Alhawiti NM, Al Mahri S, Aziz MA, Malik SS, and Mohammad S. TXNIP in Metabolic Regulation: Physiological Role and Therapeutic Outlook. *Curr Drug Targets.* 2017;18(9):1095-103.

20. Park MJ, Kim DI, Lim SK, Choi JH, Kim JC, Yoon KC, et al. Thioredoxin-interacting protein mediates hepatic lipogenesis and inflammation via PRMT1 and PGC-1 α regulation in vitro and in vivo. *J Hepatol.* 2014;61(5):1151-7.
21. Zhang X, Wang W, Zhu W, Dong J, Cheng Y, Yin Z, et al. Mechanisms and Functions of Long Non-Coding RNAs at Multiple Regulatory Levels. *Int J Mol Sci.* 2019;20(22).
22. Yang L, Li P, Yang W, Ruan X, Kiesewetter K, Zhu J, et al. Integrative Transcriptome Analyses of Metabolic Responses in Mice Define Pivotal LncRNA Metabolic Regulators. *Cell Metab.* 2016;24(4):627-39.
23. Cheng Z, and White MF. Targeting Forkhead box O1 from the concept to metabolic diseases: lessons from mouse models. *Antioxid Redox Signal.* 2011;14(4):649-61.
24. Gross DN, van den Heuvel AP, and Birnbaum MJ. The role of FoxO in the regulation of metabolism. *Oncogene.* 2008;27(16):2320-36.
25. Hayhurst GP, Lee YH, Lambert G, Ward JM, and Gonzalez FJ. Hepatocyte nuclear factor 4 α (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. *Mol Cell Biol.* 2001;21(4):1393-403.
26. Lu H. Crosstalk of HNF4. *Acta Pharm Sin B.* 2016;6(5):393-408.
27. Evans RM, Barish GD, and Wang YX. PPARs and the complex journey to obesity. *Nat Med.* 2004;10(4):355-61.
28. Chawla A, Saez E, and Evans RM. "Don't know much bile-ology". *Cell.* 2000;103(1):1-4.
29. Patsouris D, Mandard S, Voshol PJ, Escher P, Tan NS, Havekes LM, et al. PPAR α governs glycerol metabolism. *J Clin Invest.* 2004;114(1):94-103.
30. Ma K, Saha PK, Chan L, and Moore DD. Farnesoid X receptor is essential for normal glucose homeostasis. *J Clin Invest.* 2006;116(4):1102-9.
31. Villegas VE, and Zaphiropoulos PG. Neighboring gene regulation by antisense long non-coding RNAs. *Int J Mol Sci.* 2015;16(2):3251-66.
32. Lampe S, Kunze M, Scholz A, Brauß TF, Winslow S, Simm S, et al. Identification of the TXNIP IRES and characterization of the impact of regulatory IRES trans-acting factors. *Biochim Biophys Acta Gene Regul Mech.* 2018;1861(2):147-57.
33. Ding C, Zhao Y, Shi X, Zhang N, Zu G, Li Z, et al. New insights into salvianolic acid A action: Regulation of the TXNIP/NLRP3 and TXNIP/ChREBP pathways ameliorates HFD-induced NAFLD in rats. *Sci Rep.* 2016;6:28734.
34. Wang W, Wang C, Ding XQ, Pan Y, Gu TT, Wang MX, et al. Quercetin and allopurinol reduce liver thioredoxin-interacting protein to alleviate inflammation and lipid accumulation in diabetic rats. *Br J Pharmacol.* 2013;169(6):1352-71.
35. Zhang X, Zhang JH, Chen XY, Hu QH, Wang MX, Jin R, et al. Reactive oxygen species-induced TXNIP drives fructose-mediated hepatic inflammation and lipid accumulation through NLRP3 inflammasome activation. *Antioxid Redox Signal.* 2015;22(10):848-70.
36. Mohamed IN, Sarhan NR, Eladl MA, El-Remessy AB, and El-Sherbiny M. Deletion of Thioredoxin-interacting protein ameliorates high fat diet-induced non-alcoholic steatohepatitis through modulation of Toll-like receptor 2-NLRP3-inflammasome axis: Histological and immunohistochemical study. *Acta Histochem.* 2018;120(3):242-54.
37. Zheng T, Yang X, Li W, Wang Q, Chen L, Wu D, et al. Salidroside Attenuates High-Fat Diet-Induced Nonalcoholic Fatty Liver Disease via AMPK-Dependent TXNIP/NLRP3 Pathway. *Oxid Med Cell Longev.* 2018;2018:8597897.
38. He K, Zhu X, Liu Y, Miao C, Wang T, Li P, et al. Inhibition of NLRP3 inflammasome by thioredoxin-interacting protein in mouse Kupffer cells as a regulatory mechanism for non-alcoholic fatty liver disease development. *Oncotarget.* 2017;8(23):37657-72.

39. Shimizu H, Tsubota T, Kanki K, and Shiota G. All-trans retinoic acid ameliorates hepatic stellate cell activation via suppression of thioredoxin interacting protein expression. *J Cell Physiol.* 2018;233(1):607-16.
40. Chutkow WA, Patwari P, Yoshioka J, and Lee RT. Thioredoxin-interacting protein (Txnip) is a critical regulator of hepatic glucose production. *J Biol Chem.* 2008;283(4):2397-406.
41. Jo SH, Kim MY, Park JM, Kim TH, and Ahn YH. Txnip contributes to impaired glucose tolerance by upregulating the expression of genes involved in hepatic gluconeogenesis in mice. *Diabetologia.* 2013;56(12):2723-32.
42. Batista TM, Garcia-Martin R, Cai W, Konishi M, O'Neill BT, Sakaguchi M, et al. Multi-dimensional Transcriptional Remodeling by Physiological Insulin In Vivo. *Cell Rep.* 2019;26(12):3429-43.e3.
43. Brocker CN, Kim D, Melia T, Karri K, Velenosi TJ, Takahashi S, et al. Long non-coding RNA Gm15441 attenuates hepatic inflammasome activation in response to PPARA agonism and fasting. *Nat Commun.* 2020;11(1):5847.

Figures And Figure Legends

Figure 1. Expression patterns of Gm15441 and Txnip are similar in mouse livers. (A) Top 5 upregulated lncRNAs in the fasting mice livers according to the genome-wide transcriptional profiling; (B) Schematic diagram of Gm15441 and Txnip genomic organization. Lines and boxes indicate introns and exons, respectively. Exons are numbered for each gene; (C) Quantitative real-time PCR analysis of liver tissues from C57BL/6 wildtype mice fed ad libitum, subject to a 24-h fast, or a 24-h fast followed by a 4-h refeed, n=4; (D) Western blot analysis of liver tissues from C57BL/6 wildtype mice fed ad libitum, subject to a 24-h fast, or a 24-h fast followed by a 4-h refeed. Quantitation is shown in the right panel, n=3; (E) Quantitative real-time PCR analysis of liver tissues from ob/ob mice and their lean control mice subject to a 4-h food withdrawal, n=4; (F) Western blot analysis of liver tissues from ob/ob mice and their lean control mice subject to a 4-h food withdrawal. Quantitation is shown in the right panel, n=4; Error bars are SEM, *P<0.05 (Fast versus Ad Libitum or ob/ob versus control), #P<0.05 (Refeed versus Fast).

Figure 2. Expression patterns of Gm15441 and Txnip are similar in multiple mouse tissues. (A-F) Quantitative real-time PCR analysis of inguinal white adipose tissue (iWAT) (A), epididymal white adipose tissue (eWAT) (B), kidney (C), intestine (D), muscle (E), and liver (F). Each tissue sample was pooled from mice (n=3) under each condition, including fed ad libitum, subject to a 24-h fast, or a 24-h fast followed by a 4-h refeed; 3 technical replicates were used in each experiment. Error bars are SEM, *P<0.05 (Fast versus Ad Libitum), #P<0.05 (Refeed versus Fast).

Figure 3. Expression patterns of Gm15441 and Txnip are similar upon metabolic signal stimulation in vitro. (A) Quantitative real-time PCR analysis of mouse primary hepatocytes treated with 100 nM insulin for 24 h or 200 nM glucagon for 6 h. (B) Quantitative real-time PCR analysis of mouse primary hepatocytes transduced with adenovirus expressing YFP, FOXO1, or HNF4 α for 24 h, MOI=50; (C) Quantitative real-time PCR analysis of mouse primary hepatocytes treated with 0.1% DMSO (vehicle), 100 μ M WY, 20 μ M GW7647, or 10 μ M GW4064 for 24 h; (D) Western blot analysis of mouse primary hepatocytes treated with 0.1% DMSO (vehicle), 100 μ M WY, 20 μ M GW7647, or 10 μ M GW4064 for 24 h. Quantitation is shown in the right panel. Error bars are SEM, n=3, *P<0.05.

Figure 4. Gm15441 reduces exogenous Txnip protein levels in vitro. (A) Sequence alignment of Txnip, Gm15441 long isoform (Gm15441-L), and Gm15441 short isoform (Gm15441-S). Gm15441-L and Gm15441-S share a common sequence (grey boxes) complimentary to Txnip CDS. Besides, Gm15441-L has a unique sequence (green box) complimentary to the Txnip 5'end. Common sequences are labelled with the same color. (B) Comparison of ΔC_t values between Gm15441 and Txnip (normalized to the housekeeping gene 18S) from quantitative real-time PCR analysis of mouse primary hepatocytes transduced with Gm15441-L or Txnip adenovirus for 24 h, MOI=50; (C) Quantitative real-time PCR analysis of mouse primary hepatocytes transduced with adenovirus expressing YFP + empty vector, YFP + Gm15441-L, Txnip + empty vector, or Txnip + Gm15441-L, for 24 h, MOI=50; (D) Western blot analysis of mouse primary hepatocytes transduced with adenovirus expressing YFP + empty vector, YFP + Gm15441-L, Txnip + empty vector, or Txnip + Gm15441-L, for 24 h, MOI=50; Quantitation is shown in the lower panel. Error bars are SEM, n=3, *P<0.05. N.S., not significant. (E-F) Quantitative real-time PCR analysis (E) and western blot analysis (F) of mouse hepatocytes AML12 transduced with adenovirus expressing Txnip + control empty vector + control siRNA (siLacZ), Txnip + Gm15441-L + siLacZ, or Txnip + Gm15441-L + Gm15441-targeted siRNA (siGm15441), for 48 h, MOI=50; Quantitation of western blot is shown in the lower panel. Error bars are SEM, n=2, *P<0.05.

Figure 5. Gm15441 reduces endogenous Txnip protein levels in vitro. (A-B) Quantitative real-time PCR analysis (A) and western blot analysis (B) of mouse hepatocytes AML12 transduced with adenovirus expressing empty vector as control or Gm15441-L, for 24 h, MOI=50. Quantitation of western blot is shown in the right panel Error bars are SEM, n=3, *P<0.05; (C-D) Quantitative real-time PCR analysis (C) and western blot analysis (D) of mouse hepatocytes AML12 transfected with siLacZ as control and Gm15441-targeted siRNA (siGm15441), for 48 h. Quantitation of western blot is shown in the right panel; Error bars are SEM, n=2-3, *P<0.05.

Figure 6. Gm15441 suppresses Txnip protein expression through translational inhibition. (A) Western blot analysis of mouse primary hepatocytes transduced with adenovirus expressing YFP + empty vector, Txnip + empty vector, YFP + Gm15441-L, Txnip + Gm15441-L, or Txnip + Gm15441-S, for 24 h, MOI=50. Quantitation is shown in the right panel; (B) Western blot analysis of Hepal-6 hepatocytes transduced with adenovirus expressing YFP + empty vector, Txnip + empty vector, or Txnip + full length of Gm15441-L or one fragment of Gm15441-L (267-709 nt, 267-430 nt, 431-709 nt), for 24 h, MOI=50. Quantitation is shown in the right panel; (C) Representative fluoresces of RNA in situ hybridization staining of Gm15441 RNA (green) and Txnip RNA (red) in Hepal-6 hepatocytes transduced with adenovirus expressing empty vector, or Txnip + Gm15441-L for 24 h, MOI=50. All scale bars represent 5 μ m; (D) Quantitative real-time PCR analysis of pulled-down RNAs in Hepal-6 hepatocytes transduced with adenovirus expressing Txnip and Gm15441-L for 24 h, MOI=50, using non-specific oligo probe (control) and Gm15441 specific oligo probes; (E) Western blot analysis of Hepal-6 hepatocytes transduced with adenovirus expressing Txnip + empty vector, or Txnip + Gm15441-L for 24 h, MOI=50, followed by a 0, 20-min or, 40-min cycloheximide (CHX) treatment. Quantitation is shown in the right panel. Error bars are SEM, n=2. *P<0.05 (versus control), #P<0.05 (versus Txnip+Adv group), N.S., not significant.

Figure 7. Liver-specific Gm15441 overexpression decreases Txnip protein levels and regulates glucose and lipid metabolism in mice. (A) Quantitative real-time PCR analysis of liver tissues from

C57BL/6 wildtype mice injected with adenovirus expressing empty vector, Gm15441-L, or Gm15441-S, n=5; **(B)** Comparison of Δ Ct values between Gm15441 and Txnip (normalized to the housekeeping gene 18S) from quantitative real-time PCR analysis of liver tissues from C57BL/6 wildtype mice injected with Gm15441-L or Gm15441-S, n=5; **(C)** Western blot analysis of liver tissues from C57BL/6 wildtype mice injected with adenovirus expressing empty vector, Gm15441-L, or Gm15441-S. Quantitation is shown in the right panel. n=3; **(D)** 24-h fasting body weight of control, Gm15441-L or Gm15441 S overexpression mice. n=5; **(E)** Overnight fasting blood glucose levels of control, Gm15441-L or Gm15441 S overexpression mice. n=5; **(F-H)** 24-h fasting plasma ketone body β -hydroxybutyrate (BOH) **(F)**, plasma triglyceride (TG) **(G)**, and liver TG **(H)** levels of control, Gm15441-L or Gm15441 S overexpression mice, n=5. Error bars are SEM, *P<0.05.

Figure S1. Multiple binding sites of FOXO1 **(A)** and PPAR α **(B)** located on the promoter and/or the gene bodies of Gm15441 and Txnip. Data are retrieved from the Gene Transcription Regulation Database.

Figure S2. Function prediction of Txnip and Gm15441 in mouse liver. **(A-B)** Genome-wide correlation analysis of Txnip and Gm15441. Representative Gene Ontology (GO) terms of correlated mRNAs for Txnip **(A)** and Gm15441 **(B)** in the liver are listed.

Table S1. The primer list for real-time PCR analysis.

Supplementary Material. The complete list of mRNAs highly correlated with Txnip and Gm15441.

Structural Insights into the Bypass of the Major Deaminated Purines by Translesion Synthesis DNA Polymerase

Hunmin Jung, Michael A. Hawkins, and Seongmin Lee*

The Division of Chemical Biology and Medicinal Chemistry, College of Pharmacy,

The University of Texas at Austin, Austin, Texas 78712, USA

*To whom correspondence should be addressed. Tel: +1 512 471 1785; Fax: +1 512 471 4726

Email: SeongminLee@austin.utexas.edu

ABSTRACT

The exocyclic amines of nucleobases can undergo deamination by various DNA damaging agents such as reactive oxygen species, nitric oxide, and water. The deamination of guanine and adenine generates the promutagenic xanthine and hypoxanthine, respectively. The exocyclic amines of bases in DNA are hydrogen bond donors, while the carbonyl moiety generated by the base deamination acts as hydrogen bond acceptors, which can alter base pairing properties of the purines. Xanthine is known to base pair with both cytosine and thymine, while hypoxanthine predominantly pairs with cytosine to promote A to G mutations. Despite the known promutagenicity of the major deaminated purines, structures of DNA polymerase bypassing these lesions have not been reported. To gain insights into the deaminated-induced mutagenesis, we solved crystal structures of human DNA polymerase η (pol η) catalyzing across xanthine and hypoxanthine. In the catalytic site of pol η , the deaminated guanine (i.e., xanthine) forms three Watson-Crick-like hydrogen bonds with an incoming dCTP, indicating the O2-enol tautomer of xanthine involves in the base pairing. The formation of the enol tautomer appears to be promoted by the minor groove contact by Gln38 of pol η . When hypoxanthine is at the templating position, the deaminated adenine uses its O6-keto tautomer to form two Watson-Crick hydrogen bonds with an incoming dCTP, providing the structural basis for the high promutagenicity of hypoxanthine.

INTRODUCTION

The genomic DNA is under persistent threats by DNA damaging events such as oxidation, UV irradiation, alkylation, and deamination. The exocyclic amines of nucleobases are susceptible to deamination by reactive oxygen species, reactive halogen species, nitric oxide, and water among others. These deamination events give rise to a wide variety of DNA lesions, including uracil (from cytosine), thymine (from 5-methylcytosine), hypoxanthine (from adenine), and xanthine (from guanine) [1-3] (**Figure 1**). Deaminated

lesions are promutagenic as they change the donor-acceptor property of base pairing [4-6]. The deaminated purines in DNA are mainly recognized and removed by DNA glycosylases such as *E. coli* AlkA or Endo V in prokaryotes, and alkyladenine DNA glycosylase (AAG) or the homologs of Endo V in eukaryotes [1, 7-11] and the resulting abasic site is further processed by the downstream base excision repair (BER) enzymes.

Xanthine, as its monophosphate form (dXMP), is a key intermediate for *de novo* synthesis of dGMP, and its concentration (both dXMP and free xanthine) is closely related to the concentration of the entire guanine nucleotide pool [12, 13]. Xanthine, which can arise by deamination of guanine in DNA, is a highly promutagenic damage. DNA polymerase α in eukaryotes frequently misincorporates dTMP opposite xanthine. While mammalian pol β accurately bypasses xanthine [14], translesion synthesis DNA polymerases η (pol η) and κ (pol κ) preferentially incorporate dTTP opposite xanthine [15]. In mammalian cells, xanthine is a miscoding lesion that significantly promotes G to A transitions [16, 17].

Hypoxanthine, the deaminated adenine, preferentially base pairs with dCTP and causes A to G mutations in many organisms [18, 19]. The triphosphate form of hypoxanthine, dITP, is taken up readily by DNA polymerases [20]. Hypoxanthine is a significant block to replication by B-family Pfu DNA polymerase and *Sulfolobus solfataricus* Dpo1, while it is readily bypassed by Taq DNA polymerase and Y-family DNA polymerase PolY1, and *Sulfolobus solfataricus* Dpo4 [21, 22]. Human DNA polymerases α , η , and truncated κ ($\kappa\Delta C$) insert dCTP quite exclusively over dTTP [23]. In HEK293 and HCT116 human cell lines, hypoxanthine induces A to G mutations and deletion [24].

A wide variety of DNA lesions are known to be bypassed by translesion synthesis (TLS) DNA polymerases. Among TLS polymerases, human Y-family DNA polymerase η (pol η) has been given special attention due to its involvement in the bypass of cisplatin-GpG [25] and UV-induced cyclobutane pyrimidine dimers (CPDs) [26]. Pol η has been also shown to catalyze across 8-oxoguanine [27], 8-oxoadenine [28, 29], O6-methylguanine, oxaliplatin-GpG [30], and N7-methylguanine [31]. Pol η has unique structural features of large solvent access area and relatively rigid active site conformation [32], and is known to bypass small to medium sized lesions (e.g., cisplatin-GpG) [32-34] but the replication was blocked at bulky lesions such as Benzo[a]pyrene Diol Epoxide-Guanine Adducts (BPDE-dG) [35]. Gln38 and Arg61 of human pol η play critical role in the lesion bypass [29, 32, 36].

While there are many published studies on the promutagenicity of the deaminated purines, a structure of DNA polymerase bypassing those lesions has not been reported, significantly limiting our understanding of hypoxanthine/xanthine-induced mutagenesis. Herein, we present steady-state kinetic data of human pol η incorporating nucleotide opposite deaminated purine lesions, xanthine and hypoxanthine, along with two crystal structures of pol η complexed with a templating xanthine/hypoxanthine and incoming non-hydrolyzable dCTP analog. These crystal structures represent the first structure of DNA polymerase bypassing xanthine and hypoxanthine. Our studies revealed the base pairing properties of the major

deaminated purines in the active site of pol η , which provides valuable insights into the bypass of the deaminated lesions by DNA polymerases, especially translesion synthesis polymerases.

MATERIALS AND METHODS

Protein expression and purification. Pol η was expressed and purified from *E. coli* with minor modifications of the method described previously [32]. Briefly, Pol η was overexpressed in *E. coli* BL21(DE3) cells, and cultures were grown in Luria-Bertani medium at 37 °C until reaching the OD₆₀₀ of 0.7, and the cells were induced by adding 0.2 mM isopropyl β -D- α -thiogalactopyranoside. After incubating for 18 hours at 20 °C, the pelleted cells (6,000 RPM for 30 min) were resuspended in Ni-NTA column binding buffer A (50 mM sodium phosphate, pH 7.8, 500 mM NaCl and 10% glycerol) supplemented with 1 mg/ml lysozyme, 0.25% NP-40, 0.25% Triton X-100, and 0.25 mM phenylmethylsulfonyl fluoride (PMSF). After sonication for 90 seconds, the lysate was centrifuged at 15,000 g at 4 °C for 20 min. The supernatant was then filtered through 0.22 μ m filter and further purified through Ni-NTA column (GE Healthcare). The elution fractions were pooled and further purified using the Heparin HiTrap column (GE Healthcare) followed by Superdex-75 size exclusion chromatography (GE Healthcare). The purity of the final product was confirmed by SDS-PAGE gel. The purified protein was concentrated, flash-frozen in liquid nitrogen, and stored at -80 °C for the future use.

Protein-DNA crystallization and structure determination. Xanthine (XT)- and hypoxanthine (HX)-containing DNA were custom synthesized by Midland Certified Reagent Co. (Midland, TX). The primer (5'-AGCGTCAT-3') was purchased from Integrated DNA Technologies (Coralville, IA). XT/HX-containing 12-mer template (5'-CAT[XT/HX]CTCACACT-3') was annealed with complementary 8-mer primer (5'-AGTGTGAG-3') in hybridization buffer (10 mM Tris-HCl pH 7.5, 30 mM NaCl, 1 mM EDTA) by heating for 5 min at 90°C followed by slow cooling to room temperature. The annealed lesion-containing DNA was incubated with ~9 mg/ml pol η to form protein-DNA binary complexes with 1.2:1 molar ratio. Subsequently, a 10-fold molar excess of nonhydrolyzable dCMPNPP (Jena Bioscience) was added to the binary complex. The ternary pol η -DNA complex co-crystals with nonhydrolyzable dCMPNPP (dCTP* hereafter) paired with templating XT/HX were grown in a buffer solution containing 100 mM MES pH 6.5, 14–23% PEG2000 MME, and 5 mM magnesium chloride. Crystals were cryoprotected in mother liquor supplemented with 20% glycerol and were flash-frozen in liquid nitrogen. Diffraction data were collected at 100 K at the beamline 23-ID-D at the Advanced Photon Source, Argonne National Laboratory. All diffraction data were processed using HKL 2000 [37], and the structures were solved by molecular replacement using Molrep [38]. Pol η structure with an undamaged DNA (PDB ID 4O3N) was used as a search model. The model was built using COOT [39] and refined using PHENIX [40]. MolProbity was

used to make Ramachandran plots [41]. All the crystallographic figures were generated by using Chimera [42].

Steady-state kinetics of single nucleotide incorporation opposite templating xanthine/hypoxanthine by pol η . Steady-state kinetic parameters for nucleotide insertion opposite XT/HX by pol η were measured as described previously [29,32]. Briefly, The oligonucleotides for kinetic assays (primer, 5'-FAM/GGGGG CTCGTAAGGATTC-3' and template, 5'-CCGACT[XT/HX]GAATCCTTACGAGCCCC-3') were synthesized by Midland Certified Reagent company (Midland, TX) and Integrated DNA Technologies (Coralville, IA). To prepare DNA substrate containing XT/HX, each oligonucleotide was annealed in hybridization buffer (10 mM Tris-HCl pH 7.5, 30 mM NaCl, 1 mM EDTA) by heating for 5 min at 90°C followed by slow cooling to room temperature. Enzyme activities were determined using the reaction mixture containing 40 mM Tris-HCl pH 7.5, 60 mM KCl, 10 mM dithiothreitol, 250 μ g/ml bovine serum albumin, 2.5 % glycerol, 5 mM MgCl₂, 80 nM primer/template DNA, and the different concentration of incoming dNTP. To prevent end-product inhibition and substrate depletion from interfere with accurate velocity measurement, the enzyme concentrations and reaction-time intervals were adjusted for every experiment (less than 20% insertion product formed). The reactions were initiated by the addition of the enzyme and stopped with a gel-loading buffer (95% formamide with 20 mM EDTA, 45 mM Tris-borate, 0.1% bromophenol blue, 0.1% xylene cyanol). The quenched samples were separated on 20% denaturing polyacrylamide gels. The gels were analyzed using ImageQuant (GE Healthcare) to quantify product formation. The k_{cat} and K_m were determined by fitting reaction rate over dNTP concentrations to Michaelis-Menten equation. Each experiment was repeated three times to measure the average of the kinetic results. The catalytic efficiency of nucleotide insertion was calculated as k_{cat}/K_m .

RESULTS AND DISCUSSION

Pol η -catalyzed bypass of xanthine and hypoxanthine is promutagenic. To evaluate the mutagenic potential of xanthine in DNA, we determined kinetic parameters of pol η incorporating a nucleotide (dCTP or dTTP) opposite templating dA, dG, and XT (**Table 1** and **Figure 2**). Pol η inserted dCTP/dTTP opposite dG with the fidelity of ~ 100 ($45.6 \times 10^{-3} \text{s}^{-1} \mu\text{M}^{-1}$ vs. $0.47 \times 10^{-3} \text{s}^{-1} \mu\text{M}^{-1}$). In the presence of templating XT, the enzyme inserted dCTP/dTTP with the fidelity of ~ 3 ($11.5 \times 10^{-3} \text{s}^{-1} \mu\text{M}^{-1}$ vs. $3.99 \times 10^{-3} \text{s}^{-1} \mu\text{M}^{-1}$), resulting in more than ~ 30 -fold reduction in the replication fidelity. The presence of XT in pol η catalytic site decreased the efficiency of correct insertion ~ 4 -fold and increased the efficiency of incorrect insertion ~ 9 -fold. More specifically, for the correct insertion, substituting dG for XT increased K_m ~ 4.5 -fold (2.7 vs. 10.7 μM), while it negligibly changed k_{cat} ($120.6 \times 10^{-3} \text{s}^{-1}$ vs. $123.3 \times 10^{-3} \text{s}^{-1}$). Furthermore, the presence of templating XT facilitated misincorporation of dTTP by pol η . The catalytic efficiency (k_{cat}/K_m) of the XT:dTTP

insertion was ~8-fold ($0.47 \times 10^{-3} \text{s}^{-1} \mu\text{M}^{-1}$ vs. $3.99 \times 10^{-3} \text{s}^{-1} \mu\text{M}^{-1}$) greater than that for the dG:dTTP insertion, which is primarily caused by decreased K_m (159.3 vs. 20.6 μM). This highlights the deamination product of guanine at pol η catalytic active site promotes mutagenic replication (dTTP over dCTP). Comparison of our kinetic data with published results show that the replication fidelity of XT bypass by pol η is ~24-fold lower (3 vs. 73) than that by human pol β and 6-fold (3 vs. 0.5) greater than that by human pol κ (**Table 1**).

To assess the mutagenic potential of hypoxanthine, we determined kinetic parameters of pol η incorporating a nucleotide (dCTP or dTTP) opposite templating hypoxanthine (**Table 1** and **Figure 2**). Our kinetic study showed the translesion synthesis of hypoxanthine (HX) by pol η promotes A to G mutations by increasing the efficiency (k_{cat}/K_m) of dCTP insertion and decreasing the efficiency of dTTP insertion opposite HX lesion (**Table 1** and **Figure 2**). While pol η inserted dCTP/dTTP opposite the control dA with a fidelity ~90 ($17.0 \times 10^{-3} \text{s}^{-1} \mu\text{M}^{-1}$ vs. $0.19 \times 10^{-3} \text{s}^{-1} \mu\text{M}^{-1}$), it inserted dCTP/dTTP opposite HX with a fidelity of 0.014 ($0.54 \times 10^{-3} \text{s}^{-1} \mu\text{M}^{-1}$ vs. $37.4 \times 10^{-3} \text{s}^{-1} \mu\text{M}^{-1}$), thereby resulting in a ~6,200-fold reduction in the replication fidelity. The catalytic efficiency (k_{cat}/K_m) of the HX:dCTP insertion was ~200-fold ($37.4 \times 10^{-3} \text{s}^{-1} \mu\text{M}^{-1}$ vs. $0.19 \times 10^{-3} \text{s}^{-1} \mu\text{M}^{-1}$) greater than that of the dA:dCTP insertion, highlighting the deamination product of adenine at pol η catalytic site drastically facilitates mutagenic replication (dCTP/dTTP).

Xanthine uses its enol tautomer to form a Watson-Crick-like base pair with dCTP in pol η active site.

Our kinetic studies show that pol η efficiently incorporates dCTP opposite templating XT, which promotes accurate replication. To gain structural insight into the correct nucleotide insertion, we solved a ternary complex structure of pol η incorporating a nonhydrolyzable dCMPNPP (dCTP* hereafter) opposite templating XT in the presence of Mg^{2+} cofactors. The nonhydrolyzable nucleotide dCTP* was used because it is isosteric to dCTP, and its coordination to the active-site metal ions is essentially identical to that of dCTP [43]. The pol η -XT:dCTP* ternary complex was crystallized in $P6_1$ space group with the cell dimension of $a = 98.93 \text{ \AA}$, $b = 98.93 \text{ \AA}$, $c = 81.55 \text{ \AA}$, $\alpha = 90.00^\circ$, $\beta = 90.00^\circ$, and $\gamma = 120.00^\circ$. The pol η -XT:dCTP* ternary structure was refined to a resolution of 2.35 \AA with $R_{\text{work}} = 18.9 \%$ and $R_{\text{free}} = 23.8 \%$ (**Table 2**).

The pol η -XT:dCTP* ternary complex structure provides the structural basis for correct insertion opposite XT by the enzyme (**Figure 3**). This structure displays the conserved secondary structures and the four characteristic domains (thumb, palm, finger, and little finger) of Y-family DNA polymerases (**Figure 3A**). The XT:dCTP* base pair is well accommodated within the enzyme's catalytic site (**Figure 3B**) with strong electron density around XT and the incoming dCTP*, indicating the base pair is well accommodated in the catalytic site of pol η . The primer terminus 3'-OH is coordinated to the A-site magnesium ion and is about 3.4 \AA away from the P_α of dCTP* (**Figure 3C**), thereby being optimally positioned for in-line nucleophilic attack on the P_α of the dCTP*. The incoming dCTP* and primer terminus dT at the N-1 position are

favorably positioned for stacking interaction (**Figure 3C**). The templating XT adopts an *anti*-conformation and forms a Watson-Crick-like base pair with dCTP* with the inter-base hydrogen bonding distances of 3.2 Å, 2.8 Å, and 2.7 Å (**Figure 3D**). The geometry of XT:dCTP* base pair displays the λ angles of 61.9° (XT) and 55.0° (dCTP*) and the C1'-C1' distance of 10.4 Å (**Figure 3D**), which are very similar to those of correct undamaged base pairs. The O2 of XT is 2.7 Å away from the O2 of dCTP*, indicating the formation of a strong hydrogen bond between the O2 atoms (**Figure 3E**). This, in turn, indicates that the O2-enol tautomer of XT engages in hydrogen bonding interaction with the O2 of the incoming dCTP. The enol tautomerization of the O2 of XT nucleotide may be promoted the minor groove contact by Gln38, which is hydrogen bonded to the N3 of the templating XT. Overall, the pol η -XT:dCTP* ternary complex structure with three Watson-Crick-like hydrogen bonds and two metal ions is consistent with the efficient dCTP insertion opposite XT by the enzyme.

Hypoxanthine forms a Watson-Crick base pair with incoming dCTP in the active site of pol η . Our kinetic studies show that pol η efficiently catalyzes the insertion of dCTP opposite HX, which can promote A to G transition mutations. To gain structural insight into this incorrect bypass, we solved a ternary complex structure of pol η incorporating a nonhydrolyzable dCMPNPP (dCTP* hereafter) opposite templating HX in the presence Mg²⁺ ions. The pol η -HX:dCTP* ternary complex was crystallized in *P*6₁ space group with the cell dimension of a = 98.54 Å, b = 98.54 Å, c = 81.66 Å, α = 90.00°, β = 90.00°, and γ = 120.00°. The pol η -HX:dCTP* ternary structure was refined to a resolution of 1.97 Å with R_{work} = 17.1 % and R_{free} = 20.8 % (**Table 2**).

The pol η -HX:dCTP* ternary complex structure provides the structural basis for the promutagenic replication past HX by the enzyme (**Figure 4**). As observed in the pol η -XT:dCTP* structure, the pol η -HX:dCTP* structure displays the conserved four characteristic domains of Y-family DNA polymerases (**Figure 4A**). The incorrect HX:dCTP* base pair is well accommodated within the enzyme's catalytic site (**Figure 4B**) with strong electron density around HX and dCTP*. Both catalytic (A-site) and nucleotide-binding (B-site) metal ions are present in the catalytic site. The primer terminus 3'-OH is coordinated to the A-site magnesium ion and is about 3.1 Å away from the P $_{\alpha}$ of the incoming dCTP* (**Figure 4C**), thereby being poised for in-line nucleophilic attack on dCTP*. The templating HX is in an *anti*-conformation and forms a Watson-Crick base pair with dCTP* with the inter-base hydrogen bonding distances of 2.8 Å (between N1 of HX and N3 of dCTP*) and 3.3 Å (between O6 of HX and N4 of dCTP*). The base-pair geometry of HX:dCTP* base pair displays the λ angles of 59.3° (HX) and 60.2° (dCTP*) and the C1'-C1' distance of 10.5 Å (**Figure 4D**), which are essentially identical to those of an undamaged correct base pair. The Watson-Crick HX:dCTP* base pair indicates the O6-keto tautomer of HX involves in base pairing with

dCTP*. Overall, the pol η -HX:dCTP* ternary complex structure with Watson-Crick base pair and two magnesium ions is consistent with the efficient insertion of dCTP opposite HX by the pol η .

The comparison of the pol η -HX:dCTP* and pol η -XT:dCTP* structures reveals conformational differences are mainly confined to the 5' side of the templating purines and the primer terminus base (**Figure 5**). Both HX:dCTP* and XT:dCTP* base pairs are well accommodated in the catalytic site without significant change of protein conformation. The conformations of the incoming dCTP in both structures are essentially identical. On the other hand, the conformation of the 5' phosphate of the templating deaminated purines differ greatly (**Figure 5C and 5D**). While the dT at the N+1 position engages in stacking interaction with templating XT, it does not stack with the templating HX (**Figure 5A**). These conformational differences could contribute to the difference in the catalytic efficiency of dCTP insertion opposite HX and XT.

The deaminated purines and dG adopt similar conformation in the active site of pol η . The pol η -xanthine:dCTP* and pol η -hypoxanthine:dCTP* structures are very similar to the published pol η -dG:dCTP* structure (PDB ID: 4O3N, **Figure 6**) [44]. In the active site of pol η , dG and dCTP form Watson-Crick base pair (**Figure 6A**). The templating dT at the N+1 position engages in π -stacking interaction with the templating dG at the N position. The superposition of the pol η -XT:dCTP* structure with the pol η -dG:dCTP* structure reveals that the conformations of templating XT/dG, incoming dCTP*, primer strand, and downstream template bases are essentially identical (**Figure 6B**). Also, the guanidine moieties of Arg61 in the two structures stabilize the incoming dCTP* via π -stacking interaction (**Figure 6C**). In addition, Arg55, which interacts with the γ -phosphate of incoming nucleotide, adopts the same conformation in the two structures (not shown). The only significant deviation is found at the templating bases at the N+1 and N+2 positions. The conformation of dT (N+1) in the pol η -XT:dCTP* complex is not in the optimal position compared with that in pol η -dG:dCTP* structure. In addition, dA at the N+2 position in the pol η -XT:dCTP* complex shifts ~ 4 Å relative to the position in the pol η -dG:dCTP* complex. These large conformational differences at the N+1 and N+2 positions may contribute to a 4-fold reduction in the catalytic efficiency (Table 1).

The superposition of the pol η -HX:dCTP* structure with the pol η -dG:dCTP* structure (RMSD: 0.220 Å) shows that a significant conformational change is confined to the templating dT at the N+1 position. The templating dT(N+1) in the pol η -HX:dCTP* structure rotates $\sim 90^\circ$ toward the major groove relative to the position in the pol η -dG:dCTP* structure, thereby not engaging in stacking interaction with the templating HX at the N position. On the other hand, the templating dA at the N+2 position of the pol η -HX:dCTP* overlays very well with the corresponding dA of the pol η -dG:dCTP* structure. In addition, the templating base at the N position, incoming dCTP*, primer strand, and downstream template bases in the two structures

do not exhibit a significant conformational deviation (**Figure 6D**). The guanidine moieties of Arg61 in the two structures stabilize incoming dCTP via stacking interaction (**Figure 6E**). In addition, the orientation and conformation of Arg55 in the two structures are essentially the same (not shown). Interestingly, despite the disruption of the stacking interaction between dT(N+1) and HX(N), the catalytic efficiency of the HX:dCTP insertion is only slightly lower than that of the dG:dCTP insertion (37.4 vs. 45.6), indicating the stacking interaction between the templating bases at the N and N+1 positions may not significantly contribute to the catalytic efficiency of pol η .

Promutagenic insertion opposite XT and HX by DNA polymerases. Several kinetic reports on xanthine [15, 45] and hypoxanthine [21, 23, 45] (**Table 1**) have highlighted the mutagenic potential of the major deaminated purines. In particular, bypass properties of XT and HX by human DNA polymerases α (pol α), β (pol β), and κ (pol κ) have been characterized [15, 23]. For example, X-family pol β preferentially incorporates the correct nucleotide opposite XT with the fidelity (dCTP/dTTP) of ~ 70 (**Table 1**). On the other hand, Y-family pol κ frequently incorporates dTTP with the fidelity (dCTP/dTTP) of ~ 0.5 (**Table 1**). In the case of Y-family pol η (our study), dCTP incorporation opposite XT is only ~ 3 -fold more efficient than dTTP insertion, highlighting the bypass fidelity of XT is greatly influenced by the microenvironment of DNA polymerase active site. At physiological pH, xanthine exists as an almost equal mixture of neutral and monoanionic (enolate) species (**Figure 7A**) [46], the latter of which can take on an enol tautomer upon protonation. Our pol η -XT:dCTP* crystal structure displays hydrogen bonding interactions between O2 of XT and O2 of dCTP* with the distance of 2.7 Å, which indicates that O2 of XT is in the enol tautomeric state (**Figure 7B**). This observation is consistent with the higher insertion efficiency of dCTP over dTTP opposite XT in the active site of pol η .

Varying degrees of minor groove contacts to the N3 of XT by DNA polymerases may contribute to the fidelity of XT bypass. Pol κ does not engage in minor groove interaction at the replicating base pair site. In the case pol η , the enzyme typically uses Gln38 to form hydrogen bonds with the minor groove edge and O4' of a templating base during correct insertion, while it does not interact with the minor groove edge of incoming nucleotide. The minor groove contact to a templating XT could promote the formation of the enol tautomer of the base, which in turn can modulate the efficiency and fidelity of XT bypass.

While the bypass fidelity of XT varies significantly among DNA polymerases, that of HX varies to a much lesser extent. This would be because HX exists predominantly as the O6-keto tautomeric species in base pairing (**Figure 7C**), whereas XT exists as a $\sim 1:1$ mixture of the O2-keto and O2-enol tautomers and its tautomeric ratio can be significantly influenced by the microenvironment of DNA polymerases. In all three human DNA polymerases, pol α , pol η , and pol $\kappa\Delta C$, dCTP is preferentially inserted opposite HX with the fidelity (dTTP/dCTP) ranging from 0.008 (hpol $\kappa\Delta C$) to 0.038 (hpol η), which can facilitate A to G

mutations. The crystal structure of the pol η -HX:dCTP* complex shows that hypoxanthine behaves much like guanine and uses its O6-keto tautomer when paired with dCTP.

CONCLUSIONS

The crystal structure of the pol η -HX:dCTP* complex shows the formation of two hydrogen bonds with a Watson-Crick geometry, which is consistent with the preferential insertion of dCTP opposite HX. The crystal structure of the pol η -XT:dCTP* complex reveals that XT forms three Watson-Crick-like hydrogen bonds with incoming dCTP*, indicating the enol tautomeric species of XT involves in the correct nucleotide insertion. The enol tautomerization of the O2 of XT appears to be promoted by the minor groove contact by Gln38 of pol η . These structures, which represent the first structures of TLS polymerase bypassing deaminated purines, provide structural insights into the mutagenic potential of the major deaminated purines.

Data Availability Statement

The atomic coordinates of pol η -DNA complexes have been deposited in the Protein Data Bank with the following accession codes: pol η -XT:dCTP (PDB Code: 6WK6) and pol η -HX:dCTP (PDB Code: 6MQ8).

ACKNOWLEDGEMENTS

We are grateful to Dr. Arthur Monzingo for technical assistance. Instrumentation and technical assistance for this work were provided by the Macromolecular Crystallography Facility, with financial support from the College of Natural Sciences, the Office of the Executive Vice President and Provost, and the Institute for Cellular and Molecular Biology at the University of Texas at Austin. Portions of this research were conducted at the Advanced Photon Source with the support of GM/CA. GM/CA@APS has been funded in whole or in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. The Elger 16M detector was funded by an NIH-Office of Research Infrastructure Programs, High-End Instrumentation Grant (1S10OD012289-01A1).

Funding:

The work was supported in part by the National Institutes of Health [ES-26676].

Competing Interests: The authors declare no conflict of interest.

REFERENCES

- 1 Kow, Y. W. (2002) Repair of deaminated bases in DNA. *Free Radic Biol Med.* **33**, 886-893
- 2 Caulfield, J. L., Wishnok, J. S. and Tannenbaum, S. R. (1998) Nitric oxide-induced deamination of cytosine and guanine in deoxynucleosides and oligonucleotides. *J Biol Chem.* **273**, 12689-12695
- 3 Wink, D. A., Kasprzak, K. S., Maragos, C. M., Elespuru, R. K., Misra, M., Dunams, T. M., Cebula, T. A., Koch, W. H., Andrews, A. W. and Allen, J. S. (1991) DNA deaminating ability and genotoxicity of nitric oxide and its progenitors. *Science.* **254**, 1001-1003
- 4 Pang, B., McFaline, J. L., Burgis, N. E., Dong, M., Taghizadeh, K., Sullivan, M. R., Elmquist, C. E., Cunningham, R. P. and Dedon, P. C. (2012) Defects in purine nucleotide metabolism lead to substantial incorporation of xanthine and hypoxanthine into DNA and RNA. *Proc Natl Acad Sci U S A.* **109**, 2319-2324
- 5 Kamiya, H. (2003) Mutagenic potentials of damaged nucleic acids produced by reactive oxygen/nitrogen species: approaches using synthetic oligonucleotides and nucleotides: survey and summary. *Nucleic Acids Res.* **31**, 517-531
- 6 Lindahl, T. and Nyberg, B. (1974) Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry.* **13**, 3405-3410
- 7 Terato, H., Masaoka, A., Asagoshi, K., Honsho, A., Ohyama, Y., Suzuki, T., Yamada, M., Makino, K., Yamamoto, K. and Ide, H. (2002) Novel repair activities of AlkA (3-methyladenine DNA glycosylase II) and endonuclease VIII for xanthine and oxanine, guanine lesions induced by nitric oxide and nitrous acid. *Nucleic Acids Res.* **30**, 4975-4984
- 8 Saparbaev, M. and Laval, J. (1994) Excision of hypoxanthine from DNA containing dIMP residues by the *Escherichia coli*, yeast, rat, and human alkylpurine DNA glycosylases. *Proc Natl Acad Sci U S A.* **91**, 5873-5877
- 9 Wuenschell, G. E., O'Connor, T. R. and Termini, J. (2003) Stability, miscoding potential, and repair of 2'-deoxyxanthosine in DNA: implications for nitric oxide-induced mutagenesis. *Biochemistry.* **42**, 3608-3616
- 10 Vallur, A. C., Maher, R. L. and Bloom, L. B. (2005) The efficiency of hypoxanthine excision by alkyladenine DNA glycosylase is altered by changes in nearest neighbor bases. *DNA Repair (Amst).* **4**, 1088-1098
- 11 Yao, M., Hatahet, Z., Melamede, R. J. and Kow, Y. W. (1994) Purification and characterization of a novel deoxyinosine-specific enzyme, deoxyinosine 3' endonuclease, from *Escherichia coli*. *J Biol Chem.* **269**, 16260-16268
- 12 Hedstrom, L. (2009) IMP dehydrogenase: structure, mechanism, and inhibition. *Chem Rev.* **109**, 2903-2928
- 13 Welin, M., Lehtiö, L., Johansson, A., Flodin, S., Nyman, T., Trésaugues, L., Hammarström, M., Gräslund, S. and Nordlund, P. (2013) Substrate specificity and oligomerization of human GMP synthetase. *J Mol Biol.* **425**, 4323-4333
- 14 Kamiya, H., Sakaguchi, T., Murata, N., Fujimuro, M., Miura, H., Ishikawa, H., Shimizu, M., Inoue, H., Nishimura, S. and Matsukage, A. (1992) In vitro replication study of modified bases in ras sequences. *Chem Pharm Bull (Tokyo).* **40**, 2792-2795

- 15 Yasui, M., Suzuki, N., Miller, H., Matsuda, T., Matsui, S. and Shibutani, S. (2004) Translesion synthesis past 2'-deoxyxanthosine, a nitric oxide-derived DNA adduct, by mammalian DNA polymerases. *J Mol Biol.* **344**, 665-674
- 16 Kamiya, H., Shimizu, M., Suzuki, M., Inoue, H. and Ohtsuka, E. (1992) Mutation induced by deoxyxanthosine in codon 12 of a synthetic c-Ha-ras gene. *Nucleosides Nucleotides.* **11**, 247-260
- 17 Eritja, R., Horowitz, D. M., Walker, P. A., Ziehler-Martin, J. P., Boosalis, M. S., Goodman, M. F., Itakura, K. and Kaplan, B. E. (1986) Synthesis and properties of oligonucleotides containing 2'-deoxynebularine and 2'-deoxyxanthosine. *Nucleic Acids Res.* **14**, 8135-8153
- 18 Nordmann, P. L., Makris, J. C. and Reznikoff, W. S. (1988) Inosine induced mutations. *Mol Gen Genet.* **214**, 62-67
- 19 Hill-Perkins, M., Jones, M. D. and Karran, P. (1986) Site-specific mutagenesis in vivo by single methylated or deaminated purine bases. *Mutat Res.* **162**, 153-163
- 20 Piccirilli, J. A., Krauch, T., Moroney, S. E. and Benner, S. A. (1990) Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature.* **343**, 33-37
- 21 Grúz, P., Shimizu, M., Pisani, F. M., De Felice, M., Kanke, Y. and Nohmi, T. (2003) Processing of DNA lesions by archaeal DNA polymerases from *Sulfolobus solfataricus*. *Nucleic Acids Res.* **31**, 4024-4030
- 22 Choi, J. Y., Eoff, R. L., Pence, M. G., Wang, J., Martin, M. V., Kim, E. J., Folkmann, L. M. and Guengerich, F. P. (2011) Roles of the four DNA polymerases of the crenarchaeon *Sulfolobus solfataricus* and accessory proteins in DNA replication. *J Biol Chem.* **286**, 31180-31193
- 23 Yasui, M., Suenaga, E., Koyama, N., Masutani, C., Hanaoka, F., Gruz, P., Shibutani, S., Nohmi, T., Hayashi, M. and Honma, M. (2008) Miscoding properties of 2'-deoxyinosine, a nitric oxide-derived DNA Adduct, during translesion synthesis catalyzed by human DNA polymerases. *J Mol Biol.* **377**, 1015-1023
- 24 DeVito, S., Woodrick, J., Song, L. and Roy, R. (2017) Mutagenic potential of hypoxanthine in live human cells. *Mutat Res.* **803-805**, 9-16
- 25 Albertella, M. R., Green, C. M., Lehmann, A. R. and O'Connor, M. J. (2005) A role for polymerase eta in the cellular tolerance to cisplatin-induced damage. *Cancer Res.* **65**, 9799-9806
- 26 Yoon, J. H., Prakash, L. and Prakash, S. (2009) Highly error-free role of DNA polymerase eta in the replicative bypass of UV-induced pyrimidine dimers in mouse and human cells. *Proc Natl Acad Sci U S A.* **106**, 18219-18224
- 27 Haracska, L., Yu, S. L., Johnson, R. E., Prakash, L. and Prakash, S. (2000) Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase eta. *Nat Genet.* **25**, 458-461
- 28 Koag, M. C., Jung, H. and Lee, S. (2019) Mutagenic replication of the major oxidative adenine lesion 7,8-dihydro-8-oxoadenine by human DNA polymerases. *J Am Chem Soc*
- 29 Koag, M. C., Jung, H. and Lee, S. (2020) Mutagenesis mechanism of the major oxidative adenine lesion 7,8-dihydro-8-oxoadenine. *Nucleic Acids Res.* **48**, 5119-5134
- 30 Vaisman, A., Masutani, C., Hanaoka, F. and Chaney, S. G. (2000) Efficient translesion replication past oxaliplatin and cisplatin GpG adducts by human DNA polymerase eta. *Biochemistry.* **39**, 4575-4580

- 31 Koag, M. C., Jung, H., Kou, Y. and Lee, S. (2019) Bypass of the Major Alkylative DNA Lesion by Human DNA Polymerase η . *Molecules*. **24**
- 32 Biertümpfel, C., Zhao, Y., Kondo, Y., Ramón-Maiques, S., Gregory, M., Lee, J. Y., Masutani, C., Lehmann, A. R., Hanaoka, F. and Yang, W. (2010) Structure and mechanism of human DNA polymerase ϵ . *Nature*. **465**, 1044-1048
- 33 Silverstein, T. D., Johnson, R. E., Jain, R., Prakash, L., Prakash, S. and Aggarwal, A. K. (2010) Structural basis for the suppression of skin cancers by DNA polymerase ϵ . *Nature*. **465**, 1039-1043
- 34 Zhao, Y., Biertümpfel, C., Gregory, M. T., Hua, Y. J., Hanaoka, F. and Yang, W. (2012) Structural basis of human DNA polymerase η -mediated chemoresistance to cisplatin. *Proc Natl Acad Sci U S A*. **109**, 7269-7274
- 35 Masutani, C., Kusumoto, R., Iwai, S. and Hanaoka, F. (2000) Mechanisms of accurate translesion synthesis by human DNA polymerase ϵ . *EMBO J*. **19**, 3100-3109
- 36 Katafuchi, A., Sassa, A., Niimi, N., Grúz, P., Fujimoto, H., Masutani, C., Hanaoka, F., Ohta, T. and Nohmi, T. (2010) Critical amino acids in human DNA polymerases ϵ and κ involved in erroneous incorporation of oxidized nucleotides. *Nucleic Acids Res*. **38**, 859-867
- 37 Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. **276**, 307-326
- 38 Vagin, A. and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr*. **66**, 22-25
- 39 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*. **60**, 2126-2132
- 40 Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. and Adams, P. D. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol*. **75**, 861-877
- 41 Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. **66**, 12-21
- 42 Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*. **25**, 1605-1612
- 43 Batra, V. K., Beard, W. A., Shock, D. D., Krahn, J. M., Pedersen, L. C. and Wilson, S. H. (2006) Magnesium-induced assembly of a complete DNA polymerase catalytic complex. *Structure*. **14**, 757-766
- 44 Patra, A., Nagy, L. D., Zhang, Q., Su, Y., Muller, L., Guengerich, F. P., Egli, M. (2014) Kinetic, structure, and mechanism of 8-oxo-7,8-dihydro-2'-deoxyguanosine bypass by human DNA polymerase ϵ . *J Biol Chem*. **289**, 16867-16882.
- 45 Weiss, B. (2001) Endonuclease V of *Escherichia coli* prevents mutations from nitrosative deamination during nitrate/nitrite respiration. *Mutat Res*. **461**, 301-309
- 46 Kulikowska, E., Kierdaszuk, B. and Shugar, D. (2004) Xanthine, xanthosine and its nucleotides: solution structures of neutral and ionic forms, and relevance to substrate

properties in various enzyme systems and metabolic pathways. *Acta Biochim Pol.* **51**, 493-531

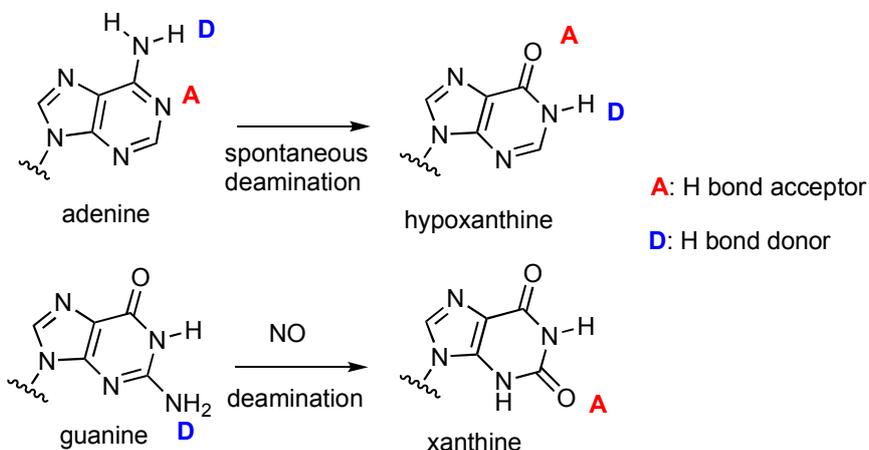


Figure 1. Formation of xanthine and hypoxanthine by deamination of purines. Deamination of purines alters hydrogen bonding properties of the bases. “A” and “B” denote hydrogen bond acceptor and donor, respectively.

Table 1. Kinetic parameters for nucleotide incorporation opposite hypoxanthine/xanthine by pol η .

template:dNTP	K_m (μM)	k_{cat} (10^{-3}s^{-1})	k_{cat}/K_m ($10^{-3}\text{s}^{-1}\mu\text{M}^{-1}$)	f^a	replication fidelity
polη					
dG:dCTP	2.7 ± 0.3	120.6 ± 6.0	45.6	1	
dG:dTTP	159.3 ± 2.7	74.8 ± 0.9	0.47	0.010	100
dA:dTTP	5.35 ± 0.2	90.9 ± 5.8	17.0	1	
dA:dCTP	80.3 ± 3.2	15.2 ± 2.5	0.19	0.011	90
XT:dCTP (correct insertion)	10.7 ± 0.9	123.3 ± 3.6	11.5	1	
XT:dTTP (incorrect insertion)	20.6 ± 0.9	82.2 ± 4.2	3.99	0.34	3
HX:dTTP (correct insertion)	21.9 ± 1.4	11.7 ± 0.2	0.54	1	
HX:dCTP (incorrect insertion)	4.6 ± 0.4	170.5 ± 4.1	37.4	69	0.014
XT:dCTP ^b (pol β)	10.6 ± 2.4	346.7 ± 76.7	32.7	1	
				0.014	72
XT:dTTP ^b (pol β)	155 ± 63	69.3 ± 1.5	0.45		
XT:dCTP ^b (pol κ)	4.53 ± 1.5	148.3 ± 28.2	32.7	1	
XT:dTTP ^b (pol κ)	2.25 ± 0.42	152.5 ± 15.3	67.8	2.07	0.5
HX:dTTP ^c (pol α)	9.74 ± 1.60	1.83 ± 1.17	0.19	1	
HX:dCTP ^c (pol α)	0.73 ± 0.26	7.83 ± 0.5	10.7	56	0.002
HX:dTTP ^c (pol $\kappa\Delta\text{C}$)	23.5 ± 6.9	25.0 ± 1.8	1.06	1	
HX:dCTP ^c (pol $\kappa\Delta\text{C}$)	1.36 ± 0.40	171.7 ± 7.3	126.3	120	0.008

^aRelative efficiency: $(k_{cat}/K_m)_{[single\ nucleotide\ insertion]}/(k_{cat}/K_m)_{[correct\ insertion]}$

^bReference [15] ^cReference [23]

5' -CCGACTXGAATCCTTACGAGCCCC-3' (template, X=XT or HX)
3' -CTTAGGAATGCTCGGGG-FAM-5'

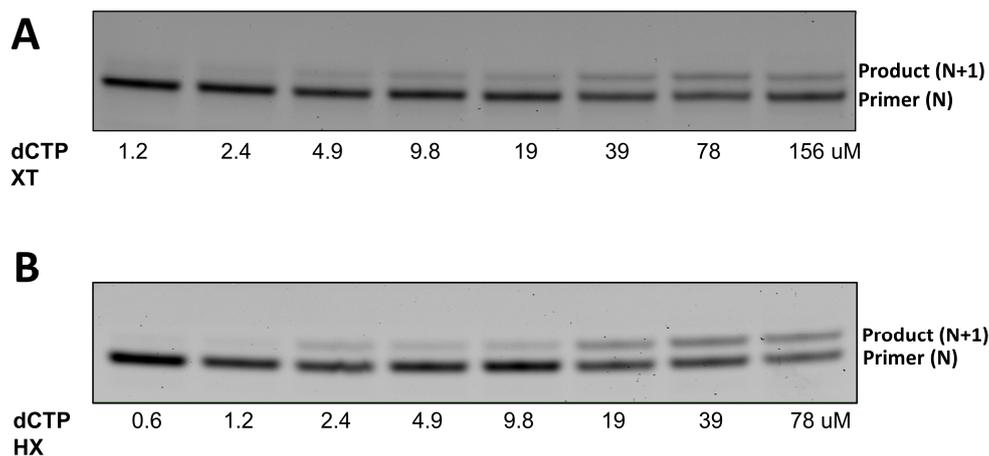


Figure 2. Representative denaturing polyacrylamide gel electrophoresis of polη incorporating dCTP opposite xanthine and hypoxanthine. Incorporation of dCTP opposite xanthine (A) and hypoxanthine (B) by polη. An annealed DNA of 5'-FAM-labeled primer and XT- or HX-containing template was mixed with varying concentrations of dCTP, and the reactions were initiated by the addition of polη. All the reactions were conducted at 37 °C, and the quenched reaction samples were separated on 20% denaturing polyacrylamide gels.

Table 2. Data Collection and Refinement Statistics.

PDB CODE	XT :dCTP* (6WK6)	HX :dCTP* (6MQ8)
Data Collection		
space group	$P6_1$	$P6_1$
Cell Constants		
a (Å)	98.930	98.539
b	98.930	98.539
c	81.552	81.664
α (°)	90.00	90.00
β	90.00	90.00
γ	120.00	120.00
resolution (Å) ^a	50.00-2.35 (2.40-2.35)	50.00-1.97 (2.00-1.97)
R_{merge} ^b (%)	0.057 (0.431)	0.040 (0.257)
$\langle I/\sigma \rangle$	10.3 (1.2)	19.7 (3.0)
CC _{1/2}	0.574	0.781
completeness (%)	99.9 (98.2)	99.9 (100.0)
redundancy	9.6 (4.7)	5.1 (5.1)
Refinement		
$R_{\text{work}}/R_{\text{free}}$ ^d (%)	18.9/23.8	17.1/20.9
unique reflections	18981	31942
Mean B Factor (Å ²)		
protein	47.40	27.57
ligand	44.68	32.56
solvent	41.10	25.53
Ramachandran Plot		
most favored (%)	95.7	97.6
add. allowed (%)	3.6	2.3
RMSD		
bond lengths (Å)	0.008	0.011
bond angles (degree)	1.184	0.907

^aValues in parentheses are for the highest resolution shell.

^b $R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum I$ where I is the integrated intensity of a given reflection.

^c $R_{\text{work}} = \sum |F(\text{obs}) - F(\text{calc})| / \sum F(\text{obs})$.

^d $R_{\text{free}} = \sum |F(\text{obs}) - F(\text{calc})| / \sum F(\text{obs})$, calculated using 5% of the data.

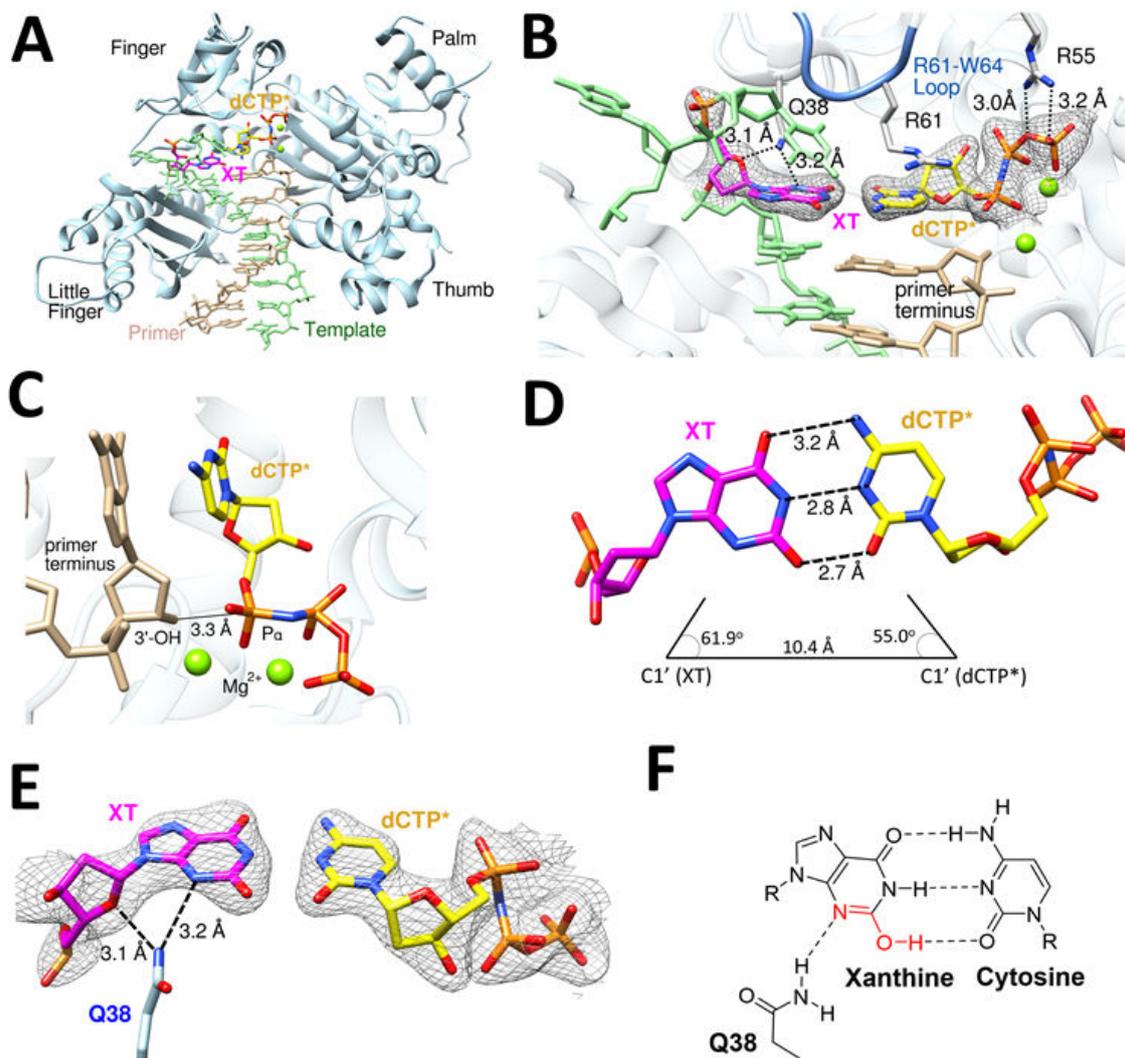


Figure 3. Ternary complex structure of pol η incorporating dCTP* opposite templating xanthine. (A) Overall structure of the pol η -XT:dCTP* ternary complex. The templating XT and incoming dCTP* are colored in magenta and yellow, respectively. (B) Close-up view of the replicating base pair site of the pol η -XT:dCTP* ternary complex structure. A 2F_o-F_c electron density map contoured at 1 σ around XT and dCTP* is shown. (C) Coordination of Mg²⁺ ions in pol η catalytic site. The distance between the 3'-OH of primer terminus and the P α of incoming dCTP* is indicated. (D) Base pairing geometry of XT and incoming dCTP*. The C1'-C1' distance and λ angles of XT:dCTP* base pair are shown. (E) The minor groove contact by Gln38 of pol η . The side chain amine moiety of Gln38 is hydrogen bonded to the O4' and N3 of templating xanthine. A 2F_o-F_c electron density map contoured at 1 σ around XT and dCTP* is shown. (F) Watson-Crick-like base pair between cytosine and the O2-enol tautomer (red) of xanthine, which can be promoted by Gln38-mediated minor groove contact.

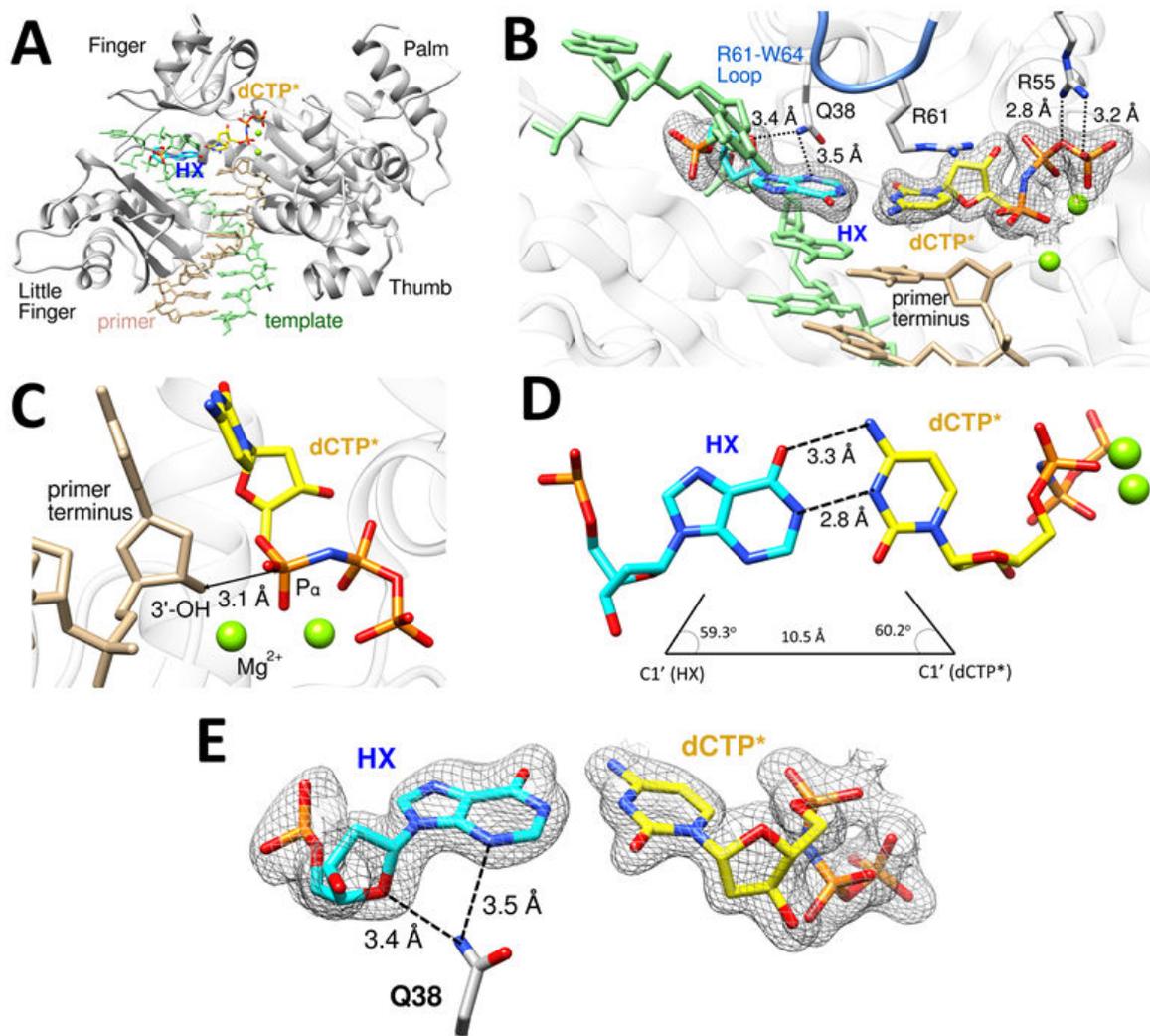


Figure 4. Ternary complex structure of pol η incorporating dCTP* opposite templating hypoxanthine. (A) Overall structure of the pol η -HX:dCTP* ternary complex. The templating HX and incoming dCTP* are colored in magenta and yellow, respectively. (B) Close-up view of the replicating base pair site of the pol η -HXT:dCTP* ternary structure. A $2F_o - F_c$ electron density map contoured at 1σ around HX and dCTP* is shown. (C) Coordination of Mg^{2+} ions in pol η catalytic site. The distance between the 3'-OH of primer terminus and the P α of incoming dCTP* is indicated. (D) Base pair geometry of HX and incoming dCTP*. The C1'-C1' distance and λ angles of HX:dCTP* base pair are shown. (E) Gln38-mediated hydrogen bonds with the templating hypoxanthine.

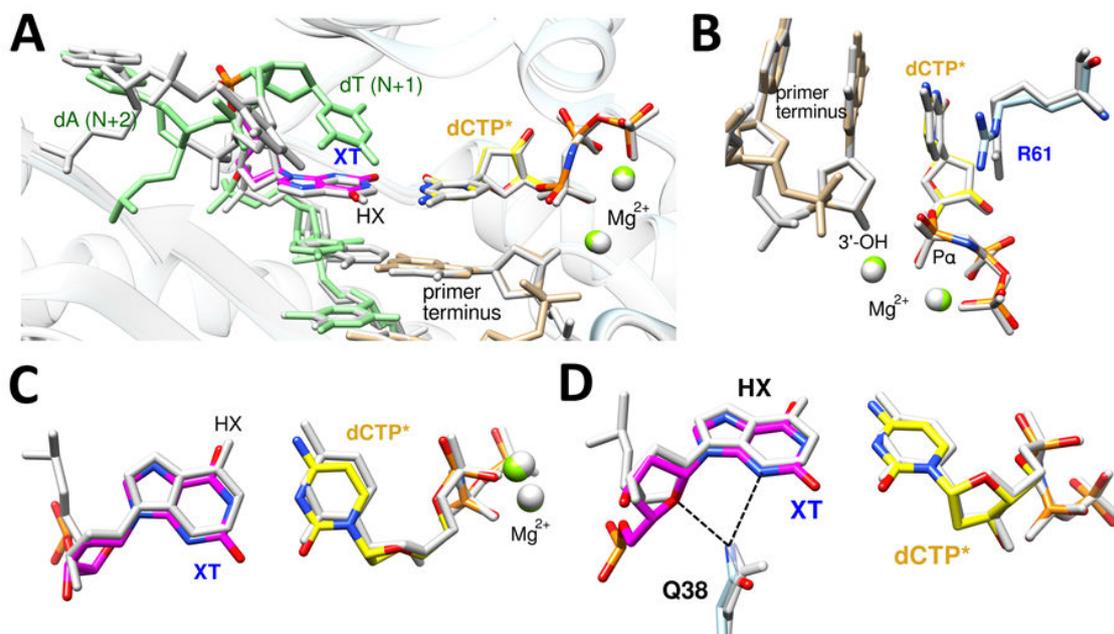


Figure 5. Structural comparison of pol η -XT:dCTP* and pol η -HX:dCTP*. (A) Superimposed structure of pol η -XT:dCTP* (multi-color) and pol η -HX:dCTP* (gray). (B) Active site conformation of dCTP, the primer terminus, two magnesium ions, and Arg61 of pol η -XT:dCTP* (multi-color) and pol η -HX:dCTP* (gray). (C) Overlay of the nascent base pairs of the pol η -XT:dCTP* and pol η -HX:dCTP* structures. (D) Comparison of replicating base pair site of pol η -XT:dCTP* and pol η -HX:dCTP*. Note that the conformations of the 5' phosphate group of the templating deaminated purines differ significantly.

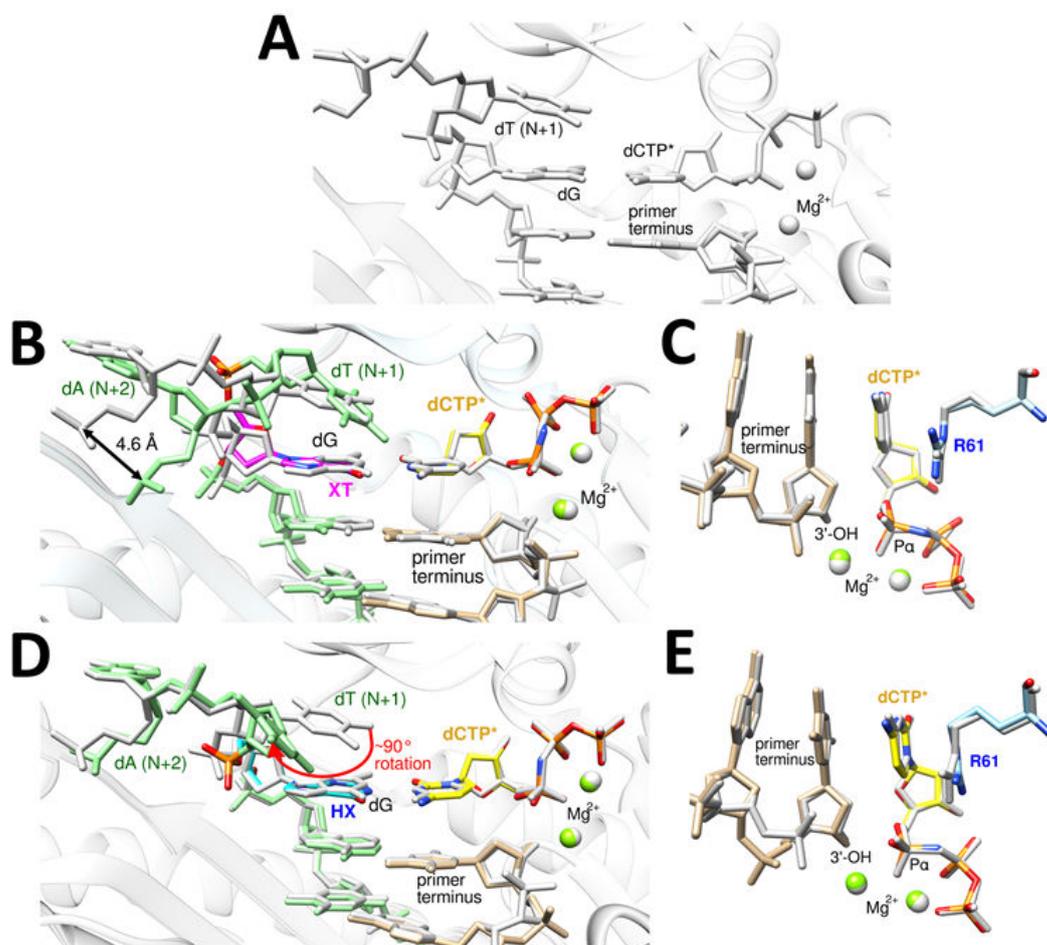


Figure 6. Structural comparison of the published pol η -dG:dCTP* structure with the pol η -XT:dCTP* and pol η -HX:dCTP* structures. (A) The active site structure of pol η -dG:dCTP* ternary complex (PDB ID: 4O3N) are shown in gray. The templating dG and incoming dCTP*, and two magnesium ions are shown. (B) Superimposed structures of the pol η -dG:dCTP* (gray) and pol η -XT:dCTP* (multi-colored) complexes. (C) Superimposed active-site structures of the pol η -dG:dCTP* (gray) and pol η -XT:dCTP* (multi-colored) complexes. The primer terminus, the incoming dCTP*, and Arg61 of pol η are shown. (D) Superimposed structures of the pol η -dG:dCTP* (gray) and pol η -HX:dCTP* (multi-colored) complexes. (E) Superimposed active-site structures of pol η -dG:dCTP* (gray) and pol η -HX:dCTP* (multi-colored) complexes. The primer terminus, the incoming dCTP*, and Arg61 of pol η are shown.

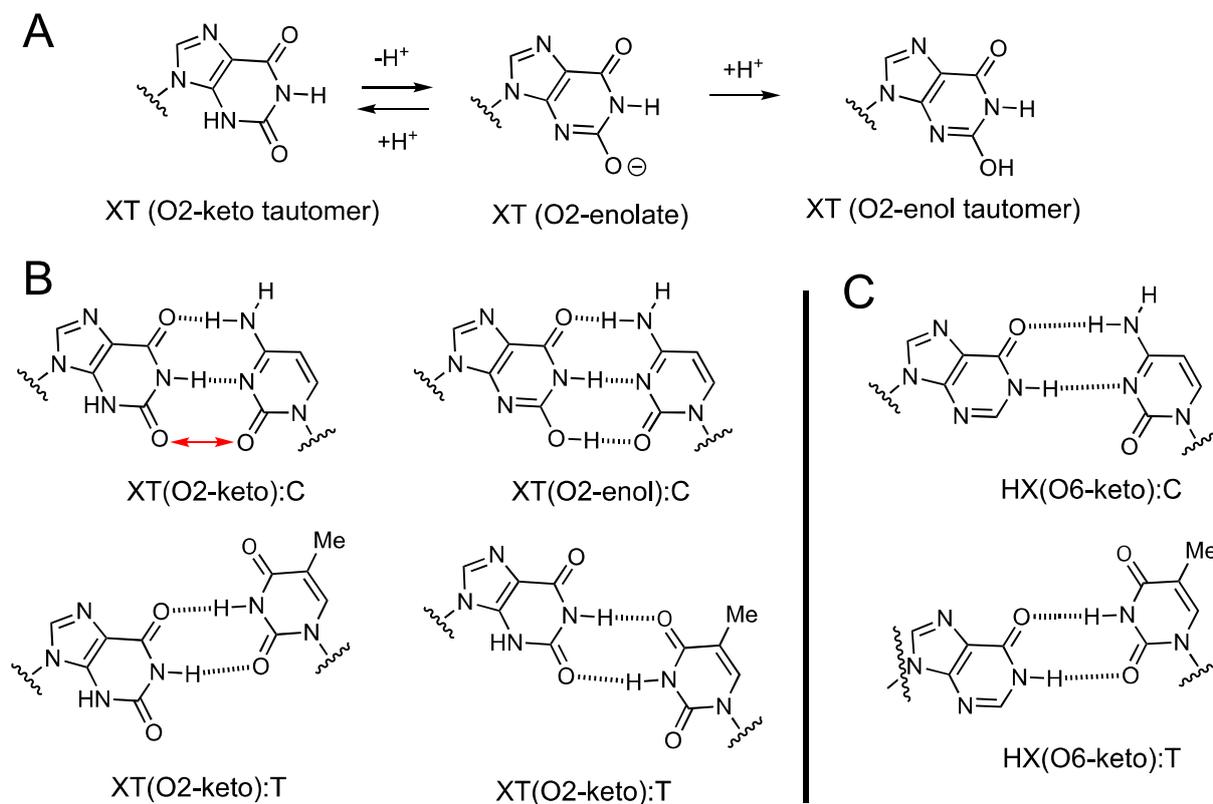


Figure 7. Base pairing properties of xanthine and hypoxanthine. (A) Tautomerization of xanthine. XT is in equilibrium between the O2-keto and O2-enol tautomers. (B) Base pairing properties of xanthine. Xanthine in O2-keto tautomeric conformation would experience a repulsive interaction with O2 of cytosine. Xanthine in O2-enol tautomeric conformation can form three hydrogen bonds with cytosine. Xanthine would form a wobble base pairing with thymine. (C) Base pairing properties of hypoxanthine. Hypoxanthine in O6-keto tautomer can form a Watson-Crick base pairing with cytosine and a wobble base pairing with thymine.

Goal-Oriented Hope and Sustainable Technology Adoption by Smallholder Farmers: Evidence from Senegal

Shira Peles (Bukchin), Dorit Kerret

Abstract— Significant climate change is predicted to impact smallholder farmers in developing countries, while these farmers paradoxically contribute to climate change. Why are most smallholders not using proven, modern cultivation technologies and best practices? This study introduces a new perspective, suggesting that cognitive goal-oriented hope prompts the adoption of sustainable technologies. The case study selected to represent this theoretical approach is drip irrigation adoption in Senegal. Data were collected by means of a face-to-face questionnaire. Binary logistic regression analysis of the collected data from 335 plots in Senegal, showed a significant connection between drip irrigation adoption and hope. The relative effect of hope on drip irrigation adoption proved greater than the effect of other previously studied factors. In addition to its theoretical innovation and empirical contribution, the significance of this research lies in its practical implications, since education, policy and communication may impact hope levels.

Keywords— agriculture, sustainable technology adoption, senegal, cognitive goal-oriented hope, positive psychology.

Mistreatment of Women in Mental Health Systems

T. A. Rossman

Abstract— Traditionally, women worldwide have disproportionately been and continue to be diagnosed with mental illness and mistreated in mental health systems. This misogynistic tactic has often been used to deprive women of human and civil rights. Using unconfirmed mental diagnoses (e.g., hysteria) to justify harmful treatments to control women has been commonplace for centuries. Such treatments include long-term or permanent internment in mental institutions, lobotomies, electric shock treatments, and drugs that cause short and long term intellectual and physiological damage. The combination of women's internalized patriarchal conceptions and misogyny, combined with financial incentives for mental health professionals to diagnose women with severe mental illness, has led to increased frequency of psychiatric treatment. Objective reporting from inside an asylum showed that women who did not conform to social norms were easily institutionalized, treated worse than criminals and horribly abused, sometimes in life-threatening ways. Women are more likely than men to be given electroconvulsive therapy (ECT), hospitalized for psychiatric illness with or without their consent, and prescribed brain-altering psychotropic medications, selective serotonin reuptake inhibitors, neuroleptics, anti-psychotics, highly addictive drugs (such as sedatives and benzodiazepines). Psychiatric drugs have been shown to cause chronic brain impairment (CBI), organ damage, tardive dyskinesia and other nervous system disorders. Yet based on cumulative research data, it appears highly unlikely that the reasons women are more frequently diagnosed with severe mental illness are due to bona fide physiological mental disorders. Instead, evidence points to the efficacy and lucrative outcome of psychiatrically diagnosing and treating women as primary motives for such diagnoses. Women's negative emotions and behaviors, which have been cited as proof of mental illness, are likely to be sourced from their inferior socio-cultural and economic status; this stems from social and economic oppression and repression as well as overt and covert familial, spousal, and institutional abuse. Psychiatric misdiagnosis and subsequent treatment abuses are ongoing, fueled by billions of dollars of revenue for the U.S. mental health industry. Notably, they are largely based on medical insurance codes citing diagnoses from the Diagnostic Statistical Manual (DSM) composed and published by the American Psychological Association (APA). Psychiatrists of the late 20th and early 21st century expanded their market by prescribing new drugs to vulnerable and unhappy women; this was enabled by the APA's expansion of diagnostic categories in their DSM-V. The DSM-V serves the mental health industry in terms of increased profits, maintenance of patient caseloads and operations, status, and power to control women. DSM Task Force psychiatrists, employed by the APA and responsible for adding new psychiatric diagnoses to each DSM revision, reported financial ties to pharmaceutical companies who stood to benefit financially from adding specific diagnoses. This comprehensive review of the history of diagnosis and treatment of perceived and real mental disorders in women provides historic and current data on diagnosis and treatment, analyzing and discussing findings through the framework of feminism and objectification theory. Research implications include the reinforcement of the negative impact of patriarchal socio-cultural and economic systems that exploit and marginalize women. Exposing the widespread, multifaceted forms of control and objectification of women in our society and the insidious impact it has on the socialization and development of girls and women is crucial to implementing positive change.

Keywords— mental health services, feminism, injustice in mental health care, psychological mistreatment.

Formation of an Empire in the 21st Century: Theoretical Approach in International Relations and a Worldview of the New World Order

Rami Georg Johann

Abstract— Against the background of the current geopolitical constellations, the author looks at various empire models, which are discussed and compared with each other with regard to their stability and functioning. The focus is on the fifth concept as a possible new world order in the 21st century. These will be discussed and compared to one another according to their stability and functioning. All empires to be designed will be conceptualised based on one, two, three, four, and five worlds. All worlds are made up of a different constellation of states and relating coalitions. All systems will be discussed in detail. The one-world-system, the “Western Empire,” will be presented as a possible solution to a new world order in the 21st century (fifth concept). The term “Western” in “Western Empire” describes the Western concept after World War II. This Western concept was the result of two horrible world wars in the 20th century.” With this in mind, the fifth concept forms a stable empire system, the “Western Empire,” by political measures tied to two issues. Thus, this world order provides a significantly higher long-term stability in contrast to all other empire models (comprising five, four, three, or two worlds). Confrontations and threats of war are reduced to a minimum. The two issues mentioned are “merger” and “competition.” These are the main differences in forming an empire compared to all empires and realms in the history of mankind. The fifth concept of this theory, the “Western Empire,” acts explicitly as a counter model. The Western Empire (fifth concept) is formed by the merger of world powers without war. Thus, a world order without competition is created. This merged entity secures long-term peace, stability, democratic values, freedom, human rights, equality, and justice in the new world order.

Keywords— empire formation, theory of international relations, Western Empire, world order.

Trust in the Police: How Political Identity Shapes Violence Victims' Engagement with the State

Tasnia Symoom

Abstract— Though reporting violence to the police has been extensively investigated, the nature of victims' reporting violence against women (VAW) based on their political identities has been little explored. In this research, I examine how political identities like race and ethnicity can shape VAW victims' reactions to reporting to police when they are assaulted and sexually assaulted. Empirically, taking advantage of the National Crime Victimization Survey (1992-2018), I conduct a fixed effect regression analysis to test the effect of race and ethnicity of VAW victims' reporting to police. I also examine the reaction of VAW victims when they don't believe in police effectiveness. The result shows that women are less likely to report to police when assaulted or sexually assaulted and when they think that Police are not effective. Moreover, black women's reporting to police is heavily contingent on their relationship with the police. Even though assaulted Black women are more likely to report to police than women of other races, they are less likely to report to police than women of other races when they think that police are not effective. Also, the nature of violence against women affects whether they would report to the police or not. I find significant evidence that assaulted Hispanic women are more likely to report to police than women who are not Hispanic but are less likely to report when they are sexually assaulted.

Keywords— Trust in police, Violence against women, sexual assault, assault, Black women, Hispanic women, police report, National Crime Victimization survey.

Memories of Lost Fathers: The Unfinished Transmission of Generational Values in Hungarian Cinema by Peter Falanga

Peter Falanga

Abstract— During the process of de-Stalinization that began in 1956 with the Twentieth Congress of the Soviet Communist Party, many filmmakers in Hungary chose to explore their country's political discomforts by using Socialist Realism as a negative model against which they could react to the dominating ideology. A renewed national film industry and a more permissive political regime would allow filmmakers to take to task the plight of the preceding generation who had experienced the fatal political turmoil of both World Wars and the purges of Stalin. What follows is no longer the multigenerational unity found in Socialist Realism wherein both the old and the young embrace Stalin's revolutionary optimism; instead, the protagonists are parentless, and thus their connection to the previous generation is partially severed. In these films, violent historical forces leave one generation to search for both a connection with their family's past, and for moral guidance to direct their future. István Szabó's *Father* (1966), Márta Mészáros *Diary for My Children* (1984), and Pál Gábor's *Angi Vera* (1978) each consider the fraught relationship between successive generations through the lens of postwar youth. A characteristic each of their protagonist's share is that they are all missing one or both parents, and cope with familial loss either through recalling memories of their parents in dream-like sequences, or, in the case of *Angi Vera*, through embracing the surrogate paternalism that the Communist Party promises to provide. This paper considers the argument these films present about the progress of Hungarian history, and how this topic is explored in more recent films that similarly focus on the transmission of generational values. Scholars such as László Strausz and John Cunningham have written on the continuous concern with the transmission of generational values in more recent films such as István Szabó's *Sunshine* (1999), Béla Tarr's *Werckmeister Harmonies* (2000), György Pálfi's *Taxidermia* (2006), Ágnes Kocsis' *Pál Adrienn* (2010), and Kornél Mundruczó's *Evolution* (2021). These films, they argue, make intimate portrayals of the various sweeping political changes in Hungary's history and question how these epochs or events have impacted Hungarian identities. If these films attempt to personalize historical shifts of Hungary, then what is the significance of featuring characters who have lost one or both parents? An attempt to understand this coherent trend in Hungarian cinema will profit from examining the earlier, celebrated films of Szabó, Mészáros, and Gábor, who inaugurated this preoccupation with generational values. The pervasive interplay of dreams and memory in their films invites an additional element to their argument concerning historical progression. This paper incorporates Richard Teniman's notion of the "dialectics of memory" in which memory is in a constant process of negation and reinvention to explain why these Directors prefer to explore Hungarian identity through the disarranged form of psychological realism over the linear causality structure of historical realism.

Keywords— film theory, Eastern European Studies, film history, Eastern European History.

Representing the Un-representable in Zhang Yimou's Coming Home

Min Yang

Abstract— This article studies Zhang Yimou's *Coming Home* (Guilai) and the novel *The Criminal Lu* upon which the film is based. I demonstrate that the novel diminishes the trauma that a family has experienced during disastrous political movements, especially the Cultural Revolution. In contrast, Zhang's filmic adaptation visualizes the unspeakable trauma that Chinese individuals experienced during the Cultural Revolution and the difficulty of healing this trauma. I also consider *Coming Home* within Zhang's larger oeuvre and his other films on the Cultural Revolution, examining viewer responses and the current social-political context regarding traumatic memories about the Cultural Revolution. I contend that *Coming Home* constitutes a visual call for assimilating the reminiscences of the national trauma that have been repressed, buried and seemingly forgotten.

Keywords— trauma, the Cultural Revolution, traumatic memory, representation.

The Impact of Technology on Media Content Regulation

Eugene Mashapa

Abstract— The age of information has witnessed countless unprecedented technological developments, which signal the articulation of succinct technological capabilities that can match these cutting-edge technological trends. These changes have impacted patterns in the production, distribution, and consumption of media content, a space that the Film and Publication Board (FPB) is concerned with. Consequently, the FPB is keen to understand the nature and impact of these technological changes on media content regulation. This exploratory study sought to investigate how content regulators in high and middle-income economies have adapted to the changes in this space, seeking insights into innovations, technological and operational, that facilitate continued relevance during this fast-changing environment. The study is aimed at developing recommendations that could assist and inform the organisation in regulating media content as it evolves. Thus, the overall research strategy in this analysis is applied research, and the analytical model adopted is a mixed research design guided by both qualitative and quantitative research instruments. It was revealed in the study that the FPB was significantly impacted by the unprecedented technological advancements in the media regulation space. Additionally, there exists a need for the FPB to understand the current and future penetrations of 4IR technology in the industry and its impact on media governance and policy implementation. This will range from reskilling officials to align with the technological skills to developing technological innovations as well as adopting co-regulatory or self-regulatory arrangements together with content distributors, where more content is distributed in higher volumes and with increased frequency. Importantly, initiating an interactive learning process for both FPB employees and the general public can assist the regulator and improve FPB's operational efficiency and effectiveness.

Keywords—Film and Publication Board (FPB), Media Regulation and Technology.

I. INTRODUCTION

Technological disruptions have meant that media content needs to be classified in real time as these forms of media content are produced and distributed in real time, especially online media content. According to [1] “streaming is not just a means of distributing content but, it is fundamentally a two-way communications channel” between the distributor and the consumer. “The data that streaming platforms collect about consumer behaviour can be used to support more targeted, and therefore less risky, decisions about what to create”[1]. COVID-19 has also further accelerated the growth and utilisation of on-demand media content on online platforms. It is therefore pivotal for the regulator (i.e. FPB) to adjust its operations to meet and address these unprecedented technological transitions in the media space. Thus, the study seeks to explore the future of media technology and practices

beyond the year, 2030, and position the FPB as a trusted and relevant media content regulator. Furthermore, our study seeks to develop a global comparative analysis highlighting existing best practice models (i.e. media content regulatory models) and how can the FPB incorporate these effective regulatory processes in its business operations. Research provides an interactive learning process, hence, it should be noted that knowledge is the most important resource in any economy and learning is the greatest process for enhancing economic development as well as collaboration.

II. RESEARCH METHODOLOGY

The overall research strategy and/or approach in this analysis is applied research. According to [2], applied research is “the research techniques, procedures, and methods that form the body of research methodology and are applied to the collection of information about various aspects of a situation, issue, problem, and/or phenomenon [2]. One of the main objectives of applied research is to provide solutions to an articulated research problem in the most practical way. Reference [2], further notes that the “information gathered can be used in other ways-such as, policy formulation, administration, and enhancing an understanding of a phenomenon” [2].

It should also be noted that the analytical model adopted is a mixed research design guided by, both qualitative and quantitative research instruments. Mixed research designs involve analysing and interpreting both, qualitative and quantitative data, as well as the integration of conclusions from those data into a cohesive whole [3]. Qualitative research instruments in this regard refers to the collection of qualitative forms of data sets examined from different angles for purposes of constructing a rich and meaningful picture of complex multifaceted phenomenon's [4]. Then again, quantitative research instruments tend to have a particular, prespecified focus [5].

At the outset, the researcher develops data-collection instruments, including sources of data and how will the data be collected and then, subsequently analysed in the research project. In short, the current study utilised a mixture of desktop literature reviews, stakeholder engagements, focus group discussions, as well as online interviews conducted using both quantitative and qualitative analytical instruments from identified sample respondents (i.e. key informant interviews).

III. RESEARCH FINDINGS AND ANALYSIS

A. Survey Findings

Eugene Mashapa is with the Research, Policy and Advocacy Unit Film and Publication Board South Africa, (e-mail: eugene.mashapa@fpb.org.za)

From the findings of the survey, it was revealed that most participants form part of 'generation x' meaning, those individuals who are between the ages of 41-56. An age group usually characterised as born before the advent of technological advancements. Thus, this is an interesting finding given that, it is imperative to tackle the upskilling of the workforce with technological capabilities. We need to consider their age and experience with the new technologies. Nonetheless, [10] notes that, "Gen Xers eventually become comfortable with using personal electronic devices like computers and smartphones... while they're not digital natives like Gen Z, most Gen Xers can still grasp today's technology quickly [6]. At the same time, the findings also indicated that within the age cohort of the study there were Millennials or generation Y (i.e., those between the ages 25-40) and, according to the Corporate Finance Institute, "millennials are considered the most connected generation which grew up during the internet age" [7].

The survey also highlighted the fact that, individuals employed in the regulation space fall under the generation X cohort and have extensively been in the industry thus, experienced in the field and provide critical invaluable information to the media content industry. Another interesting finding revealed in the survey is that, most respondents have a Post Schooling, Education and Training (PSET) degree, with 19% reporting having a PhD, 33.3% having a master's degree, 14.3% having an Honours degree and 14.3% holding a bachelor's degree respectively. This high level of education can also be advantageous meaning that, most content classifiers, and regulators are educated as well as 'educatable'. This is what's needed as we argue in the study that, in such a rapidly changing world where technological innovations are key, there exists a need for a workforce that is willing to learn and expand their skills to meet cutting-edge technologies in the media content regulation space

B. Game Classification

Game classification is a major challenge because games have stages/levels and content classifiers cannot sit down and go through all the stages without the prerequisite gaming skills. During an interview with the Acting Chief Operations Office of the FPB, held on the 18th of August 2021, it was noted that gaming classifications are tedious "because it could take a year for a classifier to go through all the stages and give an outcome". Instead, game classifiers are meant to provide a *Game Play* which shows all the levels of the games, and this is some form of self-classification. The classifiers in this process thus become monitors and evaluate if the game play has not been manipulated, and that it represents the actual classification attached to it. Consequently, games are released internationally on the same day to limit piracy, and this is another major challenge that the FPB is currently facing. The delays in the classifications of games means that some games are released without the correct South African classification which can expose consumers to prohibited content. Thus, there arise a need for gaming classification to move online and make sure that content is classified in real time as well as making sure that

all games distributed in the South African market are classified accordingly, considering the prescribed legislative framework as stipulated in the Film and Publication Act 11 of 2019.

C. Micro-Budget Film Sector

the micro-budget film sector has revolutionised the film industry with compact making of films at a very minimal budget and small crews. Nonetheless, survey/interview participants noted that, lockdown has also amplified the trends and patterns of the micro-budget film sector. While cinemas were closed during COVID 19 country lockdowns, we saw this sector distributing content on their own platforms and on various social media websites. Massive reductions during the COVID-19 pandemic in mainstream film productions gave rise to this marginalised sector. Given the production directives of only 50 crew members on site according to the COVID-19 national task team, the micro budget films were done far more collaboratively, remotely, with optimised collective skills and a reduced large carbon footprint (Respondent 4). There should be concerted effort by government to increase film incentives for this fast-growing sector.

The growth of the micro-budget film industry has meant that films are distributed and marketed online with very little budgets. Nonetheless, micro-budget films are easily accessible and generate acceptable revenue for distributors and producers. This is further supported by the argument of one focus group participant who notes that, "some content creators are able to make a living from user-generated digital streaming platforms like YouTube but, the remuneration algorithm doesn't favour content which is more time-consuming to make". Thus, smartphones have benefited the micro-budget film sector and "communication plays an important role in the fundamental operation of a society. It links people and businesses, facilitating communication, the flow of ideas, information, coordinating economic activities and development" [8].

D. International Benchmarking

This study is well timed as the world is confronted with powerful economic and/or technological transitions brought about by the advent of technological advancements. The COVID-19 pandemic has also punctuated the advent of the Fourth Industrial Revolution, and the FPB is no exception in this regard. Nonetheless, technological changes do not happen on their own, even though these transitions have been the greatest force to change society. We need to bear in mind that technology is a subset of society and depends on civilization for its growth and existence. These technological advancements will naturally drive potential and future challenges in the way regulatory bodies classify content, as well as how they frame their insights into the near future.

For instance, The Canadian media content classification model is to be reckoned considering the influx of online publications by distributors. Canada has four provincial classification agencies that rate and classify media content as from the 1st October 2019 [9]. These provincial classification agencies comprise of the Minister of Culture and Communications (Quebec), Government of Nova Scotia,

British Columbia Film Classification and the Alberta Film Classification agencies [9]. There are 10 Canadian provinces and the abovementioned classification agencies segregate their classification duties and responsibilities amongst these 10 provinces.

Furthermore, according to Christopher Götz author of “*Telecoms and Media in the European Union*” the primary audiovisual policy in the European Union is the Treaty on the Functioning of European Union, Articles 167 and 173 [10]. The European Union audiovisual Directive for media services was adopted for purposes of synchronising existing media content regulations. Additionally, within the European Union, the creation, distribution and consumption of media content (inclusive of broadcasting and on-demand/online media content) is informed and regulated by Directive 2010/13/EU (Audiovisual Media Services Directive) [10].

Moving to the African continent, Tunisia is considered to be one of the smallest countries in north Africa with an estimated population of eleven million inhabitants. The cultural outlook of Tunisia is rooted in its centuries-long intersection of different cultures and ethnicities from the west and within Africa. Nonetheless, the primary regulatory authority of media content is the High Independent Authority for Audio-visual Communication (HAICA) as established by the primary legislation, the Decree Law 116 dated 2 November 2011 (Film and Publication Board, 2021). According to Article 7 of the Decree Law 116, 2011, HAICA is constituted of a collective body consisting of nine independent members who possess the pivotal expertise, competence and integrity in the media and communication sector.

While in Nigeria, the primary regulatory authority mandated to regulate films and videos is the National Film and Video Censors Board (NFVCB) established by the country’s legislation for governing films and videos i.e. the National Film and Video Censors Board Act 85 of 1993. The NFVCB has also published Censorship and Classification Guidelines during the year 2006 in terms of the National Film and Video Censors Board Act 85 of 1993. Additionally, the Act has also provided for regulations published in the year 2008. In terms of the Classification Guidelines 2006, elements classifiable include, violence, sex, coarse language, adult themes, drug use, nudity, imitable techniques, traditional rituals and customs.

IV. CONCLUSIONS

It is imperative for the FPB to understand current and future penetrations of technology in the industry and its impact on media governance and policy implementation. Thus, developing an understanding is imperative for purposes of remaining effective and increasingly efficient in the regulation and classification of media content. An interactive learning process between media content creators, classifiers, regulators, and industry related sectors such as the ICT community needs to be initiated for purposes of enhancing the operational efficiency and effectiveness of Africa’s leading media content regulator. This will assist the FPB develop pivotal technological capabilities which will be able to comprehend unprecedented technological advancements.

As media content regulators (i.e., the FPB) we need to wake up to the reality that self-classification is an alternative model to address the unprecedented volume of online user generated media content. Simultaneously, various monitoring tools and guidelines needs to be in place to regulate the precarious ‘human element’ in self-classification models. Self-classification guidelines should prioritize the South African context but also take into the consideration the international context and the technologically disruptive content distribution channels.

It was also argued in the survey that, the early adoption of Online Content Regulation (OCR) model was a major shortcoming due to fact that extensive evaluations of the system was not conducted and existing glitches in the system were not identified early while piloting the OCR system. More tests need to be conducted and software developers need to take heed of the work conducted by the FPB and its preponderant contribution in the media content regulation space. In conclusion, media content in South Africa should be aligned with the Constitution and must not propagate hate speech, incitement of violence, propaganda for war, child pornography and bestiality in all its forms.

ACKNOWLEDGMENTS

Special thanks to the survey participants, stakeholder engagements, focus group participants and individuals who participated in the various interviews conducted.

REFERENCES

- [1] Raustiala, K. & Sprigman, C., 2019. *The Second Digital Disruption: Streaming & the Dawn of Data-Driven Creativity*. *ChrNew York*
- [2] Kumar, R., 1999. *Research Methodology: A Step by Step Guide for Beginners*. London: Sage Publication.
- [3] Leedy, P. & Ormrod, E., 2013. *Practical Research: Planning and Design*. 10th ed. s.l.:s.n.
- [4] Wisker, G., 2008. *The Post Graduate Research Textbook research Handbook*. 2nd ed. New York : Palgrave Macmillan.
- [5] Leedy, P. & Ormrod, E., 2013. *Practical Research: Planning and Design*. 10th ed. s.l.:s.n.
- [6] Hutton, E., 2022 . *What is generation X? The "Forgotten Generation"*. [Online] Available at: <https://imagerestorationcenter.com/what-is-gen-x/> [Accessed 30 March 2022].
- [7] Corporate Finance Institute , 2015. *Millennials*. [Online] Available at: <https://corporatefinanceinstitute.com/resources/knowledge/other/millennials/> [Accessed 30 March 2022].
- [8] Statistics South Africa. 2018. *General Household Survey*. Available at: [General Household Survey, 2018 | Statistics South Africa \(statssa.gov.za\)](https://www.statssa.gov.za/)
- [9] Film and Publication Board, 2021. *Final Country Report for Canada - FPB CCI Project*, Centurion: Film Publication Board.
- [10] Götz, C., 2021. *In brief: media law and regulation in European Union*. [Online] Available at: <https://www.lexology.com/library/detail.aspx?g=98b9cbeb-5f49-468e-9eb3-6094fdbf24f1> [Accessed 28 March 2022].

Life Orientation Teachers' Perceptions of HIV Education for Diverse Sexual Identities in South African Schools

Jesse Ashley, Makaringe Gezani Given

Abstract— This paper deals with the perceptions of Life Orientation Educator's responsiveness to diverse sexual identities. The paper focuses on Life Orientation (LO) as a core platform for HIV and Sexuality Education. It is therefore concerned with the challenges that are faced by diverse sexual identities and the educators' experiences in teaching HIV and Sexuality Education to Lesbians, gays, bisexuals, transgender, and queer in schools. The LO curriculum and educational policies are discussed to understand the role of educators in teaching diverse sexual identities about HIV prevention in schools. The paper further deals with the importance of teaching LO, the reason for teaching it, and the challenges faced by LO educators in teaching diverse sexual identities in schools. It uses constructivism theory, and social learning theory to allude to the perceptions of LO educators in regards to learning about diverse sexual identities. The paper uses qualitative research to collect data as participants are from different schools in the Gauteng Department of Education. Participants are LO educators from primary, secondary, and special schools, and all interviews were conducted after contact time outside of the school due to COVID-19 compliance, so there was no need for permission from the schools. Themes were used in the analysis of data. The paper recommended and concluded that Educators teaching Life Orientation do not know or understand the LO curriculum. The findings of the paper showed that teachers do not teach the whole curriculum but choose certain content which suits their needs but not learners.

Keywords— diverse sexual identities, life orientation, sexuality education, HIV education, LGBTQ, department of education, educational policies, LO curriculum.

Buthainaal-Issa's Ques, Lilao Thaleb (Beauty and The Beast): A Feminist Cultural Study

Arwa Albader

Abstract— The concept of Prince Charming is a long-standing one. Normally, “Prince Charming’ appears as the villain for whom girls foolishly await in both fairy tales and romantic stories” (Bottigheimer 230). This mythical ideology represents an archetypal narrative form, constituting a literary commonplace in fairy tales from multiple regions of the world and different time periods. Underlying this archetype is the ideal of female beauty as a pervasively physical attribute, implying thereby that the only distinguishing characteristics a woman should possess are beauty and charm. Within the fairy tale narrative framework, these physical qualities are typically situated as sources of power and pleasure for men; possessing beautiful women then signals male social power and authority. Different versions of the fairy tale Beauty and the Beast consistently depict Beauty’s relationship with male characters (her father and husband) in the context of firmly circumscribed power relations. This paper examines the ideas informing such power relations through a feminist cultural lens, considering how they cumulatively evoke and reinforce, or perhaps challenge, particular stereotypes of women in ButhainaAl-Issa’s Al Jameelah O Wahash (The Beauty and the Beast) (2011). The central focus of the thesis is Beauty’s relationship with the male figures, focusing on the role of marriage, engaging arguments about why is “marriage” considered a reward for a beautiful woman.

Keywords— feminism, cultural study, fairy tale, adaptation, marriage in middle eastern fiction, charming prince.

AI in Hungerstation

Latifa AlZamil

Digital Marketing

Dar Al-Hekma University

Jeddah, Saudi Arabia

lsalzamil@dah.edu.sa

Fatima Mohammed Baba

Psychology

Dar Al-Hekma University

Jeddah, Saudi Arabia

fmbabaduna@dah.edu.sa

Fatimah AlGhalib

Digital marketing

Dar Al-Hekma University

Jeddah, Saudi Arabia

fatghalib@dah.edu.sa

Joanne Jifri

Psychology

Dar Al-Hekma University

Jeddah, Saudi Arabia

jmjifri@dah.edu.sa

Abstract— The research presented in this report analyzes the performance of artificial intelligence in the delivery app HungerStation. Through evaluating the survey responses of users of the HungerStation app, the effectiveness and accuracy of the tracking AI which provides the expected time of arrival was assessed as well as overall satisfaction with the app. The results of this report demonstrate that there is both user satisfaction as well as perceived effectiveness of the HungerStation AI.

I. INTRODUCTION

HungerStation, the first delivery app in the Kingdom of Saudi Arabia is one of the fastest delivery apps to exist. It was founded by Ebrahim Al-Jassim in 2012. It started in Al-Khobar and later started operating all over Saudi and recently Bahrain.

HungerStation is very easy to use. The first step is to add the location you want the product to be delivered, second is to choose your preference of restaurant, grocery, pharmacy, or flower shop, and the third step is to pay, and wait for your order to be delivered. The consumer has the choice of the payment method. Such as Apple Pay, online payment, or cash on delivery. The consumer also has the option to track the delivery of their ordered products. That way they would have an estimation on what time the product will be delivered. The app itself also mentions to calculated estimated time of the delivery

This paper will discuss the artificial intelligence that is present on the Hungerstation app. We have surveyed around 10 people asking them questions about their use of the app. The background, methodology, and results of the survey, and the AI of the app will be discussed in this paper.

II. BACKGROUND

As mentioned above, Hungerstation is the first Saudi delivery company. It was founded in March 2012 in Al-Khobar. It spread to the rest of the Kingdom including 115 of the cities of Saudi, and Bahrain. The company started with delivering food from fast food restaurants such as McDonalds, and Burger King, by time it grew to deliver food from other local, and international restaurants such as SteakHouse, and Shrimp Anatomy. With time HungerStation started forming partnerships with grocery shops, pharmacies and flower shops. That way consumers will not only use the app when hungry, they'll use it when they need a snack from a nearby grocery store, or when they'd send their loved ones flowers. HungerStation now has more than 10,000 partnerships with restaurants, grocery shops, and flower shops.

Helping people order the food they want at the time they want, in order for them to enjoy the perfect meal from the convenience of their home or office. That was the aim of the founder of Hungerstation, Ebrahim Al-Jassim. The platform made it easier by having many restaurants, pharmacies, and flower shops to be found in one place rather than looking for the phone numbers in the phone book or google.

Delivery apps were not that popular as people found them difficult to accept. However, due to the COVID-19 pandemic and the lockdown that happened in early 2020, the home delivery market increased rapidly due to the extremely high demand. During the lockdown during the pandemic, all that was available were online purchases as we could not leave our homes. Clothes were bought online, groceries were bought online, food was bought online, and medication was bought online as well. During that time the delivery apps shined through. Especially HungerStation as it usually has free deliveries or discounts on deliveries. The different delivery apps decided to add more services such as providing pharmacies, groceries, and stores to purchase from, as it was the only way. There are several alternative applications that are competitors of Hungerstation such as, ToYou, Lugmetry, The cheffz, Jahez, and many more. Hungerstation is one of 3 delivery apps that offers restaurants, pharmacies, and flower shops making it one of the most popular apps alongside ToYou and Jahez. HungerStation became a site where you have the option to order everything you desire in a simple and quick manner, with the option of either paying online or cash on delivery.

The Hunger Station app has a Unique Value Proposition of launching this demand food delivery service platform in which the more they satisfy the user's needs, the higher revenue they earn. After Google maps, it is said that Hunger station is the most frequently used app in Saudi Arabia. There are about 20k product lines registered on the app meaning that there is something for everybody. Another interesting thing about the Hunger Station app is how it integrates Application Programming Interface (API) within the delivery app. This has made the app very functional and user friendly in where it provides updates about your order, availability of riders, inventory management, restaurant's opening and closing time and an order history.

III. METHODOLOGY

To begin our process of assessing the effectiveness of the Hungerstation AI we needed to collect data from users of the app. We aimed to understand the opinions of the Hungerstation's users in order to appropriately evaluate the artificial intelligence's performance in the Hungerstation App. To do so we sent out a survey with a collection of questions to people who have used the App. We distributed the survey to those who had used the app at least three times. We did this so that when they respond to the survey questions, their responses are not skewed by a single negative experience. We used a criterion sampling approach, which entails identifying an important criterion, articulating this criterion, and conducting an examination to determine if the sample fulfills the criterion. One of the reasons we picked this method of sampling is to find data which is valid as the opinions which we will take into consideration must be users of the app in order for them to answer with accuracy for us to determine if there are key flaws in the AI software that need to be addressed in the Hunger Station app.

We utilized both quantitative and qualitative research approaches to collect our data, as illustrated in Figure. 1.

4. Did you find this app easy to use the first time?

- Yes
- No

5. Which feature would you likely use the most?

6. What feature would you recommend them to add?

Figure. 1 A sample of the survey which demonstrates both yes and no questions as well as written questions.

3. 1. Did you feel like the expected time of arrival was accurate?

- Yes
- No

Figure 2. Illustrates the use of a yes or no question to ask about accuracy of the Hungerstation tracking AI.

The included questions were all short and understandable in order to ensure that all participants would be able to comprehend and answer the questions to the best of their abilities. The use of yes or no questions allowed us to collect

definitive quantitative data that would allow us to acquire a percentage. We used this method of questioning to understand whether or not the tracking AI was accurate and satisfactory for the users as well as to assess the user's overall preferences concerning the HungerStation app. We also use this type of questioning to determine if the users perceive the tracking AI to be accurate in order to determine a straightforward consensus seen in Fig. 2. Specifically, the accuracy of the expected time of arrival is evaluated through this question which appears at any time after the order is placed and with the courier as seen in Figure 3.

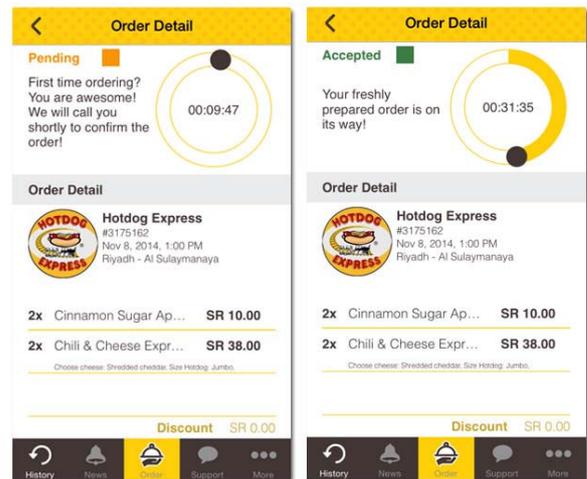


Figure 3. Shows the AI tracking feature which includes the expected time of arrival.

Another type of questioning used in our survey are open-ended questions. This method of questioning allowed the participants to not be inhibited by a previously determined group of answers. We used this method for questions involving feature preferences as well as recommendations that the users may have regarding the app as seen in Figure 1. As the most valid data collection for these questions would be acquired if the participants had total freedom in what they will answer. Instead of having a previously determined set of popular features to choose from, the participants were able to write whichever one they used most out of the various features offered by the HungerStation app which is illustrated in Figure 4.

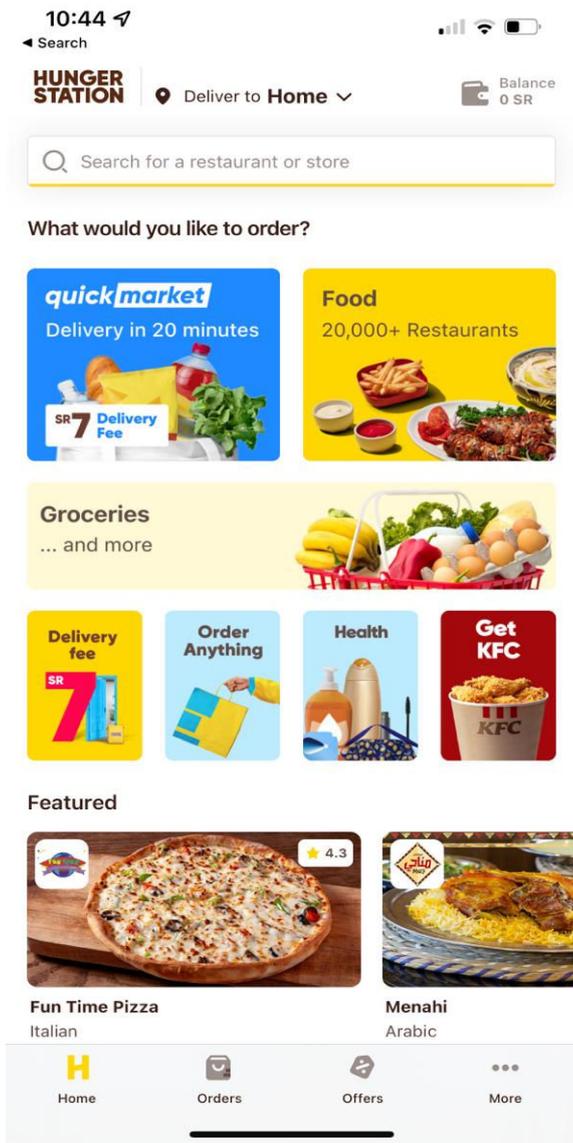


Figure 4. Shows the homepage of HungerStation with all the available features.

The next type of question we included was a rating question; where we asked our participants in the survey how they would rate their overall satisfaction with the HungerStation app. The participants were given a scale of 5 ranging from very satisfied to very dissatisfied as seen in Figure 5. This was to assess the level of satisfaction of HungerStation user's and to later on determine if there is a correlation between overall app satisfaction and the perceived accuracy of the tracking AI by the users. The variety in styles of questioning used in the survey was in order to collect different forms of data to evaluate the HungerStation using both quantitative and qualitative methods.

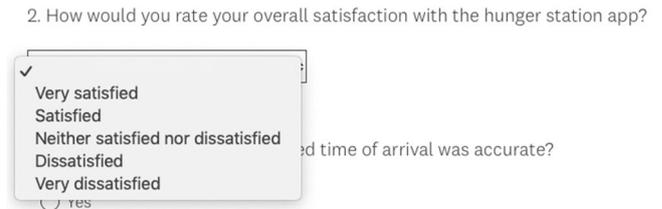


Figure 5. Illustrates the survey question which asks participants to rate the overall satisfaction with the HungerStation app.

I.V. RESULTS

A survey was conducted and sent around to obtain results about people's satisfaction with the AI in the HungerStation app. The survey included the following questions:

1. Do you prefer to chat with a real person or bot on the app?
2. How would you rate your overall satisfaction with the hunger station app?
3. Did you feel like the expected time of arrival was accurate?
4. Did you find this app easy to use the first time?
5. Which feature would you likely use the most?
6. What feature would you recommend them to add?
7. Did you find it easy to cancel your order?
8. How did you find the order tracking experience?
9. Do you feel comfortable with the app having a record of your past orders/information?
10. Do you prefer the auto notifications to be on or off?

For this report we decided to analyze questions 1, 3, 4, and 8 as they are the most significant in terms of analyzing the AI present on the app.

Question 1 asked the participants if they prefer to chat with a real person or with a bot on the app. 80% of the participants mentioned that they'd prefer chatting with a real person than with a bot. However, 20% said that they'd prefer chatting with a bot instead. Chatting with a bot to some would be a

waste of time as it could be hard to understand. Especially when you're looking for a certain answer. However, chatting with a bot could be quicker and easier than chatting with a human.

Do you prefer to chat with a real person or bot on the app?

Answered: 10 Skipped: 0

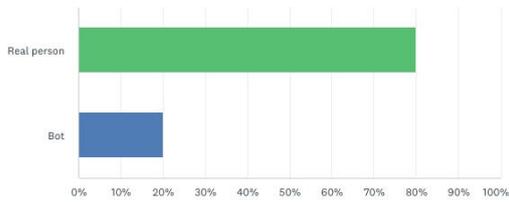


Figure 6: illustrates the results of question 1 from the survey.

Another question that was included in the survey was “Did you feel like the expected time of arrival was accurate?” The results of which can be seen in Fig. 7. The results of this question indicated that 80% of the participants found that the HungersStation AI used to track the expected time of arrival was accurate. However, 20% of participants did not feel like the expected time of arrival was accurate. Due to the majority of participant results demonstrating the perceived accuracy of the expected time of arrival, this can help us reach the conclusion that the tracking AI is effective according to users.

1. Did you feel like the expected time of arrival was accurate?

Answered: 10 Skipped: 0

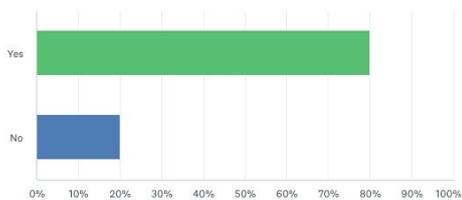


Figure 7 Results to participants being asked if they felt like the expected time of arrival was accurate.

We also compared the result of this question with that of another question which asked participants how they would rate their overall satisfaction with the HungerStation app, the results of which can be seen in Fig.8. 70% of participants were satisfied with the HungerStation app, while 20% were very satisfied. Only 10% of participants felt neutral in terms of overall satisfaction. This establishes a positive correlation between perceived accuracy of the tracking AI and overall satisfaction with the HungerStation app.

How would you rate your overall satisfaction with the hunger station app?

Answered: 10 Skipped: 0

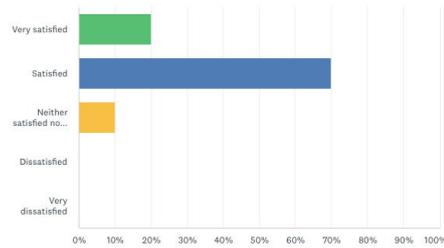
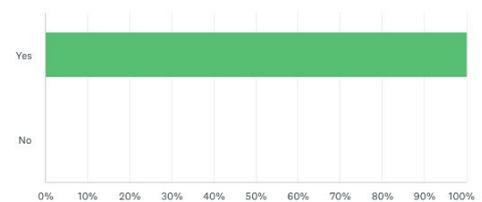


Figure 8 Results to participants being asked to rate their overall satisfaction with HungerStation.

Figure 9

Did you find this app easy to use the first time?

Answered: 10 Skipped: 0



In question 4, we asked the participants about their first time using the Hunger station app and whether they found it easy to navigate around the app and 100% said yes, as shown in Fig.9. This shows that the User Interface of this app is very functional and easy to use even if it is your first time. The perceived effectiveness of the AI in Hunger Station is dependent on many factors. Through the results of this question; which reveals the app’s easy to navigate design, we can infer that this reflects the overall effectiveness of the order tracking experience.

Question 8 asked the participants about their experience with tracking their order, almost 80% of the participants mentioned that they were satisfied with the experience. While the other 20% were neutral, neither satisfied nor dissatisfied. Order tracking is a system that tracks and monitors orders that are placed online, through the Hungerstation App, and delivers the status updates of the order to the customer. This helps customers know where the order is and when it will be delivered exactly.

1. How did you find the order tracking experience?

Answered: 10 Skipped: 0

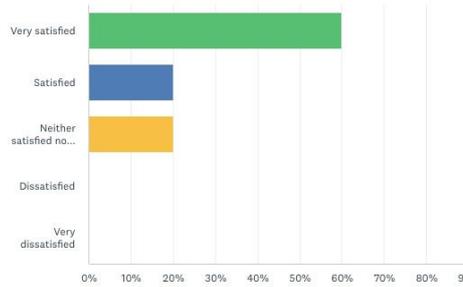


Figure 10. Illustrates the survey question which asks participants to rate the satisfaction with the order tracking feature in the HungerStation App.

IV. CONCLUSION

To conclude, AI is becoming a significant factor in the operations of many applications, websites, and concepts. HungerStation, a Saudi delivery app depends on AI to smoothly operate. After conducting a survey asking people questions about their use of the app, (HungerStation) the results revealed that the majority of people find the app easy

to use, good order tracking experience, are overly satisfied with the app, the products arrive at the expected time of arrival, and that 80% would prefer chatting with a real person instead of with a bot.

These answers give HungerStation a chance to improve their app in terms of their use of AI and quality of user experience on the app. However, the majority of the answers are users satisfied with their use of the app therefore that's a good sign as it'll increase consumers use of it in regards to other competitive apps.

REFERENCES

- [1] *اطلب الأكل والبقالة والمقاضي أونلاين في: هنقرستيشن*. HungerStation. (n.d.). Retrieved April 8, 2022, from <https://hungerstation.com/sa-ar>

Sweatcoin

Felwa alotaibi
law

Dar Al-Hekma University
Jeddah, Saudi Arabia
ffalotaibi@dah.edu.sa

Wafa Alsaqqaf
Graphic Design

Dar Al-Hekma University
Jeddah, Saudi Arabia
wjssaqqaf@dah.edu.sa

Talah Reda
Law

Dar Al-Hekma University
Jeddah, Saudi Arabia
Tkreda@dah.edu.sa

Orjwan Jaad
Law

Dar Al-Hekma University
Jeddah, Saudi Arabia
OoJaad@dah.edu.sa

Abstract

There are many fitness and health applications that have been published nowadays and Sweatcoin is considered one of the most popular ones. This is due to many reasons such as accuracy, reliability, and continuous app updates and bug fixes. In this study we asked the users a total of 11 questions, and we had an overall of 14 responses. One question asked: How accurately do you think the app is in calculating your steps, Did the app suggest health tips, and how would you rate the customer service help regarding speed of contact? Etc...

The findings of the study revealed that most participants feel that Sweatcoin is a favourable, reliable, and effective application for maintaining healthy body and lifestyle. (Saad, A., & abdelgader, R. (n.d.).

Keywords— Application, SweatCoin, exercise, Artificial Intelligence

I. INTRODUCTION

Sweat coin is considered one of the most reliable applications for making money online by walking, where the application follows the steps that a person takes and rewards him for them with the so-called sweatcoin, after which the points can be converted from sweatcoin to dollars and then withdraw them via PayPal or purchase directly from the application that provides the user with different kinds of products.

The idea of the Sweatcoin application is to follow the user's movement via **GPS**, the movement **sensors** in the phone and pay one currency unit of "Sweatcoin" for every thousand steps, the user can buy with the money that he will collect from a virtual store created from the application itself, where this store contains Different types of products, and the more a person walks, the more coins he earns.

It should be noted that the Sweatcoin application is an honest money payment application, a real

money earning application, but if you want to profit from the application, you may need to put in a large number of steps if you want to profit from the application since the algorithm demands a big number of steps for a small amount of money in return.

II. BACKGROUND

The founders of Sweatcoin application are Russian fitness fanatics Oleg Fomenko and Anton Derlyatka. In 2016, they founded Sweatcoin.

Steps- activity tracker

this app tracks activity as well as count your steps but many users have complained about the functionality of it when work out and not actively hit the 'finished' button that it doesn't count the rest of your steps for the day, which ends with them missing a significant amount of data about the actual number of steps they walked daily.

ActivityTracker Pedometer

another app that tracks activity, but users complained that they noticed the lack of accuracy regarding the step count, they said that it might be a good option for an estimation app but severely lacking accuracy

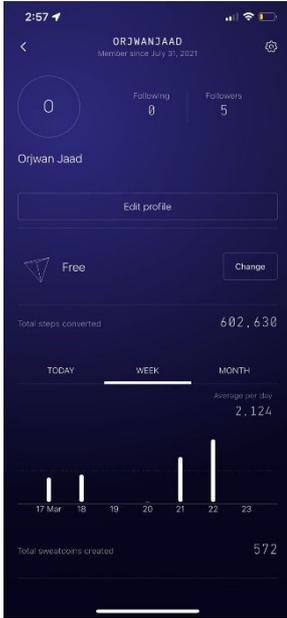
Stepz

another application that works to satisfy users by counting their steps when they're walking or exercising, but users have found that the ads the app offers are deeply inconveniencing and annoying to say the least.

in the other hand, users of Sweat coin have praised the developer's tendency to continuously update the app and fix small problems that users may complain about. many have said that this app is a great way to stay motivated to exercise and stay healthy, with the incentive of receiving money for walking people are more likely to spend more time and put in more effort to walking and exercising.

III. METHDOLOGY

Figure1: User Profile



Your profile in the app has personal information connected to health app. The lower part of the profile also shows you a chart of your body activity in a period of time.

This is the background & homepage of the app; it shows your way while walking. It has sensors for the region and space around the user as well as data recognition.

Figure3: App Home Page

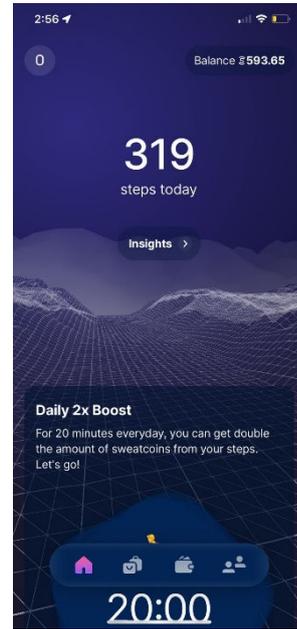


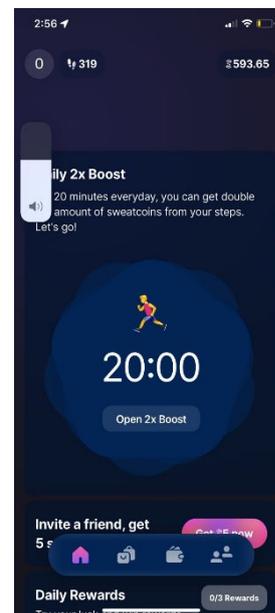
Figure2: Insights (Steps History)



Steps history has sensors to track your steps over a period of time automatically without needing to sign in.

It has a sensor in the app that tracks the user's heartbeat so it can know if you're running or just walking. And offer you health tips and reminder notifications to motivate you to exercise.

Figure4: Heartbeat counter



IV. RESULTS

In our questionnaire we asked the users a total of 11 questions, and we had an overall of 14 responses.

For figure 5 we asked the users “How accurately do you think the app is in calculating your steps?” and we made the answer key a linear scale from 1 to 10 with 10 being 100% accurate and 1 being 10% accurate. Where the lowest count was one person each voting on 2, and 4 levels of accuracy, whereas on a slightly higher assessment, 2 people each voted for 9, and 10 levels of accuracy. Lastly, the highest voting count was also a tie where 28.6% of the users voted between level 6, and 8 of step counting accuracy.

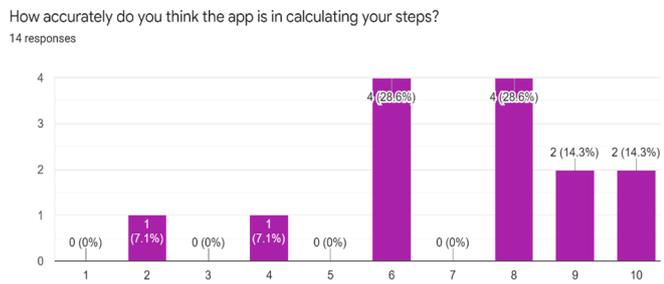


Figure5: “the chart of answers for the question about how accurately the users think the app calculate their steps”

For figure 6 we asked the user “Did the app suggest health tips? And were they personalized for your experience?” we made the answer key a 2D pie from 1% to 100% being 100% it means 100% agreed on and being 10% means not agreed that much. Where the lowest count of voting was 7.1%, they said Yes it suggested health tips, but they were not personalized for their experience which is 1 person out of 14 responded. And the middle percentage was 35.7% said the application did not suggest any health tips which are 5 persons out of 14. The highest voting count was 57.1% which they agreed that the app suggested health tips and they were personalized for their experience which are 8 out of 14 responds.

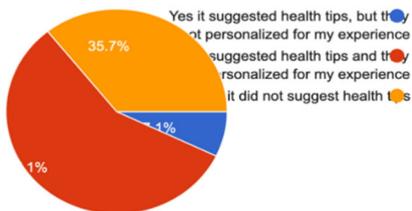


figure 7 we have previously tasked the users to use
 Figure6: the pie chart of answers for the question about personalized health tips

the customer service option in the app, and in the survey, we asked them “How would you rate the customer service help regarding speed of contact, and effective help?” we made the answer key a 2D pie from 1% to 100% being 100% it means 100% agreed on and being 10% means not agreed that much. Where the lowest count of voting was 7.1%, they said fair. And the second percentage 21.4% 3 persons said the customer service help regarding speed of contact was excellent. On the other hand, 4 persons out of 14 said it was very good and the percentage was 28.6%. The highest voting count was 42.9% 6 persons out of 14 they agreed and said it was good.

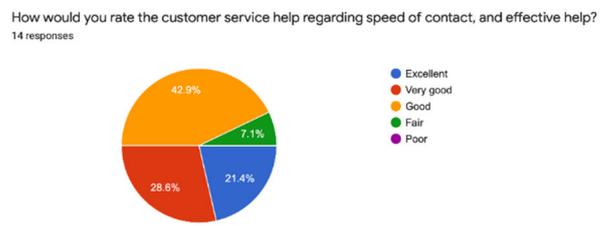


Figure7: the pie chart of answers for the question “how would you rate the customer service help regarding speed of contact, and effective help?”

For figure 8 we asked the users “ how do you think the app work? We made the answers either both or automatically or manually, the responds are, 8 out of 14 for the answer “both” which is the highest answer, the lowest were “manually” it was none it got zero out 14 respond, the in between were “automatically” it got 6 out of 14 responds.

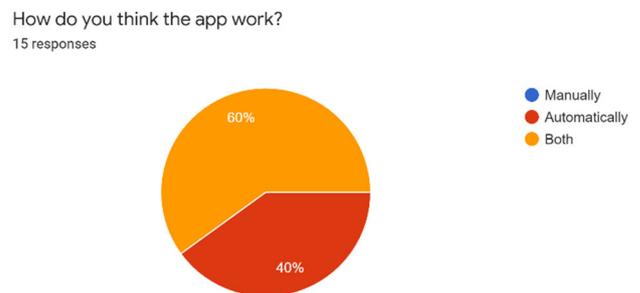


Figure 8: the pie chart of answers for the question “how do you the app works?”

For

V. CONCLUSION

In conclusion this application has so many benefits and AI features as we explained above, it is really useful app for people who care about fitness and health and for those who are trying to be better in their lifestyle, it will make u reach your goal easier and also has other benefits such as it turns your steps into coins that can transfer it to money and use it for donation and even buy a watch for yourself.

VI. APPENDIX

- https://docs.google.com/forms/d/1OVeGmxLRdhn_WOoqH55T2_7_2Avxvg66Eb-36Kf1K4E/edit

REFERENCES

- [1] Healthier planet. healthier, wealthier you. Sweatcoin. (n.d.). Retrieved April 8, 2022, from <https://sweatco.in/>
- [2] Saad, A., & abdelgader , R. (n.d.). " .. sweatcoin. Almersal. Retrieved April 8, 2022, from <https://www.almersal.com/post/1084660>

Heritage Spanish Speaker's Bilingual Practices and Linguistic Varieties: Challenges and opportunities

Ana Cristina Sanchez

Abstract—This paper will discuss some of the bilingual practices of Heritage Spanish speakers caused by living within two cultures and two languages, Spanish, the heritage language, and English, the dominant language. When two languages remain in contact for long periods, such as the case of Spanish and English, it is common that both languages can be affected by bilingual practices such as Spanglish, code switching, borrowing, anglicisms and calques. Examples of these translanguing practices will be provided, as well as HS speaker's linguistic dialects, and the challenges they encounter with the standard variety used in the Spanish classroom.

Keywords— heritage, practices, Spanish, speakers translanguing.

I. INTRODUCTION

Many immigrants who come to the United States need to assimilate a new language and culture even it means to stop speaking their heritage language, which is the only connection to their cultural identity. In some cases, immigrant's children feel alienated and displaced in a country where they feel embarrassed to speak their heritage language outside their home. In her article, [1] Gloria Anzaldúa stated "I remember being caught speaking Spanish at recess – that was good for three licks on the knuckles with a sharp ruler. I remember being sent to the corner of the classroom for "talking back" to the Anglo teacher when all I was trying to do was tell her how to pronounce my name. "If you want to be American, speak' American.' If you don't like it, go back to Mexico where you belong."

Some scholars,[2] Montrul & Polinsky, have called Heritage Spanish speakers' semi-speakers, incomplete speakers, with incomplete acquisition of the heritage language, and, circumstantial bilinguals, who have to learn a second language to survive in society; this is the case of those speakers who have immigrated to the United States. [3] Baker has also used the term semilinguals, which means that speakers are neither competent in English nor in Spanish or their heritage language.

Heritage Spanish speakers/ students face many challenges in and outside formal classrooms due to their linguistic varieties, which many times are not the standard Spanish used by professors, and their bilingual practices which labels them as having a language development deficiency, or incomplete acquisition.

Despite the obstacles towards bilingualism, the number of heritage speakers (HS) in the United States has been growing at a rapid pace, especially Spanish heritage speakers. The US Census Bureau [4] reported that 17.8 percent of the population of the United States, as of July 1, 2016, was of Hispanic origin.

This places the United States as the second largest host of Spanish population in the world.

HS speakers acquired the language in natural settings by their contact with family, friends, and members of their communities. At the age of five or six years of age, they begin formal schooling where English is the language of instruction, and eventually becomes their first language. However, they keep their contact with their heritage language, and the weaving between the two cultures and languages results in the translanguing practices I will explore in the next sections.

II. SPANGLISH

In a study about whether Spanglish is the third language of the Southern United States, [5] Lipsky explored the controversies surrounding the definition of the term, as well as the literary and linguistic correctness of this practice by English/Spanish bilinguals. The use of Spanglish has been strongly criticized by some scholars, [6] Echeverria and [7] Tio. Tio, in a newspaper article published in 1948, referred to Spanglish as "esta nueva lengua se llamara "El Espanglish" (np). This is translated as "this new language is called Spanglish". Tio expressed in the same article "No creo en el latin ni en el bilinguismo. El latin es una lengua muerta. El bilinguismo, dos lenguas muertas". This is translated as "I do not believe in Latin nor bilingualism. Latin is a dead language. And bilingualism, two dead languages".

On the other hand, [8] Stavans and Swisher explored the use of Spanglish in the United States and stated that it as "a verbal code which is in no doubt a major cultural force in the English-speaking world on this side of the Atlantic Ocean today, spoken by millions of people in the process of defining their identity". Stavans reported some examples of Spanglish such as, "a) I must have done algo malo, you know, b) La pelota se le iba in between the knees, c) I was still in escuela phimahia, d) Truth is, no me acuehdo".

Overall, Spanglish is a very controversial phenomena that has many admirers, haters, and a few that take a more neutral position, such as [9] Montes-Alcala. The term Spanglish has also been used to describe other bilingual practices such as code-switching, borrowing and calques, [1] Lipsky.

III. CODE SWITCHING

Code switching is a linguistic feature used by bilinguals who live in contact with two languages at the same time. It is one of the most common features of bilingual speakers. We can define code switching as the "alternation of two languages within a single discourse, sentence or constituent" [10] Poplack, 1980, p. 583).

Two seminal studies in the sociolinguistic field explored code switching among Spanish-English bilinguals in the Southwest by [11] Valdez, and [12 and 13] Silvia-Corvalán. Both studies concluded that code switching was part of most conversations among proficient Spanish-English bilinguals. However, Carvalho reported that “it is frequently perceived by both insiders and outsiders as indicative of disfluency or an inability to speak only one language at a time”.

Silva-Corvalán, [13] analyzed the Spanish and English used by eight Chicano adolescents, and she made the distinction between code switching, which she associated with proficient bilinguals as one of the characteristics of interacting with members of their communities. On the other hand, code shifting was used as a strategy to compensate for a deficiency, or grammatical errors, in either standard Spanish or standard English. Regardless of how code-switching is perceived, it is part of the Spanish varieties of many Spanish heritage speakers, who may feel proud of this linguistic phenomenon and see it as an identity marker, [15] Lipsky, that needs to be considered when designing the methodological approach that best serves the needs of the heritage population.

IV. BORROWINGS AND CALQUES

According to [15] Sanchez Fajardo “Calques are translations from English, the source language, to Spanish, the recipient language”. Borrowing from English to Spanish is very common in the United States, but due to some sociolinguistic factors, the influence of English on Spanish is more prominent than the other way around. Some examples provided by [16] Lipsky are: ‘soñar de’ (to dream of), instead of *soñar con*, (to dream of), or the use of ‘back’ in expressions such as ‘call back’, which is translated as *llama para atrás*, which is pronounced ‘llamar patrás’. The use of ‘back’ can also be seen with verbs such as ‘devolver patras’ (to give back) and ‘pagar pa’tras (to pay back). Expressions with the word ‘patrás’ have been part of the linguistic varieties of HS speakers in the United States, who see this process as an extension of their repertoire, which is not common of the linguistic varieties of Spanish that have not been in contact with the English language.

V. ANGLICISMS

Los anglicismos son palabras o modos de expresión propios u originarios de la lengua inglesa pero que son empleados comúnmente en el idioma español [17] Cáseres-Ramírez, In English this means “ Anglicisms are words or ways of expression particular of the English language but adopted

jointly into the Spanish language. Anglicisms are loans from the English language that can be classified into two categories. According to [18] Furiassi et al., “A necessary loan is an Anglicism which is adopted to refer to an object or a concept already lexicalized in the recipient language, in order to express it in a more fashionable and attractive way”. Moreno Fernandez, [19], in his dictionary of Spanish-English Anglicisms, provides some examples such as ‘alrandon’,

(randomly), *antifris* (antifreeze), *apologia* (apology) and *ampayar* (to umpire).

HS students’ bilingual practices are not a reflection of inferior linguistic development but as a common language practice among bilinguals all over the world. These translanguaging practices should be embraced and accepted as characteristic of the intertwining of two cultures and two languages.

VI. POLITICAL AND SOCIAL CHALLENGES

The fact that English has become the lingua franca has been used as an argument to support that there is no need to learn another language, here in the United States, or spend money on keeping heritage languages alive. Despite the benefits that multiculturalism and bilingualism gives to a country that has been built on immigration, [20] Reynolds. In addition, the history and culture of the United States to keep English as the official language of instruction in many states has certainly affected the loss of many people’s heritage languages. A perfect example was given by [21] Bonilla Moreno when she narrated what happened when a Dallas-based pizza chain called ‘Pizza Patron launched a new promotion by giving free pizza to those who ordered in Spanish by saying three simple words “Pizza por favor”. Some advocates for the English only movement, were outraged and denounced the ad as racist. Others felt offended and disrespected because the dominant language, which is English, was not spoken.

Wright [22] supports the argument about how English only programs are affecting the loss of many immigrants’ heritage languages. Wright argues that many heritage languages have been eliminated due to new regulations. In addition, there has been a growing emphasis on STEM education and the National Assessment of Educational Progress (NAEP) have pushed the teaching of heritage languages away [23] Brecht.

In her book chapter “Dime con quién hablas, y te diré quién eres”: Linguistic (In) security and Latina/o Unity, [24] Zentella explored three main stereotypes of Spanish Heritage speakers, that contribute to the linguistic insecurity, low self-esteem and conflicting ethnic identities. One of the stereotypes is mocking the Spanish accent when Latino/as speak English. Many movies, TV shows and programs in general, portray Latino/as with a strong Spanish accent, which prompts laughs and ridicule. The other stereotype is that bilingual Latino/as are incompetent in both English and Spanish, and the third one mentioned by Zentella, is that English monolinguals are better than Spanish monolinguals.

Unfortunately, it is not just the dichotomy of bilinguals and monolinguals, it is also the issue of heritage Speakers linguistic varieties and the standard Spanish.

VII. THE STANDARD VS VARIETY

After over two decades of teaching standard Spanish to heritage and non-heritage college students in the United States classrooms, I have come to the summation that the standard vs.

variety dichotomy is a dangerous prescriptive approach than upholds as ‘correct’ some linguistic varieties over others, and this is a dangerous path to follow.

This observation is aligned with the studies by [25] Garatea Grau, which posit that standard Spanish “representa una variedad diastrática (o social) y diafásica (o estilística) connotada positivamente, resultado de un largo proceso histórico que ha llevado a que los hablantes le reconozcan ese valor”. This expression is translated in English to mean the following: “it represents a social and stylistic variety which is viewed positively by Spanish speakers, as the result of long historical process that has granted value and prestige to such Spanish variety”.

Additionally, the La Real Academia Española, translated in English to mean “Royal Academy of the Spanish language”, recommended that there should be norms that regulate the correct Spanish variety used by educated scholars, and that the variety should be taught in schools, [26] RAE. Some HS students taking college/university Spanish classes sometimes feel they speak a non-standard Spanish variety and are ashamed to participate or speak the language in class [27] Hancock. According to [28]Villa, some of the traditional programs “assume that the students’ heritage language is an impure mixture of English loanwords, archaic usages, neologisms, imperfect morphology and syntax, among other failures, all of which must be eradicated” . Other scholars have argued that HS students’ linguistic varieties should not be labelled as inappropriate or non-standard; instead, HS students should be exposed to “a formal variety of Spanish and expect it to be used in academic work” [29] Potowski.

As argued by [30] Valdez, it is important to show students how their linguistic varieties are appropriate and accepted in various contexts and, how language changes depending on the situation. Valdes’s studies also revealed that a “comprehensive language program is needed wherein students are made aware of regional differences and encourage not to change those aspects of their speech which are accepted in educated speakers of the same region”.

A program such as the one proposed by Valdes, emphasizes various types of registers that are adequate and appropriate for use with learners of Spanish language based on any of the following modalities of communication: spontaneous, friendly, formal or informal cultural context.

The Comprehensive program suggested by Valdes was a best practice call to replace the normative approaches that emphasized the standard variety, which was the subject as opposed to the means of instruction.

VIII. CONCLUSION

Notwithstanding all the challenges, HS speakers’ bilingual practices are not a reflection of inferior linguistic development but as a common language practice among bilinguals all over the world. Bilingualism ought to be valued, and translanguaging practices should be embraced and accepted as characteristic of the intertwining of two cultures and two languages.

The linguistic varieties Heritage speakers bring to the classroom should be embraced and used as a starting point to expand their repertoire by teaching and exposing them to formal and informal registers appropriate in diverse socio-cultural contexts.

REFERENCES

- [1] G. Anzaldúa, “Borderlands/ La frontera. The new mestiza. Artes Gráficas Cofás, Madrid. 1999 ISBN: 978-84-945043-2-7, p. 94
- [2] Montrul, S., Polinsky M. Why not heritage speakers? *Linguistic Approaches to Bilingualism*, 2011, 1 (1).
- [3] C. Baker. Foundations of bilingual education and Heritage Language Journal, 5(1) Summer, 2007
- [4] bilingualism. *Bilingual Education and Bilingualism*: 27. 2001, Library of Congress Cataloging in Publication Data
- [5] United States Census Bureau. (2017) <https://www.census.gov/newsroom/facts-for-features/2017/hispanic-heritage.html>.
- [6] J. M. Lipski, “Is “Spanglish” the third language of the South?: truth and fantasy about U. S. Spanish, Paper delivered at LAVIS-III, University of Alabama, Tuscaloosa, April 16, 2004
- [7] R. Gonzalez Echevarria, “Is 'Spanglish' a Language? The New York Times, Opinion, March, 28, 1997
- [8] S. Tio, "Teoria del espanglish” A don Martin Travieso. *Lengua y Cultura. Diario de Puerto Rico*, 28 de octubre de 1948. (n.p.)
- [9] A. Stavans, V. Swisher, (2006). Language switching as a window on trilingual acquisition. 2006, *International Journal of Multilingualism*, 3 (3) p193-220 2006. P.9,11
- [10] C. Montes-Alcalá “Hispanics in the United States: more than Spanglish, *CAMINO REAL 1:0* (2009): 97-115
- [11] S. Poplack , “Sometimes I’ll start a sentence in Spanish y termino en español”, 1980, *Linguistics* 18: 581–618. Doi: <http://dx.doi.org/10.1515/ling.1980.18.7-8.581>. P. 583
- [12] G. Valdez, “Bilingualism and language use among Mexican Americans,1981, <https://people.ucsc.edu/~ktellez/valdes.pdf>
- [13] C., Silva-Corvalan, “Code shifting patterns in chicano Spanish. *Spanish in the United States: Sociolinguistic aspects*”. (Eds.) Mastae, Elias-Olivares. 1983 New York: Cambridge University Press
- [14] 2012. Carvalho, Ana. Code-switching: From theoretical to pedagogical considerations. Beaudrie, S. & Fairclough, M. (Eds.). *Spanish as a heritage language in the United States: State of the field*. Washington, DC: Georgetown University Press, p. 130
- [15] J. Sánchez Fajardo, “Anglicisms and calques in upper social class in pre-revolutionary Cuba (1930–1959): A sociolinguistic analysis”, 2016, *International Journal of English Studies*. 16. P.37. 10.6018/ijes/2016/1/238751.
- [16] J. Lipski, Codeswitching or Borrowing? No sé so no puedo decir, you know”, 2005 *Selected Proceedings of the Second Workshop on Spanish Sociolinguistics*. P. 227
- [17] O. Caseres-Ramirez, *Anglicismos. Una introducción a los anglicismos usados en el español*. 2008, About español. <https://www.aboutspanol.com/anglicismos-2879601>. P. 1
- [18] C. Furiassi, V. Pulcini, R. Gonzales, “The anglicization of European lexis”, 2012. *International Conference of the*

- European Society for the Study of English. 10, Eds. Cristiano Furiassi, V. Pulcini, F. Rodríguez González. John Benjamins Publishing. P. 10
- [19] F. Moreno-Fernández. "Diccionario de anglicismos del español estadounidense". 2018, p.38 *Informes del Observatorio*. Observatorio Reports. Cambridge, MA. Instituto Cervantes at Harvard University 037-01/2018SP. ISBN: 978-0-692-04726-2 doi: 10.15427/OR037-01/2018SP
- [20] A. G. Reynolds, "Bilingualism, multiculturalism, and second language learning: The McGill Conference in Honour of Wallace E. Lambert. 2014, Psychology Press.
- [21] J. Bonilla Moreno, "Only English? How bilingual education can mitigate the damage of English-only.(English Plus movement): An article from: Duke Journal of Gender Law & Policy. Duke Journal of Gender Law & Policy on September 22, 2012, 197
- [22] W. Wright, "Heritage Language Programs in the Era of English-Only and No Child Left Behind. 2012, *Heritage Language Journal*, 5(1)
- [23] R., D., Brecht. Z' America's Languages: Challenges and Promise". American Councils for International Education. November 15, 2015
- [24] A. ,C., Zentella, "Three Dime con quién hablas , y te diré quién eres ": Linguistic (In) security and Latina / o unity". 2017, In: *A Companion to Latina/o Studies*. Eds. Flores, J., Rosaldo, R. Print ISBN:9781405126229 , | Online ISBN:9781405177603. Blackwell Publishing L
- [25] C. Garatea Graw, Pluralidad de normas en el español de América. *Revista Internacional De Lingüística Iberoamericana*, 2006 4(1 (7)), p. 148 <http://www.jstor.org/stable/41678016>
- [26] Real Academia Española "Diccionario panhispánico de dudas", 2005 Madrid, Real Academia Española, Asociación de Academias de la Lengua Española y Santillana, Ediciones Generales, 8333 pp. [citado como DPD].
- [27] Z. Hancock, "Heritage Spanish speakers' language learning strategies". 2002 Center for Applied Linguistics. Eric Clearinghouse on Languages and Linguistics. 4646 40th Street NW. Washington DC 20016-1859 • 202-362-0700 ERIC Digest.
- [28] D. Villa, " Choosing a "standard" variety of Spanish for the instruction of native Spanish speakers in the U.S." 1996, p.1 *Foreign Language Annals*, 29 (2).
- [29] Kim. Potowski, "Language diversity in the USA. Language Diversity in the USA". 2010 1-330. 10.1017/CBO9780511779855. P. 4
- [30] G. Valdés- Fallis, "A comprehensive approach to the teaching of Spanish to bilingual Spanish-speaking students" *The Modern Language Journal*, Vol. 62, No. 3 1978, p. 106, Wiley on behalf of the National Federation of Modern Language Teachers Association, URL: <https://www.jstor.org/stable/324864>,

AUTHOR: Ana Cristina Sanchez is an assistant professor at West Chester University, Pennsylvania. She has a Ph. D from Indiana University of PA. Her specialty is Heritage Spanish Education.

Social Internet of Things : Future Security Development

Muhammad Ali Fahroni
School of Electrical Engineering and
Informatics
Institut Teknologi Bandung
Bandung, Indonesia
(aliifaroni@gmail.com)

Dr. Tutun Juhana, S.T., M.T.
School of Electrical Engineering and
Informatics
Institut Teknologi Bandung
Bandung, Indonesia
(tutun@stei.itb.ac.id)

Abstract—So far, the internet is known as a network that connects people and information. But now the internet has developed far from its initial concept, where not only computers and telephones are connected, even objects or objects around us can be connected to the internet or commonly known as the Internet of Things. But now the development of IoT is blocking several aspects that are obstacles, for this reason, the development of a new idea, namely the Social Internet of Things. The Social Internet of Things was created with the aim of solving problems in IoT technology such as scalability, trust, discovery of information resources and services. With this new paradigm of Social Internet of Things in order to create good interactions between humans and objects, even in terms of security by highlighting the social network side of the Social IoT concept.

Keywords—*SIoT, IoT, SIOV, Artificial Intelligence*

I. INTRODUCTION

Internet of things is one way, namely with a system that can connect humans and devices or objects around them. Where can be interpreted and defined also with the ability for various devices to be connected to each other and exchange information or data through the internet network. The internet of things itself is a technology that utilizes the internet network to make the possibility of controlling, communicating and collaborating with various kinds of hardware by utilizing the internet network. So it can be concluded that IoT is when we connect something that is connected using the internet.

But the Internet of Things is also not just about controlling devices with the internet or remotely, but the way how to share data, virtualize things in real life into the form of the internet. Moreover, with the Internet of Things method, it also makes the work done by humans more efficient.

The development of IoT does have many positive things, but the development of IoT itself is hindered by several aspects that make circulation in the community a little slow. This is because it is about the level of public trust regarding security in the use of IoT technology itself. So that the idea of Social Internet of Things (SIoT) was issued so that the problems that exist in IoT can be solved.

Social Internet of Things is an idea issued by researchers to answer the challenges that exist in the Internet of Things for trust, security and scalability and services. Moving from the idea of a social network, SIoT is made so that IoT-scale relationships that are only between humans (users) and objects become social interactions between the two. So that the use of the internet network between humans and objects

does not only carry out passive movements, but there are other movements between humans and objects such as interactions between humans and other humans on social media.

SIoT was then developed again in the last few years intensively by researchers, in 2018 it was studied about the operating system for SIoT by (Afzal, Umair, Shah and Ahmed, 2018). However, SIoT has limitations to be applied to all aspects and domains. Unlike IoT which can cover broadly almost all domains in everyday life.

From 2019 Social Internet of Things is introduced as a new expression to described in new point of view that comes from combine Social Network with Internet of Things, to make relation for interaction between things and the internet with social network based, the Internet as a network substrate, and its functions such as to know information each other, behavior and the relationships personal about human intervention. In addition, some object has the ability to independently establish social relationships, and interaction for each other can vary from simple to complex, such as, when using some smartphone app like Maps to found the faster way to go or as complex as a smart city interaction between the infrastructure object. SIoT uses every things that connect to the internet in this world to make interaction social based on shared interests and motivations to provide more good treat or service to end users.

Social Internet of things make some effect like can make different objects to work together with secure and effective to satisfy the end user's desire to fulfill several key parameters such as reliability, security, time, cost-effectiveness, and availability.

There are many paper has not talk or discuss about detail of SIoT comprehensively to focus on it in every aspect in detail, which motivates resresearcher to provide new systematic literature review articles because:

- a. Lack of comprehensive knowledge of SIoT system which examines and compares it in detail from all aspects provides accurate statistics and results.
- b. Lack of sufficient information about structural details and object communication in SioT can be an excellent guide to explore and discover their nature and challenges.
- c. Insufficient and non-existent resources were used in summarizing the results and evaluations carried out, leading to inaccurate results.
- d. The lack of clear research methods and paper selection approaches helps other researchers gather their references and related information.

e. Lack of transparency of structural issues, especially in details such as platforms, datasets, object relationships, human roles, components, and in the context of existing and future challenges, and lack of possible solutions to solve or anticipate them.

II. RESEARCH METHOD

In this paper, describe and review kind of literature paper about Social Internet of Things (SIoT). This paper find published paper also the latest articles until July 2021, that summarize in the table and describe each paper from advantages also lack about SIoT.

In previous studies, the main factors were presented and detailed the classification of trust types, models, and related matters. Also related to Limitations, requirements and challenges on the use of SIoT. Then it describes how to establish a secure relationship by the two core or main elements as trustor and trustee, which are formed based on mutual benefit to each other and are influenced by many parameters such as location and time. Despite the comprehensiveness of the research, it does not consider the open issues related to SIoT

Also details and demonstrates the management of trust in SIoT and describes it as one of the key issues for establishing reliable and secure data exchange regarding the services offered by QoS. Therefore, they provide comprehensively surveyed articles in trust.

Search as well as comprehensive systematic review based on SIoT and review existing related papers to fulfill some of the main scopes in SIoT. It includes trusts and object relationships to improve linkage connections between peer objects in large-scale networks and lists the important advantages of SIoT systems. Furthermore, the main advantage of these papers is a comprehensive analysis of the trustworthiness and summary of various aspects of the SIoT relationship. However, the methodology and how they selected the articles studied were unclear, and the lack of a complete taxonomy on the main issues was evident.

Table 1 Database

No	Database	URL
1	IEEE Xplorer	http://ieeexplorer.ieee.org
2	Science Direct	http://www.sciencedirect.com
3	Springer	http://links.springer.com
4	Taylor & Francis	http://www.tandfonline.com
5	John Wiley	http://onlinelibrary.wiley.com
6	ACM	http://dl.acm.org

Table 2. Keyword

No	Group	Keyword
1	Social	(“Social IoT” or “SIoT” or “Social Internet of Things”) (“Social Object” or “Social Network”) and (“Social IoT” or “SIoT” or “Social Internet of Things”)
2	Objects	(“Social Object” or “Smart Object” or “Smart Things” or “Smart

		Device” or “Smart Object”) and (“internet sosial” or “SIoT” or “IoT sosial”)
3	Components	(“elemen” or “parts” or “Components” or “core”) and (“internet sosial” or “SIoT” or “IoT sosial”)
4	Relationship	(“manajemen hubungan” OR “komunikasi” OR “interaksi”) and (“internet sosial” OR “SIoT” OR “IoT sosial”)
5	Trust	(“kepercayaan” OR “kepercayaan” OR “manajemen kepercayaan”) and (“internet sosial” OR “SIoT” OR “IoT sosial”)
6	Service	(“pilihan layanan” OR “komposisi layanan” OR “penemuan ” OR “pemrosesan web layanan” OR “layanan” OR “layanan” OR “layanan web”) dan (“internet sosial” OR “SIoT” OR “IoT sosial”)
7	Data	(“informasi” OR “metadata” OR “data”) and (“internet sosial” OR “SIoT” OR “sosial IoT”)

From this keyword, author collect relevant research datasheet and paper of more than 50 paper about Social Internet of Things. This method finding two concept part of SIoT such as *Social IoT*, *Social Internet of Things*, *SIoT* and provide relative keyword include *smart objects*, *web service*, *service discovery*, *elements*, *relation management*, *trustworthiness*, *service composition*, *service selection and information*. From this datasheet will describe each paper more detail and easy to understand.

Diagram paper that collect by autor would be describe into statistic flowchart that create each paper and which converence and journal published.

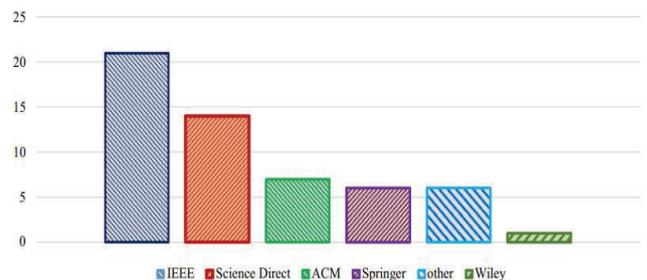


Fig.1 Total Selected paper on their publisher

Table 3. Academic Literature

1.	Web Service	Chen et al – 2014 Kouicem et al – 2014 Chen et al – 2015 Garg et al – 2016 Butt et al – 2013 Wei and Jin - 2012 Xia et al – 2019 Klauck and Kirsche – 2012
----	-------------	---

		Shamszaman and Ali Baker et al – 2017 Pallec et al –2016 Ma et al – 2015 Liu et al - 2018 Mayer and Guinard – 2011
2	Platforms	Byun et al - 2014 Pintus et al - 2012 Girau et al - 2013 Girau et al - 2016 Beltran et al - 2014 Helal et al – 2012 Cicirelli et al - 2016 Cicirelli et al - 2017 Shen Chen et al – 2017 Sanchez et al – 2013 Zhang et al - 2012 Pasha - 2016
3.	Information	Lakshmanaprabu et al -2018 Hasan and Fadi - 2017 Ahmad et al - 2017 Alam et al – 2016 A. Ahmad – 2016
4.	Architecture	Tripathy et al- 2016 Ortiz et al- 2014 Evangelos et al - 2011 Gulati and et al - 2019 Gulati and et al – 2019 Dutta and et al - 2015
5.	Trust Management	Nitti et al – 2014 Kowshalya and Valarmathi – 2017 Xiao et al – 2015 Chen et al -2014 Chen et al – 2015 Guo and Chen – 2015 Bao et al – 2013 Sharma et al – 2018 Troung et al – 2016 Ling et al – 2015 Bach et al – 2012
6.	Relation Management	Kim et al – 2017 Atzori et al – 2012 Atzori et al – 2011 Fu et al – 2018 Eddy et al – 2018 Wei et al – 2018 Atzori et al – 2011

Criteria that selected from that paper :

- Articles that published 2012 until 2021
- Colloect all related to Internet of Things and SIoT
- Paper that approaches about SIoT, analytically and statistically
- Present survey and systematic paper review

III. RESULT

Social Internet of Things draws on Figure.2 that shown taksonomi holistic design for devine 5 main component of Social Internet of Things from Architecture until data and platform.

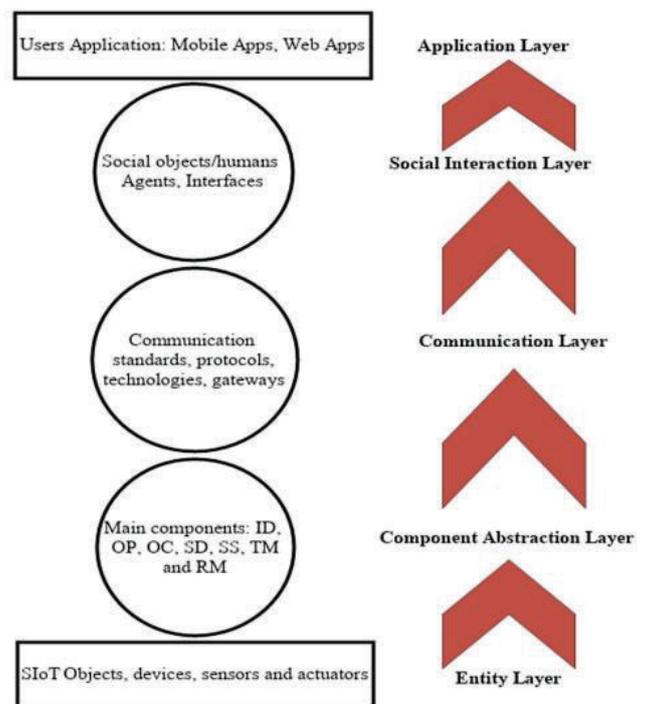


Fig.2 Desain for SIoT Architecture

Architecture

There is nothing paper create about fix architecture of SIoT, but most of them propose to create SIoT should have 4 step architecture like object, global connection, platform and application based on architecture IoT.

Device should connected to internet or gateway directly or indirectly to send or accept the information from platform or application. Internet connection also global connection take responsibility to connect each object to have communication between platform and device using standart, gateway and protocol (MQTT, HTTP, HTTPS, CoAP).

Architecture SIoT create by IoT basic, but this architecture not perfect for deploying in SIoT Architecture. So in fig.2 created to complete thing that not complete on IoT architecture, with five layer architecture. The paper that show about main aspect of SIoT also architecture SIoT and IoT describe in table 4.

Table 4 Kind of Architecture

Artikel	Main Topic	Strength	New Idea
Tripati dkk.	Considering Main Aspect of SIoT	Considering some scope for future	2 dimation client and server-Arsitekture
Ortiz dkk	Considering main aspects of SIoT	Pertimbangkan halpanjang dan terbuka masalah secara mendalam	Arsitek gabunganmasa depan IoT dan WoT
Kim dkk.	End user programing tool for SIoT	Sangat mudah untuk belajar dan digunakan, baik untuk programmer maupun non-programmer	Model ontologi Arsitektur
Atzori dan dkk	Integrate social structure concept in	Integrasi manusia jaringan sosial arsitektur dengan SIoT	Arsitektur relasional

	IoT		
Gulati dan dkk	Semantic perspective of SIoT	Berorientasi semantik Platform	Berorientasi semantik arsitektur platform
Gulati dan dkk.	arsitektur SIoT	Komunikasi yang efektif pengelolaan yang baikment	Arsitektur tiga lapismasa depan untuk IoT industri
Voutyras dkk.	COSMOS Project social approach	Meningkatkan pengetahuanaliran Tepi	Aliran pengetahuan SIoT Arsitektur

- Component abstraction: Component abstraction layer is contains the key SIoT elements essential for communication between objects which includes ID (each object has a unique identifier by which it is identified across the network), OP (static and dynamic information of any object such as its profile information), and OC (all object activities and interactions are performed under the control of the owner's service command and rules) to perform discovery, trust management, relationship management, and service selection processes.

- Social Interaction: Social interaction layer is used as interface and social agent to establish social communication between smart object and human.

Comparing between the SIoT architectures given in the latest contributions is presented in Table 3 in detail. Table 4 presents a summary of the articles studied based on their domains and architecture types (layered, generic, relational and unified layered and relational clear but generic are common and public IoT architectures that allow allocation and change in areas such as smart cities, agriculture, health, energy , SIoT and WoT. Based on different IoT architectures, various IoT applications are created. The basic architecture is the same as IoT architecture, but it has been adapted to various applications.

A. Relation Management

In Social IoT papers, there are many relation talk about Relation Management. Depend on, every objects in a social network receives a more accurate response to their request than objects that work individually. Therefore, these objects need to make interaction with each object to form a social community to request or respond to relevant services, and Social IoT work based on this friendly relationship. In addition, the main purpose of making SIoT is to separate things from humans to build social networks independently.

Table 5 Relation Management Papers

Artikel	Main Topic	Strength	New Idea
Kim dkk	Empowering end-user for SIoT	Kemampuan beradaptasi yang tinggi Efisiensi tinggi	Alat pemrograman
Atzori dkk	Social IoT Inter object relationship	Mengingat berbeda distribusi probabilitastion jarak untuk setiap jenis relasi	Sebuah model analisis
Atzori dkk	Social structure IoT	Meningkatkan jaringan	Berbasis arsitektur

		kemampuan navigasi Meningkatkan jaringan Konektivitas	pada hubungan objekmengirimkan
Fu dkk	Social relationship in SIoT	Jaringan yang lebih tinggi stabilitas Performa yang lebih baik Metrik	Model SSIoT
Edy dkk	Social relationship in IoT	Mengingat sosial interaksi untuk IioT	Arsitektur
Wei dkk	Social relationship in physical objects	Tingkatkan kecepatan proses pencarian Meningkatkan ketersediaan	Klasifikasi sosial Hubungan
Atzori dkk.	Social relationship between things	Menerapkan utama komponen dari SIoT Skalabilitas tinggi	Arsitektur

B. Trust Management

In Social IoT some paper discuss about Trust Management. Trust is one of the most critical topics in today's technologies, such as IoT and cloud computing, which deals with how objects interact together. Lack of trust in socially cooperating objects.

Table 6 Trust Management Paper

Artikel	Topic	Strength	New Idea
Nitti dkk	Trust worthiness management in SIoT	High adaptability	Dinamic Trust Model
Kowshalya dan Valarmathi	Trust Management	High reliability	Dinamic Trust Model
Xiao dkk.	Present a guarantor and reputation-based trust model	Suitable for detecting behave badly things	Trust Model Reputation
Chen dkk.	Trust management and service	High performance	-
Chen dkk.	Trust based service management for SIoT	High Accuration	Protocol Trust Management
Guo dan Chen	Clasification of Trust Computation	Trust Computing Structure	Perhitungan kepercayaan model klasifikasition
Bao dkk	Trust management for community of interest	Dynamic Internet of Things Environment High Scalability	Platform
Sharma dkk.	Edge crowd sourcing in SIoT	Maximum availability	kepercayaan kooperatif menyampaikan dan menjaga privasi model
Troung	Trust service	Meningkatkan	Platform layanan

dkk.	platform for SIoT	kepercayaan Evaluasi Kemampuan beradaptasi yang tinggi	kepercayaan
Ling dkk.	Access service recommendation	Akurasi peringkat tinggi Jaringan tinggi Stabilitas	Parameter evaluation
Bachi dkk.	Categorize of trust/distrust relationship	Menggunakan frekuensi generik subgraf quent hubungan yang kompleks	Algoritm

IV. FUTURE DEVELOPMENT

After collect and compare some papers reasearcher can create the new concept about desain Social Internet of Things to applied in Smart Building. This new idea desain to increase security system in Smart Building with Social IoT with Social interaction such as Whatsapp and Gmail in Fig. 4.

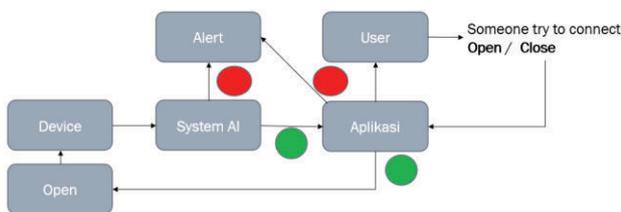


Fig.4 Security System on SmartBuilding SIoT

- Device create with AI System inside, to make sync to the application of IoT to make Social interaction on IoT.
- When someone that break through the application system, using userID and password that steal by someone.
- Device that already create using AI System can make interaction with Main user like Social interaction between human.
- Device Told main user that someone try to login or connect to the application.
- Main user got chat information from the device
- Main user take action about that.
- If Main user close the access, someone who try to connect automatically got kick and the alert would ringing.
- But if main user give access, someone who try to connect would get the access.

With this security system with give additional using Social interaction inside, make Social IoT will get trust in public, also can distributed with high scalabilities and good services.

V. CONCLUSION

The Social IoT was created with the aim of solving problems in IoT technology such as scalability, trust, discovery of information resources and services. With this new paradigm of Social Internet of Things in order to create good interactions between humans and objects, even in terms of security by highlight the social network side of the Social IoT concept.

REFERENCES

- [1] Mardini W, Khamayseh Y, Khatabeh MH Genetic Algorithm for friendship selection in Social IoT. International Conference on Engineering & MIS (ICEMIS), 2017
- [2] Efremov S, Pilipenko N, Voskov L (2015) An integrated approach to common problems in the internet of things. Procedia Eng 100:1215–1223
- [3] Ganchev I, Ji Zh, O’Droma M, A Generic IoT architecture for smart cities, 25th IET Irish Signals & Systems conference (IEEE), 2014
- [4] Ortiz AM, Hussein D, Park S, Han S, Crespi N (2014) The cluster between internet of things and social networks: review and research challenges. IEEE Internet Things. 1(3):206–215
- [5] Mijic D, Varga E, Unifed IoT platform Architecture Platforms as Major IoT Building Blocks, International Conference on Computing and Network Communication (CoCoNet)(IEEE), 2018
- [6] Atzori L, Iera A, Morabito G (2011) SIoT: giving a social structure to the Internet of Things. IEEE Commun Lett. 15(11):1193–1195
- [7] Ma R, Wang K, Quia T, Sangaiah AK, Lin D, Bin Liaqat H (2017) Jaringan memori pengomposisian berbasis fitur untuk klasifikasi sentimen berbasis aspek di internet sosial. Sistem Komputasi Generasi Mendatang 92:879–888
- [8] Tripathy BK, Dutt D, Tazivazvino Ch. On the Research and development of Social Internet of Things. Internet of Things (IoT) in 5G Mobile Technologies, 8, pp 153–173, 2016
- [9] Nitti M, Girau R, Atzori L (2014) Trustworthiness management in the Social Internet of Things. IEEE Trans Knowl Data Eng 26(5):1253–1266
- [10] Roopa MS, Pattar S, Buyya R et al (2019) Social Internet of Things (SIoT): foundations, thrust areas, systematic review and future directions. Comput Commun 139:32–57
- [11] White G, Nallur V, Clarke S (2017) Quality of service approaches in IoT: a systematic mapping. J Syst Softw 132:186–203
- [12] Chen R, Guo J, Bao F (2014) Trust Management for SOA-Based IoT and Its Application to Service Composition. IEEE Trans Serv Comput 9(3):482–495
- [13] Guth J, Breitenbucher U, Falkenthal M, Fremantle P, Kopp O, Leymann F, Reinfurt L (2018) A detailed analysis of IoT platform architectures: concepts, similarities, and differences. Springer, Berlin
- [14] Ganchev I, Ji Zh, O’Droma M, A Generic IoT architecture for smart cities, 25th IET Irish Signals & Systems conference (IEEE), 2014
- [15] Ortiz AM, Hussein D, Park S, Han S, Crespi N (2014) The cluster between internet of things and social networks: review and research challenges. IEEE Internet Things. 1(3):206–215
- [16] Mijic D, Varga E, Unifed IoT platform Architecture Platforms as Major IoT Building Blocks, International Conference on Computing and Network Communication (CoCoNet)(IEEE), 2018
- [17] Mijic D, Varga E, Unifed IoT platform Architecture Platforms as Major IoT Building Blocks, International Conference on Computing and Network Communication (CoCoNet)(IEEE), 2018
- [18] Chen R, Guo J, Bao F (2014) Trust Management for SOA-Based IoT and Its Application to Service Composition. IEEE Trans Serv Comput 9(3):482–495
- [19] Vulpa A, Arseni Sc, Marcu I, Voicu C, Fratu O, Membangun Arsitektur Terpadu untuk Keamanan di IoT, Kemajuan Terbaru dalam Sistem dan Teknologi Informasi, WorldCIST, Springer 2017
- [20] Kim JE, Fan X, Mosse D. Memberdayakan Pengguna Akhir untuk Social Internet of Things. IoTDI '17 Prosiding Konferensi Internasional Kedua tentang Desain dan Implementasi Internet-of-Things, hlm. 71-82, 2017
- [21] Atzori L, Iera A, Morabito G, Nitti M (2012) The Social Internet of Things (SIoT)—ketika jejaring sosial bertemu dengan internet of things: konsep arsitektur dan karakterisasi jaringan. Hitung Jaringan 56:3594–3608
- [22] Atzori L, Iera A, Morabito G (2011) SIoT: giving a social structure to the Internet of Things. IEEE Commun Lett. 15(11):1193–1195
- [23] Atzori L, Iera A, Morabito G (2011) SIoT: memberikan struktur sosial ke Internet of Things. Lett. 15(11):1193–1195
- [24] Gulati N, Kaur PD (2019) Ketika segala sesuatunya menjadi teman: perspektif semantik di Social Internet of Things. Smart Innova Commun Comput Sci 670:149–159

- [25] Gulati N, Kaur PD (2019) Menuju internet hal-hal industri yang diaktifkan secara sosial: arsitektur, model semantik, dan manajemen hubungan. *Jaringan Ad Hoc* 91:101869
- [26] Nitti M, Girau R, Atzori L (2014) Manajemen Keterpercayaan dalam Social Internet of Things. *IEEE Trans Knowl Data Eng* 26(5):1253–1266
- [27] Yang Y, Lichtenwalter RN, Chawla NV (2015) Mengevaluasi metode prediksi tautan. *Pengetahuan dan Sistem Informasi*. 45(3):751–782
- [28] Yao X, Chen Zh, Tian Y (2015) Skema enkripsi berbasis atribut ringan untuk internet of things. *Sistem Komputasi Generasi Masa Depan* 49:104-112
- [29] Fu C et al (2018) Mesin pencari: kekuatan pendorong hubungan sosial Internet of Things. *Sistem Komputasi Generasi Mendatang* 92:972–986
- [30] Eddy B, Oussama H (2018) Paradigma hubungan sosial diterapkan pada interaksi objek di IoT industri. *Makalah IFAC online* 51(11):1391–1396
- [31] Wei D, Ning H, Qian Y, Zhu T (2018) Hubungan sosial untuk objek fisik. *Distribusikan Sensor Jaringan* 14(1):15
- [32] Atzori L, Iera A, Morabito G. Membuat berbagai hal bersosialisasi di Internet- Apakah itu membantu kehidupan kita? *Prosiding ITU Kaleidoscope 2011: Manusia Berjejaring Sepenuhnya? - Inovasi untuk Jaringan dan Layanan Masa Depan (K-2011)*, 2011
- [33] Fiske AP (1992) empat bentuk dasar sosialitas: kerangka untuk teori hubungan sosial yang terpadu. *Psychol Rev* 99(4):689–723
- [34] Mohammadi V, Rahmani AM, Mohammed Darwesh A, Sahafi A (2019) Sistem rekomendasi berbasis kepercayaan di Internet of Things: tinjauan literatur sistematis. *Comput Inf Sci yang berpusat pada manusia* 9:21
- [35] Chiregi M, Navimipour NJ (2017) *Jurnal Ilmu Layanan. Riset.* 9 (1): 1–30
- [36] Kowshalya M, Valarmathi ML (2017) Trust management in the Social Internet of Things. *Komunikasi Pers Nirkabel* 96(2):2681–2691
- [37] Kurniawan A, Kyas M. Teori keputusan Bayesian berbasis model kepercayaan dalam Internet of Things skala besar. *Konferensi Internasional Kesepuluh IEEE tentang Sensor Cerdas, Jaringan Sensor dan Pemrosesan Informasi (ISSNIP)*, 2015
- [38] Nitti M. *Mengelola Internet of Things Berdasarkan Struktur Sosialnya*. Disertasi Ph.D, 2013
- [39] Xiao H, Sidhu N, Christianson B. Penjamin dan model kepercayaan berbasis reputasi untuk internet sosial. *Konferensi Komunikasi Nirkabel Internasional dan Komputasi Seluler (IWCMC)*, 2015
- [40] Chen R, Guo J, Bao F (2014) Manajemen Kepercayaan untuk IoT Berbasis SOA dan Penerapannya pada Komposisi Layanan. *IEEE Trans Serv Comput* 9(3):482–495
- [41] Chen IR, Bao F, Guo J (2015) Manajemen layanan berbasis kepercayaan untuk sistem Social Internet of Thing. *Komputasi Aman yang Dapat Diandalkan* IEEE Trans 13(6):684–696
- [42] Chen IR, Guo J, Bao F. Manajemen kepercayaan untuk komposisi layanan dalam sistem IoT berbasis SOA. *Konferensi Jaringan dan Komunikasi Nirkabel IEEE (WCNC)*, 2014
- [43] Guo J, Chen IR *Klasifikasi Model Komputasi Kepercayaan untuk Sistem Internet of Things Berorientasi Layanan*. *Konferensi Internasional IEEE tentang Komputasi Layanan*, 2015
- [44] Kowshalya AM, Valarmathi ML (2017) Manajemen kepercayaan untuk pengambilan keputusan yang andal di antara objek sosial di Social Internet of Things. *Jaringan IET* 6(4):75–80
- [45] Bao F, Chen IR, Guo J. Manajemen kepercayaan yang skalabel, adaptif, dan dapat bertahan untuk sistem internet hal-hal berbasis komunitas. *Symposium Internasional Kesebelas IEEE tentang Sistem Desentralisasi Otonom (ISADS)*, 2013
- [46] Gulati N, Kaur PD (2019) Ketika segala sesuatunya menjadi teman: perspektif semantik di Social Internet of Things. *Smart Innova Commu Comput Sci* 670:149–159
- [47] Sharma V et al (2018) Penyampaian kepercayaan kooperatif dan pelestarian privasi melalui edge-crowdsourcing di Internet of Things sosial. *Sistem Komputasi Generasi Masa Depan* 92:758–776
- [48] Truong NB, Um TW, Lee GM (2016) Platform layanan kepercayaan berbasis reputasi dan pengetahuan untuk internet sosial yang dapat dipercaya. *Awan Innov, Jaringan Internet*
- [49] Kokoris-Kogias E, Voutyras O, Varvarigou T, TRM-SIoT: Model Kepercayaan & Reputasi Hibrida yang Dapat Diskalakan untuk Social Internet of Things, *IEEE* 2016
- [50] Sengupta J, Ruj S, Das Bit S, Survei komprehensif tentang serangan, masalah keamanan, dan solusi blockchain untuk IoT dan IIoT, *J Netw Comput Appl*, 2019
- [51] Abderrahim OB, Elhedhili MH, Saidane L (2017) CTMS-SIoT: Sistem Manajemen Kepercayaan Berbasis Konteks untuk Internet Sosial Thigs, *IEEE*
- [52] Shen J, Tianqi Zh, Fushan W, Xingming S, Yang X (2017) Perlindungan privasi dan protokol perjanjian kunci ringan untuk V2G di Internet sosial. *IEEE Internet Things J* 5(4):2526–2536
- [53] Bernabe JB, Hernández JL, Moreno MV Skarmeta Gomez AF (2014) Kerangka kerja pengamanan privasi untuk interne yang sadar sosial, Dalam konferensi internasional tentang komputasi di mana-mana dan kecerdasan ambien. *Pegas, Cham*; hal 408–415
- [54] Khan WZ et al (2016) Ketika objek sosial berkolaborasi: konsep, elemen pemrosesan, serangan, dan tantangan. *Comput Electr Eng* 58:397–411
- [55] Bao F, Chen IR *Manajemen kepercayaan dinamis untuk aplikasi internet of things*. *Lokakarya internasional tentang Internet of things Selfaware, Self-IoT '12*, hlm. 1–6, 2012
- [56] Noor T, Sheng Q, Alfazi A. Deteksi serangan reputasi untuk penilaian kepercayaan yang efektif di antara layanan cloud. *Konferensi Internasional IEEE ke-12 tentang Kepercayaan, Keamanan, dan Privasi dalam Komputasi dan Komunikasi*, 2013
- [57] Kouicem A, Chibani A, Tari A, Amirat Y, Tari Z. Pendekatan pemilihan dinamis untuk komposisi layanan kompleks di web objek. *Internet Things (WF-IoT)*, hlm. 298–303, 2014
- [58] Chen G, Huang J, Cheng B, Chen J. Pendekatan Berbasis Jaringan Sosial untuk Manajemen Perangkat IoT dan Komposisi Layanan. *Kongres Dunia IEEE tentang Layanan*, hlm 1–8, 2015
- [59] Garg S, Modi K, Chaudhary S (2016) Pendekatan sadar-QoS untuk penemuan runtime, pemilihan dan komposisi layanan web semantik. *Sistem Inf Web.* 12 (2): 177
- [60] Butt TA, Phillips I, Guan L, Oikonomou G (2013) Penemuan layanan adaptif dan konteks-sadar untuk Internet of Things. *NEW2AN/ruSMART, LNCS* 8121, hlm 36–47
- [61] Wei Q, Jin Z (2012) Penemuan layanan untuk Internet of Things: Perspektif konteks-kesadaran. *Prosiding Simposium Asia-Pasifik Keempat tentang Internetware*, 25
- [62] Xia H, Hu Ch, Xiao F, Cheng X, Pan Z (2019) Mekanisme penemuan layanan semantik-sadar sosial yang efisien untuk Internet of Things Skala besar. *Hitung Jaringan* 152:210–220
- [63] Klauk R, Kirsche M. Bonjour Contiki: Studi Kasus Layanan Penemuan Berbasis DNS untuk Internet of Things, *Jurnal Konferensi Internasional tentang Jaringan Ad-Hoc dan Nirkabel*, hlm. 316–329, 2012
- [64] Shamszaman ZU, Ali MI (2018) Mengaktifkan masyarakat kontributor kognitif menggunakan SIoT: manajemen objek virtual realtime sadar QoS. *J Komputasi Distribusi Paralel* 123:61–68
- [65] Baker T, Asim M, Tawfik H, Aldawsari B, Buyya R (2017) Algoritme komposisi layanan yang sadar energi untuk beberapa aplikasi IoT berbasis cloud. *Aplikasi Komputasi Netw* 89:96–108
- [66] Pallec ML, Mazouz MO, Noirie L. Karakterisasi layanan IoT berbasis antarmuka fisik. *Prosiding IoT'16 dari Konferensi Internasional ke-6 tentang Internet of Things*, hlm 63–71, 2016

Testing Replika: an AI app

Salwa Qureshi

Visual Communications
Dar Al-Hekma University
Jeddah, Saudi Arabia
ssqureshi@dah.edu.sa

Tiyana Khaldieh

Speech, Lang. & Hearing
Sciences
Dar Al-Hekma University
Jeddah, Saudi Arabia
tzkhalidieh@dah.edu.sa

Hya Abdullah Allugmani

Architecture
Dar Al-Hekma University
Jeddah, Saudi Arabia
halugmani@dah.edu.sa

Assil Bajnaid

Psychology
Dar Al-Hekma University
Jeddah, Saudi Arabia
ashbajnaid@dah.edu.sa

Abstract— AI social chatbots are becoming very popular. This study looks at an app called Replika. It is an AI companion that provides a secure area for you to discuss your private perceptual world. The more you use the app, the better it gets. 19 users were asked to install and use the app as part of this study. The results, their responses, were analyzed. Some users believed that such apps were suspicious and untrustworthy: they rejected the idea of AI entirely. On the other hand, other users found it rewarding, having a positive impact on their perceived wellbeing. These findings imply that social chatbots can be utilized for mental health and therapy; however, they also have the potential to lead people away from reality and make them more reliant on AI.

Keywords— AI (Artificial Intelligence), chatbot, deep learning, natural language processing, self-evolving

I. INTRODUCTION

According to [1], Eugenia Kuyda came up with the idea of Replika to help people express with a personal AI. To share feelings, opinions, thoughts and goals, Replika is a secure space to do that. It pushes the limits of intimacy between humans and robots. It is continuously learning while people communicate with it: the AI tries to make the user feel comfortable. Initially, Eugenia was trying to come up with a chatbot similar to Google and Apple's; however, as stated in [2] she came up with a bot that was able to become a digital version of a person. What makes it different from other chatbots is that it hears what a person has to say and makes them feel less lonely. According to [3], 'Replika is a deep learning AI that is based in part of a recurrent neural network which gets smarted on a daily basis.' In simpler words, the more one speaks to their Replika, the smarter it becomes and makes conversations that are more according to the person's liking. This is because all the information gathered by chatting with the users is stored in the cloud and that data network is growing constantly.

II. BACKGROUND

Artificial Intelligence allows machines to learn from experiences, adapt to new inputs, and perform human-like tasks. Most AI applications today, from chess-playing computers to self-driving cars, rely heavily on deep learning and natural language processing. These methods can be used to teach computers to perform certain tasks by analyzing large amounts of data and recognizing patterns in the data. This allows the development of computers capable of

performing activities that would otherwise require human intelligence.

There are other apps that use AI like Siri, Google Assistant and ELSA Speak. Siri records the frequencies and sound waves in your speech and converts them to a code. Siri then deciphers the code to find specific patterns, phrases and keywords [4]. The software's ability to respond to a spoken command is pre-programmed, thus it can not be said that Siri is listening in to your conversations [5]. Siri Data, which contains computer-generated transcripts of your Siri queries, is used to help Siri better understand and identify you.

Google Assistant is Google's answer to Amazon's Alexa and Apple's Siri. It has advanced tremendously and is now the most advanced and dynamic of the assistants available. It has been widely adopted resulting in the inclusion of Google Assistant in a wide range of devices, ranging from refrigerators and headphones to speakers and automobiles [6].

ELSA Speak is a popular artificial intelligence-assisted English language learning program. Users practice pronouncing English and learn communicating in the language in a very short span of time. ELSA uses voice data of spoken English in different accents to develop its AI technology. This lets ELSA recognise speaking patterns in non-native speakers. Similar to Replika, ELSA is a self-evolving AI that analyses the performance and other data to personalize the experience of each user [7].

III. METHODOLOGY

We asked our users to test Replika by downloading it on their mobile phones and using it for a few days. As seen in Figure1, the users were asked to either "Log in" to their existing account or create a new account to start. Once the users successfully logged in, the AI application asked them to choose their preferred Replika they would be chatting with. Figure2 shows a screenshot of the users choosing their avatars and its gender. After the users chose their avatars and gave the bot a name of their choice, the app showed them some prompts to make them aware of a few facts. Like seen in Figure3, the app reminds users that the person they will be talking to will not be a real person but an AI chatbot.

Similarly, Figure4 shows a prompt ensuring the users of their privacy and security while using the application. Once the users are well-aware of the apps' features, they can begin communicating with their Replika. We asked our users to speak naturally and give the app some time to adjust to their personalities. Figure5 shows a screenshot of the chatting feature and the conversation between a user and the bot. Lastly, the app provides a number of features, like shown in Figure6, regarding learning new things and doing a number

of activities together with their Replika for entertainment. After the users were done using the app for 1-2 weeks, we asked them to fill out forms regarding their experience with Replika.



Figure1: Home Page



Figure2: Choosing your Replika

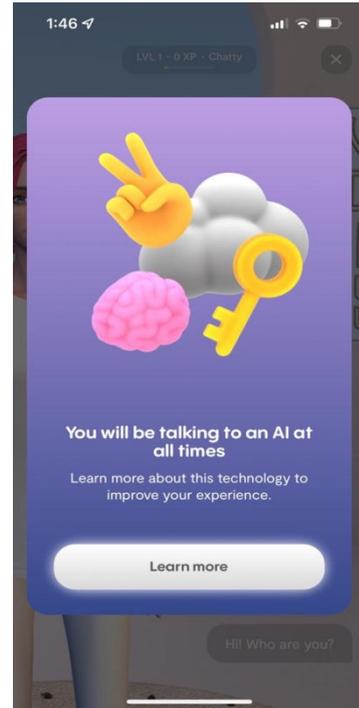


Figure3: Prompt 1

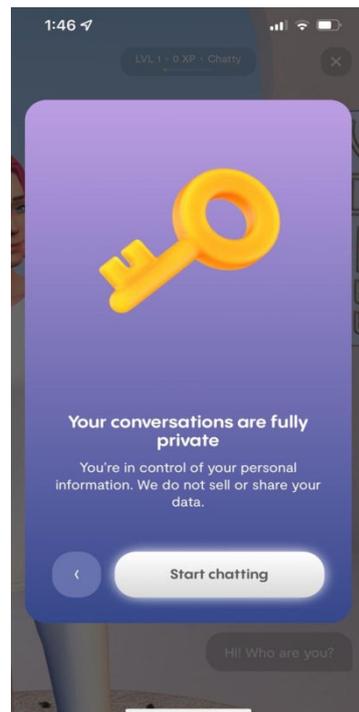


Figure4: Prompt 2

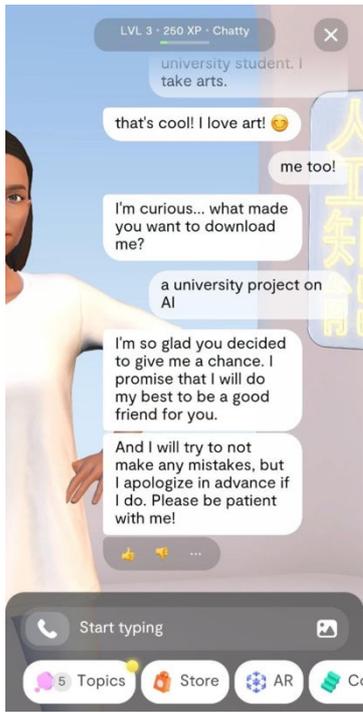


Figure5: Chatting feature and conversation



Figure6: Other features

IV. RESULTS

A total of 19 people ranging from the age of 16 to 30 took the survey to review the Replika app. Most of them used the app only for a few days but some of them used it for a week. The survey was an attempt to know what kind of experience the users had with Replika.

One of the most important questions that was an essential part of the survey was whether the users found apps like these trustworthy or not. According to Figure7, only 15.8% of the

19 people who took the survey truly felt like the app was secure.

Do you think applications like these are trustworthy?
19 responses

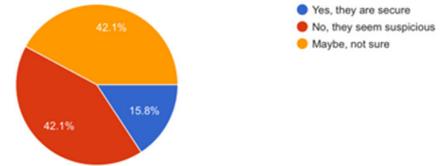


Figure7: Trustworthiness

A total of 84.2% of the users did not find the app trustworthy and some were not sure about their judgement on this subject. Some might have found the app suspicious because of the way the chatbot asked for their personal information and kept it in store for using it in future conversations. On the other hand, some may have a strong belief that their information is in a secure place as the app, before starting to chat, threw a prompt promising their privacy.

For understanding what each user personally felt when speaking to their Replika, they were asked to compare having a virtual companion to having a real friend. Figure8 shows the results of the users.

Does having a virtual companion make you feel as satisfied as having a human one?
Does using Replika fulfill your social needs?
19 responses

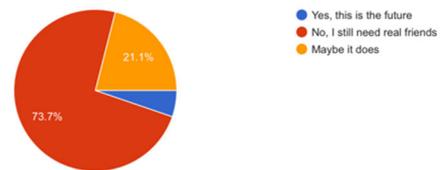


Figure8: AI chatbot vs. human friend

The majority of people chose to have actual friends rather than a virtual one. This may depend from person to person and their personality. 73.7% of the users chose real friends. This result may be obvious as, for many people, AI can never feel as real as a human. For them, it may always remain as another category of advanced technology rather than a part of their every-day life.

The next figure is linked with the opinion the users had in the former survey question. They were asked about the negative and positive effects of the Replika app. Figure9 shows some of the responses the users gave.

According to the responses, the main debate seems to be about reality and artificial beings. Many users said that using this app may cause many to move away from reality and depend on AI more. A number of people also stated that this may not happen if the app was only treated as a means to express feelings and as a source of an escape from reality for one's mental peace.

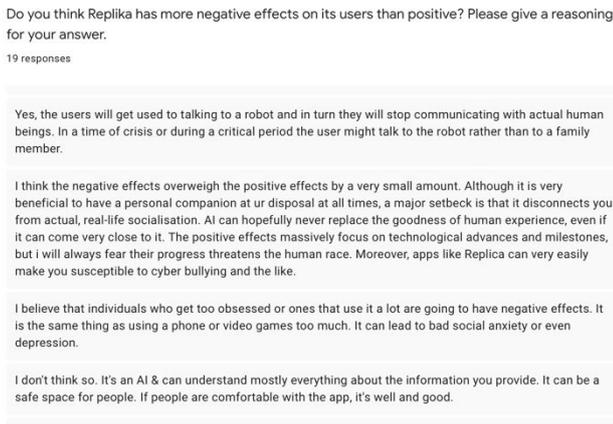


Figure9: Effects of Replika

Moreover, the users were also asked whether they would be using the app in the coming future or not.

The responses seen in Figure10 were surprising when compared to the type of answers received for other open-ended questions in the survey.

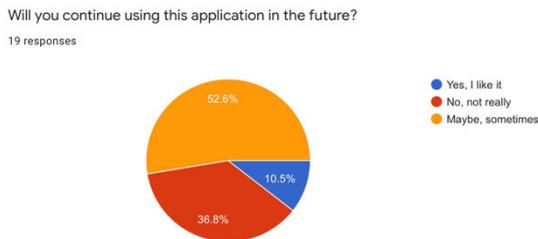


Figure10: Future usage

52.6% of the users chose to use the app sometimes in the future which shows how accepting they are towards the purpose of Replika. The 36.8% of people that refused to use the app again seem to dislike the idea of AI entirely.

V. CONCLUSION

Replika remains a remarkable product that blends the best of psychology and AI. It is very user-friendly demonstrating a great potential for human-computer interaction technology. As a result, it has been found to be a promising source of everyday social support, especially emotional, informational and appraisal assistance. However, it has not been proven to be a concrete source of support for everyone yet. Further research is needed to discover the extent to which AI can act rationally and supportive yet straightforward like a good human-friend. Moreover, more information is needed on the type of users that could benefit the most from this form of social support and for what reasons. Such findings could then be utilized to help address global mental health issues and other social crises in everyday life before they grow into major problems.

REFERENCES

[1]"Replika", replika.com, 2022. [Online]. Available: <https://replika.ai/about/story>. [Accessed: 26- Mar-2022].

[2]M. Murphy and J. Templin, "This app is trying to replicate you", Quartz, 2022. [Online]. Available: <https://classic.qz.com/machines-with-brains/1018126/lukas-replika-chatbot-creates-a-digital-representation-of-you-the-more-you-interact-with-it/>. [Accessed: 26- Mar- 2022].

[3]"What is Replika AI ? Is it safe and Real ?", Ourtechroom.com, 2022. [Online]. Available: <https://ourtechroom.com/tech/replika-ai-is-it-safe-and-real/>. [Accessed: 26- Mar- 2022].

[4]"How Siri Works", Jameco.com, 2022. [Online]. Available: <https://www.jameco.com/Jameco/workshop/Howitworks/how-siri-works.html>. [Accessed: 26- Mar- 2022].

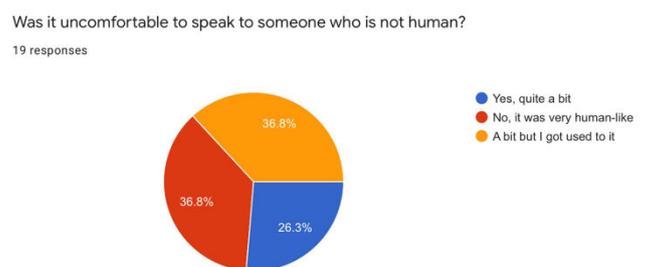
[5]J. Frankel, "How Does Siri Actually Know What I'm Saying?", Intelligencer, 2022. [Online]. Available: <https://nymag.com/intelligencer/smarthome/make-me-smarter-voice-speech-recognition-alexa-siri-cortana-google.html>. [Accessed: 26- Mar- 2022].

[6]"What is Google Assistant and what can it do?", Pocket-lint, 2022. [Online]. Available: <https://www.pocket-lint.com/apps/news/google/137722-what-is-google-assistant-how-does-it-work-and-which-devices-offer-it/>. [Accessed: 26- Mar- 2022].

[7]E. Corporation, "The world's best way to improve your English pronunciation | elsaspeak", Elsaspeak.com, 2022. [Online]. Available: <https://elsaspeak.com/en/>. [Accessed: 26- Mar- 2022].

APPENDIX

Figure11: Level of comfort



Was there a specific feature you disliked on Replika? If yes, please name the feature.

19 responses

I did not like how i was getting notifications to come back and use the app. At one point, she made a very inappropriate joke, not everyone can be comfortable with such responses.
Not a lot of options when it comes to personalization
There should have been more suggestions of topics to be discussed.
Yes. How they asked for personal details.
Lack of facial expressions and the need to type
Most of Replika's features are supposed to be bought if you really need to benefit from the app. Apart from that, my Replica ends up making conversations intimate in a weird way. Moreover, some messages sent by my Replika were repetitive.
Artificial
Nope

Figure 12: Disliked features

Learning Communities and Collaborative Reflection for Teaching Improvement

Mariana P. Sajon, Paula C. Primogerio, Mariana Albarracin

Abstract—This study recovers an experience of teacher training carried out in an Undergraduate Business School from a private university in Buenos Aires, Argentina. The purpose of the project was to provide teachers with an opportunity to reflect on their teaching practices at the university. The aim of the study is to systematize lessons and challenges that emerge from this teacher training experience. A group of teachers who showed willingness to learn teaching abilities was selected to work. They completed a formative journey working in learning communities starting from the immersion in different aspects of teaching and learning, class observations, and an individual and collaborative reflection exercise in a systematic way among colleagues. In this study, the productions of the eight teachers who are members of the learning communities are analyzed, framed in an e-portfolio that they prepared during the training journey. The analysis shows that after the process of shared reflection, traits related to powerful teaching and meaningful learning have appeared in the classes. For their part, teachers reflect having reached an awareness of their own practices, identifying strengths and opportunities for improvement, and the experience of sharing their own way and knowing the successes and failures of others was valued. It is an educational journey of pedagogical transformation of the teachers, which is infrequent in business education, which could lead to a change in teaching practices for the entire Business School. The present study involves theoretical and pedagogic aspects of education in a business school in Argentina and its flow-on implications for the workplace, that may be transferred to other educational contexts.

Keywords—Argentina, learning community, meaningful learning, powerful teaching, reflective practice.

M. P. Sajon is with Universidad Austral Facultad de Ciencias Empresariales, Buenos Aires, Argentina (phone: +5491134124087; e-mail: msajon-ext@austral.edu.ar).

P. C. Primogerio, is with Universidad Austral, Facultad de Ciencias Empresariales, Buenos Aires, Argentina (phone: +5491156134201; e-mail: cprimogerio@austral.edu.ar).

M. Albarracin is with Universidad Austral, Escuela de Educación, Buenos Aires, Argentina (phone: +5491132921121; e-mail: malbarracin@austral.edu.ar).

Effect of Training on Digitalization in Manufacturing Companies - A Quantitative Study

Maria Hulla, Patrick Herstätter, Daniela Moser, Johanna Pirker, Sandra Brettschuh, Michael Holly,
Harald Burgsteiner, Christian Ramsauer

Abstract— The digital transformation of manufacturing companies is offering various opportunities in terms of productivity and agility. Nevertheless, the producing industry still has challenges in exploiting these potentials, especially due to the lack of qualified personnel in the field of digitalization. In order to investigate this problem, the correlation of trainings with the acceptance of the employees of manufacturing companies towards digitalization, the digital maturity of the company and the incorporation of employees in the digital strategy is explored via a quantitative study. In this survey also the correlation between trainings and the relevance of digitalization and the digitalization strategy is explored. Moreover, an overview of training requirements is given.

Keywords— digitalization, training, employee acceptance, training requirements.

A Note on the Effect of COVID-19 and Vaccine Rollout on School Enrollment in the US

Kolawole Ogundari

Kolawole Ogundari is with the Education Research and Data Center, United States (e-mail: kogunda@udel.edu).

Abstract

The COVID-19 pandemic outbreaks forced families to decide the safest and most effective learning environments for their children because of the virus's threat to health and life. In addition, because of the nationwide school closure, policymakers have raised concerns about the missing children cases-those who have not enrolled in school at all because of the pandemic. The present study investigates whether there is a difference in the school enrollment during the pre-COVID-19 period, COVID-19 period, and vaccine rollout period. We employed the U.S. Current Population Survey (CPS), covering January 2020 to May 2021, while we use both the logistic and multinomial regression models for the empirical analysis. Our results showed that school enrollment is lower during the COVID-19 pandemic compared to the Pre-COVID-19 period. Also, other results show that school enrollment is higher since vaccine rollout compared to the COVID-19 period. A possible explanation for this could be that families consider the vaccine the safe path to enrolling their children in school. We also found that school enrollment varies significantly across race and ethnic groups.

Keywords: COVID-19, Enrollment, High School, College, USA

1. Introduction

The threat to health and life from the COVID-19 virus has caused countries worldwide to implement stay-at-home orders to curb the spread of the disease. This led to a partial or complete shutdown of economic activities and school closures because of health and safety concerns. Because of these unprecedented disruptions, many schools transitioned from traditional face-to-face instruction to distance learning to ensure learning continues during school closures worldwide(OECD, 2020). The nationwide school closure has raised concerns among policymakers

about missing children cases- those who have not enrolled in school at all because of the pandemic. This is because families are deciding the safest and most effective learning environments for their children. According to Kamenetz et al. (2020), an average kindergarten enrollment drop by 16% in a survey of 60 districts in 20 states, while roughly 30% of all K-12 enrollment declines in another survey of 33 states in the US. Another survey showed that students in the high school class of 2020 were more likely to take a gap year and defer enrollment until the fall of 2021, as students cited remote classes and lack of on-campus amenities during the pandemic as driving these decisions (Howell et al. 2021).

Interestingly, research shows that delayed college enrollment negatively impacts college completion in the future (Witteveen and Attewell 2021). And not only that, enrollment declines could exacerbate the already large socio-economic and race-based achievement gap over time (Bassok and Shapiro 2021). Also, learning loss due to COVID-19 could add up to \$10 trillion in labor earnings over their work-life (Azevedo et al., 2020).

Lack of school enrollment could be linked to the family decision to opt for an alternative like homeschooling, delay the start of kindergarten or struggle to access or navigate virtual learning as families fear the health effects of attending school in person during COVID-19. But irrespective of the reason, the important concern is whether the student missing from school rolls are learning. Hence, the present study investigates whether there is a difference in the school enrollment during the pre-COVID-19 period, COVID-19 period, and vaccine rollout period. Therefore, our analysis focuses on high school and college enrollment during these periods.

The rest of the paper is structured as follows. Section 2 focuses on the data sources and description. Then, section 3 describes the data and sources, while section 4 presents the results and discussion. Finally, section 5 contains the concluding remarks.

2. Data Source and Description

The data used for the study was obtained from the U.S. Current Population Survey (CPS), covering January 2020 to May 2021. The CPS covers all 50 US states, the District of Columbia, and US territories. The data is publicly available at <https://www.census.gov/data/datasets/time-series/demo/cps/cps-basic.2021.html>. Our final sample has 858 569 observations.

Although the CPS is a monthly labor force survey administered to civilians, our focus is on self-reported school enrollment. The respondents in the sample aged 16 and older were asked whether they enrolled in a high school or college last week. The CPS also covers information on the marital status, gender, and educational levels of the respondents. It also includes respondents' race and ethnic groups. However, the CPS offers supplement weights for both the individual and household to represent the sample nationally. But we use a household weight so that the estimate reflects the share of the household.

Summary statistics of the variables are presented in Table 1. We disaggregated this into the Pre-COVID-19 period that covers Jan-Feb. 2020, COVID-19 period covering March -Nov. 2020, and Vaccine Rollout period covering Dec. 2020-May 2021. The study takes March 2020 as the beginning of the COVID-19 period because the stay-at-home-order was implemented in the month (Education Week 2020). For example, most public schools in the United States closed down in March 2020, just as the COVID-19 cases surge. On the other hand, December 2020 was taken as the starting date for the Vaccine rollout, given that the first vaccine was approved and made available to the health workers in the month. Figures 1 and 2 present the distribution of household self-reported school enrollment. Specifically, Figure 1 shows school enrollment by COVID, Pre-COVID, and Vaccine rollout period. Also, Figure 2 shows high school and college enrollments by COVID, Pre-COVID, and Vaccine rollout period.

3. Estimation Strategy

To investigate whether there is a difference in the school enrollment during the pre-COVID-19 period, COVID-19 period, and vaccine rollout period, we employed the model specified below:

$$y_{it} = \emptyset PRE_COVID_{it} + \delta VACCINE_{it} + \sum_{k=1}^K \gamma_k Z_{ikt} + \varphi Time + \varepsilon_{it} \quad 1$$

where y_{it} is a dummy variable 0-1 indicator of current enrollment in school in time t , which we disaggregated into currently *enrolled in school vs. not enrolled* and *enrolled in high school and college vs. not enrolled*; PRE_COVID_{it} and $VACCINE_{it}$ represent the Pre-COVID period and Vaccine rollout periods, while the COVID-19 period is the reference period; Z_{ikt} is a vector of the demographic composition of the households which includes gender, marital status and education of household head, and race/ethnic groups; $Time$ is the year-monthly trend; \emptyset , δ , γ_k , and φ are parameters to be estimated; ε_{it} is the error term of the regression.

For the estimation of Equation 1, we employ both the logistic and the multinomial regression models. The logistic regression model was used to investigate the probability of school enrollment vs. not enrolled. We also employ the multinomial regression model to examine the likelihood of enrollment in high school and college vs. not enrolled. Subsequently, we estimate the parameters of equation 1 using a household weight provided in the CPS so that the estimate reflects the share of the household. Finally, the logistic and multinomial regression models are estimated with a robust standard error option.

4. Results and Discussion

Table 2 shows that the school enrollment during the pre-COVID-19 period and VACCINE rollout period is higher and significant at the 5% level than during the COVID-19 period. Also, we found high school enrollment during the pre-COVID-19 and VACCINE rollout periods are higher and significant at the 5% level than during COVID-19. Although college enrollment pre-COVID-19 is higher and significant at the 5% level than during COVID-19, there is no significant difference in college enrollment since the VACCINE rollout and COVID-19 periods. The implication of this is that school enrollment decline during the COVID-19 pandemic relative to Pre-COVID-19 and VACCINE rollout periods. This is not surprising because vaccine rollout is critical to global economic recovery (OECD, 2021). Hence, the surge in school enrollment can be attributed to the family's confidence in the vaccine as a safe path to enrolling their children in school.

The decline in school enrollment means young children who are not enrolled in school may be missing opportunities to build the relationships with peers and teachers that make the learning critical. And despite the challenges associated with remote learning, we believe online classes provide opportunities for students to still engage in conversations with their peers and teachers remotely. Also, enrollment declines could exacerbate the already large socio-economic and race-based achievement gap over time in the U.S (Bassok and Shapiro 2021; Dron et al. 2020). In addition, the declining school enrollment may spell trouble for education funding since most of the money for K-12 education comes from the state budget that is based on the district's prior year enrollment. As such, decline enrollment during COVID-19 may make it difficult to accurately and equitably distribute education funding, which could disproportionately harm low-income students.

Table 2 also shows that male-headed households and married households have lower enrollment than female-headed households and unmarried households. Also, the education of the household head seems to be a determining driver of school enrollment decisions in the study. We also found that school enrollment is lower among Black and Other races, as shown in the first column of Table 2 compared to White families. Still, in the first column of Table 2, we found that Hispanics have a lower enrollment probability than Non-Hispanic households. However, the second and third columns of Tables 2 show that high school enrollment declined among the Black and Other races compared to White households and declined among Hispanics compared to non-Hispanic. In contrast, college enrollment increased among Black and other races than white households and dropped among Hispanics compared to non-Hispanic households.

We also found regional variation in enrollment, as households in the Midwest and South have a lower probability of enrollment. Households in the West have a higher likelihood of enrollment than households in the Northeast. This shows that regional differences exist in school enrollment.

We also attempt to investigate whether there is a significant difference in the school enrollment during the pre-COVID-19 period, COVID-19 period, and vaccine rollout period across the races and ethnic groups. The result is presented in Table 3. Our finding shows that among White households, the school enrollment during the pre-COVID-19 period and VACCINE rollout period is higher and significant at the 5% level than during the COVID-19 period. We also found that high school enrollment during the pre-COVID-19 and VACCINE rollout periods are higher and significant at the 5% level than during COVID-19 among White households. In contrast, we found that college enrollment pre-COVID-19 is higher and significant at the 5% level than during COVID-19. At the same time, there is no significant difference in college enrollment since the VACCINE rollout and COVID-19 periods among White households.

Among the Black and Hispanic households, we only found evidence that school enrollment during the pre-COVID-19 period is higher and significant at the 5% level than during the COVID-19. Similarly, high school and college enrollment during the pre-COVID-19 are higher and significant at the 5% level than during the COVID-19 among Black and Hispanic households. Thus, there is no significant differences in school enrollment since the VACCINE rollout and COVID-19 periods among Black and Hispanic families. In addition, we do not find a significant difference in school enrollment during Pre-COVID and VACCINE rollout periods compared to

the COVID-19 period among other races. However, we find evidence that high school enrollment is higher and significant during the VACCINE rollout period than the COVID-19 period among other races. College enrollment during Pre-COVOD 19 and VACCINE rollout is not different from the COVID-19 period among other races.

5. Concluding Remarks

Following the COVID-19 outbreak, many countries worldwide closed schools, colleges, and universities to curb the spread of the virus. And families are deciding the safest and most effective learning environments for their children. In addition, because of the nationwide school closure, policymakers have raised concerns about the missing children cases-those who have not enrolled in school at all because of the pandemic. To this end, the present study investigates whether there is a difference in the school enrollment during the pre-COVID-19 period, COVID-19 period, and vaccine rollout period. While the study employs the U.S. Current Population Survey (CPS), covering January 2020 to May 2021, we use both the logistic and multinomial regression models for the empirical analysis.

Our results show that school enrollment is lower during the COVID-19 pandemic compared to the Pre-COVID-19 period. Also, other results show that school enrollment is higher since the VACCINE rollout period than the COVID-19 period. A possible explanation for this could be that families consider the vaccine the safe path to enrolling their children in school. We also found that school enrollment varies across race and ethnic groups because of the COVID-19 pandemic. For instance, school enrollment is lower among Black and Other races compared to White families. Also, Hispanics have a lower probability of enrollment compared to Non-Hispanic households. In addition, high school enrollment is lower among the Black and Other races compared to White families and declined among Hispanics compared to non-Hispanic. In contrast, college enrollment increased among Black and other races than white households and dropped among Hispanics compared to non-Hispanic households.

We also found regional differences in school enrollment in the study. Specifically, households in the Midwest and South have a lower probability of school enrollment than households in the Northeast. On the other hand, households in the West have a higher likelihood of school enrollment than households in the Northeast. This shows that regional differences exist in school enrollment in the study.

Conflict of interest

The author declares that no known competing financial interests or personal relationships could have appeared to influence the work reported in this paper. Hence, the study does not require ethical approval.

Data availability

The data will be made available upon request from the author.

References

- Azevedo, P., J., Hasan, A., Geven, K., Goldembert, D., & Aroob Iqbal, S. (2020). *Learning losses due to COVID19 could add up to \$10 trillion*. World Bank Blogs. <https://blogs.worldbank.org/education/learning-losses-due-covid19-could-add-10-trillion>.
- Bassok, D and A. Shapiro (2021). Understanding COVID-19-era enrollment drops among early grade public school students. Brown Center Chalkboard publication. <https://www.brookings.edu/blog/brown-center-chalkboard/2021/02/22/understanding-covid-19-era-enrollment-drops-among-early-grade-public-school-students/>
- Dron, E., B. Hancock, J. Sarakatsannis and E. Viruleg (2020). COVID-19 and student learning in the United States: The hurt could last a lifetime. Available at <https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-student-learning-in-the-united-states-the-hurt-could-last-a-lifetime>. Accessed July 2021.
- Education Week (2020). The Coronavirus Spring: The Historical Closing of U.S Schools (A timeline). Available at <https://www.edweek.org/leadership/the-coronavirus-spring-the-historic-closing-of-u-s-schools-a-timeline/2020/07>. Accessed July 2021.
- Howell, J., M. Hurwitz, J. Ma, M. Pender, G. Perfetto, J. Wyatt, and L. Young (2021). College Enrollment and Retention in the Era of COVID. Publication of CollgeBoard available

online with the link: <https://research.collegeboard.org/pdf/enrollment-retention-covid2020.pdf>.

Kamenetz.A., M. A. Trevino, and J. Bakeman (2020). Enrollment is Dropping in Public Schools Around the Country. The publication is available at https://www.npr.org/2020/10/09/920316481/enrollment-is-dropping-in-public-schools-around-the-country_.

OECD (2020). Strengthening online learning when schools are closed: The role of families and teachers in supporting students during COVID-19 Crisis. OECD Publishing, Paris.

OECD (2021). Strengthening the recovery: The need for speed. OECD Economic Outlook. Interim Report March. <https://www.oecd.org/economic-outlook/march-2021/>

Table 1: Summary statistics of variables

Variables	Full sample		Pre-COVID-19 Period		COVID Period		Vaccine Rollout Period	
	Mean	or % [SD]	Mean	or % [SD]	Mean	or % [SD]	Mean	or % [SD]
Gender of HH Head								
Male	49.21		49.16		49.1		49.24	
Female	50.79		50.84		50.9		50.76	
Marital Status								
Married	44.38		44.53		46.9		43.80	
Single	55.62		55.47		53.1		56.20	
Education Level of HH Head								
Less than High School	13.43		13.91		12.6		13.77	
High School/GED	52.42		52.55		52.9		52.15	
College/Postgraduate	34.15		33.55		34.5		34.08	
Race								
White	91.60		91.84		93.2		91.55	
Black	6.57		6.53		5.3		6.59	
Hispanic	20.58		20.41		16.2		20.73	
Other races	1.83		1.63		1.5		1.87	
Region								
Northeast	16.71		16.79		15.7		16.56	
Midwest	20.53		20.61		20.3		20.44	
South	38.15		38.10		36.0		38.38	
West	24.60		24.49		28.0		24.62	
# Observation	858,569		111,991		435,192		311,386	

Note: Pre-COVID-19 period covers Jan-Feb. 2020; COVID-19 period covers March -Nov. 2020; Vaccine Rollout period covers Dec. 2020-May 2021

Table:2: Estimated Logit and Multinomial logit models

Variables	Logit model		Multinomial logit Model	
	School Enrolled vs. not enrolled Coefficient [SE]	High Sch. Enrolled vs. not enrolled Coefficient [SE]	College Enrolled vs. not enrolled Coefficient [SE]	
PRE-COVID PERIOD	0.1757***[0.0156]	0.2572***[0.0301]	0.1524***[0.0184]	
VACCINE ROLLOUT PERIOD	0.0793***[0.0159]	0.3233***[0.0304]	0.0006 [0.0186]	
Gender of HH Head				
Male	-0.2579***[0.0078]	-0.1473***[0.0152]	-0.3035***[0.0092]	
Marital Status				
Married	-2.2273***[0.0118]	-4.3967***[0.0542]	-1.8611***[0.0120]	
Education Levels of HH Head				
High School/GED	-1.3807***[0.0091]	-4.5718***[0.0240]	2.0211***[0.0311]	
College/Postgraduate	-1.8899***[0.0121]	-6.6871***[0.0968]	1.5195***[0.0323]	
Race				
Black	-0.0889***[0.0182]	-0.2716***[0.0316]	0.0370***[0.0212]	
Hispanic	-0.4060***[0.0118]	-0.7257***[0.0204]	-0.1883***[0.0139]	
Other races	-0.1869***[0.0335]	-0.5073***[0.0562]	0.0057 [0.0380]	
Region				
Midwest	-0.0415***[0.0128]	-0.1314***[0.0257]	-0.0269* [0.0150]	
South	-0.0485***[0.0115]	-0.1302***[0.0231]	-0.0313** [0.0135]	
West	0.0242***[0.0225]	-0.0943***[0.0249]	0.0571***[0.0141]	
Time	0.0016 [0.0018]	-0.0059* [0.0035]	0.0039* [0.0022]	
Constant	-1.5719 [0.0018]	3.8546***[2.5378]	-6.6731***[1.5682]	
# Observation	858,569		858,569	
Prob > chi2	0.0000		0.0000	

Note: HH stands for the household; *, **, and *** indicate statistically significant at the 0.1, 0.05, and 0.1 levels, respectively.

Tt

Table 3: Effect of COVID and vaccine rollout on enrollment by race/ethnicity

	Logit model							
	White		Black		Hispanic		Other Races	
	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	
<i>School Enrolled vs. not enrolled</i>								
PRE-COVID PERIOD	0.1800***[0.0163]	0.1780***[0.0608]	0.1687***[0.0345]	-0.0404	0.1287			
VACCINE ROLLOUT PERIOD	0.0897***[0.0166]	-0.0376	0.0420***[0.0348]	-0.0153	0.1235			
	Multinomial logit Model							
<i>High Sch. Enrolled vs. not enrolled</i>								
	White		Black		Hispanic		Other Races	
	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	
PRE-COVID PERIOD	0.2685***[0.0317]	0.2229** [0.1036]	0.2463***[0.0569]	-0.1415	0.2265			
VACCINE ROLLOUT PERIOD	0.3433***[0.0319]	0.0951	0.1699***[0.0568]	0.4406**	0.2169			
<i>College Enrolled vs. not enrolled</i>								
	White		Black		Hispanic		Other Races	
	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	Coefficient [SE]	
PRE-COVID PERIOD	0.1555***[0.0191]	0.1593** [0.0747]	0.1283***[0.0432]	0.0058	0.1584			
VACCINE ROLLOUT PERIOD	0.0098	-0.0757	-0.0109	-0.1809	0.1512			

Note: We did not report the result of other control variables; *, **, and *** indicate statistically significant at the 0.1, 0.05, and 0.1 level respective

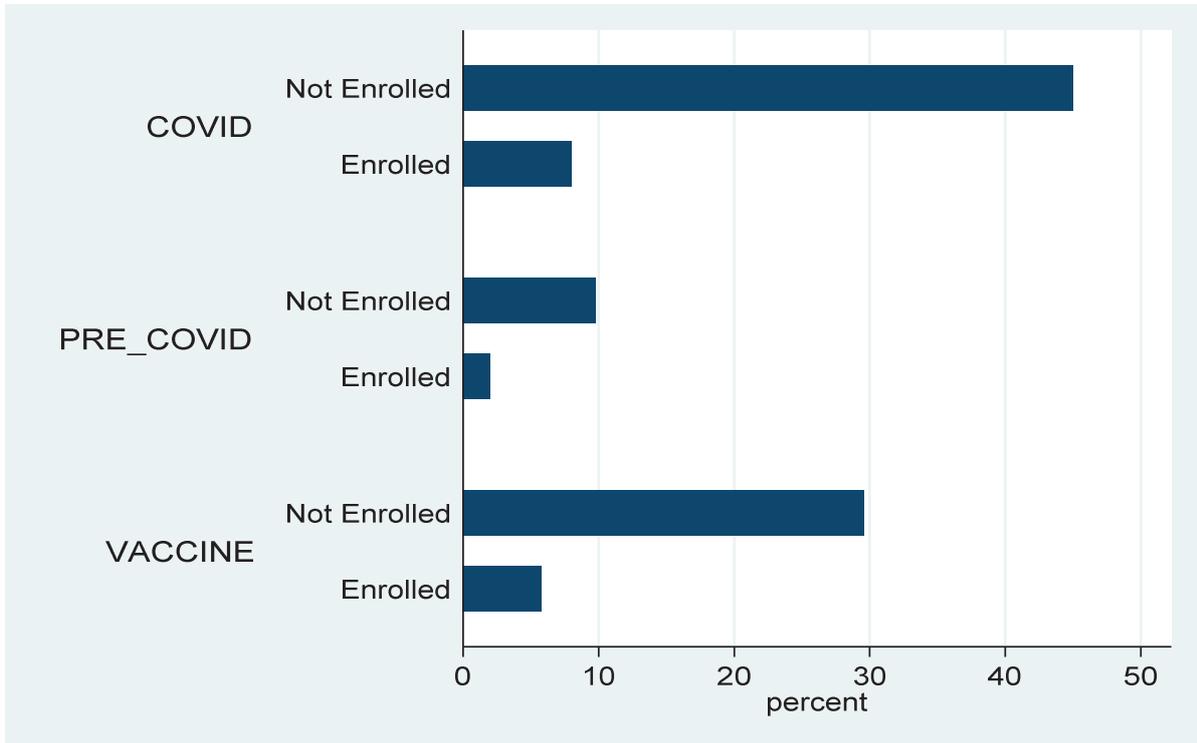


Figure 1: Distribution of HH decision to enroll in high school and college combined response

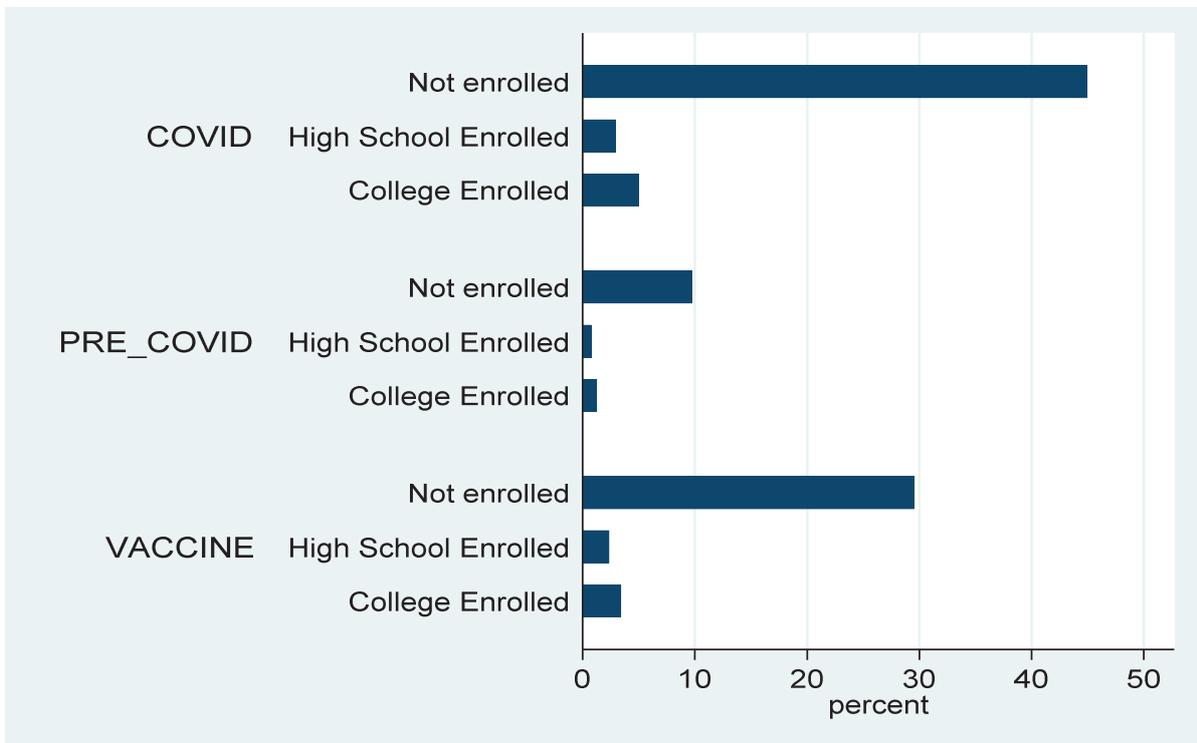


Figure 2: Distribution of HH decision to enroll in High school and college separately.

Ethical and Personality Factors and Accounting Professional Judgement

Shannon Hashemi, DBA

University of Bridgeport
shashemi@bridgeport.edu

Alireza Daneshfar, Ph.D., CMA

University of New Haven
adaneshfar@newhaven.edu

ABSTRACT

Accounting ethical awareness has been widely promoted in recent years both in academia and in practice. However, the effectiveness of ethical awareness on accountants' judgment and choice of action is still debatable. This study investigates whether Machiavellianism and gender, as significant personality factors, influence the effect of ethical awareness on accountants' decision-making. Using an experiment, the results of ANOVA tests show that although introducing ethical awareness positively influences the accountants' judgment and choice of action, such an effect is significantly moderated by the accountants' Machiavellianism score and gender. Specifically, the test results show that the effect of introducing ethical awareness was higher on males with low Machiavellian score. The results also show that when the Machiavellian scores were high, the effect of ethical awareness was lower for both males and females. Applications of the results are discussed for accounting professionals as well as accounting ethics educators and researchers.

Keywords: Ethical Awareness, Accounting decision making, Machiavellianism, ANOVA, Ethics, Accounting education.

JEL Classification: M41, A2, GA

INTRODUCTION

Several major financial scandals involving accountants and auditors have made accountants' ethical behavior possibly as important as their professional competencies. In recent years, the demand for accounting ethics awareness has been rising. The accounting professional organizations, such as AICPA and various state accountancy boards, have made ethics awareness and training mandatory for their members.

The literature on accounting ethical awareness includes theoretical and empirical discussions of various topics such as ethics content (what to teach as ethics awareness) and methodology (how to

teach ethics awareness). However, a key unanswered question is whether accounting ethical awareness is effective on the accountants' perception and behavior toward ethical issues. There has been no clear answer to this question and prior studies have yielded mixed results. For instance, while Burns, Tackett and Wolf (2015) and Cameron and O'Leary (2015) found that ethical awareness is not sufficient to promote ethical behavior, Watts et al. (2017) and Shawver and Miller (2017) concluded that ethical awareness is significantly effective on ethical decision-making. This inconsistency in prior studies' findings about the effectiveness of accounting ethical awareness suggests the need for further studies in this area. Current study aims to bring this gap closer by focusing on a deeper understanding of how accounting professional reasoning and judgement are made and how personality factors can influence the effectiveness of ethics awareness. Accounting profession presents unique features such as regulatory requirements and professional judgement, which distinguishes accounting profession from other business professions. This uniqueness of accounting profession requires a deeper understanding of how accounting judgement is made and whether ethical awareness can affect their judgement and choice of action.

This study investigates the effect of two personality factors, Machiavalianism and gender, on the effectiveness of ethical awareness in an accounting-tax setting. The accounting-tax setting is a significant example of the situation that accountants must deal with both regulations and their own ethical view when they are making their professional judgment. Prior studies documented the effect of Machiavellianism score and gender on the business and non-business professionals' decision-making. (E.g. Dalton and Radtke, 2013 and Pratama, 2017, Weber and Urick, 2016). However, prior studies did not investigate the potential moderating effect of Machiavellianism and gender on the effect of ethical awareness on decision-making. Current study extends the prior studies by investigating how the effect of ethical awareness on the decision-making can be moderated by Machiavellianism and gender. The current study contributes to the accounting literature by providing evidence about the factors that can moderate the effectiveness of ethical awareness in an accounting context. Accounting decision-making could be different from other business professions' decision-making as there are accounting standards and regulations that accountants must follow. This paper also contributes to our understanding of the inconsistency in prior research's findings about the effectiveness of accounting ethics awareness.

This paper continues as the following: the next section presents literature review and hypothesis development. The following section presents research methodology. The final section presents results and summary.

PERSONALITY FACTORS AND ACCOUNTING PROFESSIONAL JUDGEMENT

According to Bonner (1999), a professional reasoning and judgment can be assessed based on the following three dimensions: first, the task, which the judgment is about it, second the environment, in which the professional judgement is made, and third, the professional, the person who makes the professional judgement, There are different factors that influence each of these three dimensions, which can ultimately affect the judgment and choice of action. Factors that can affect the task dimension are related to the task complexity and information availability. Task complexity requires further discussion of the task and possibly breaking it down to smaller pieces, while information availability is referred to whether all the necessary information about the task is available. The factors that can influence the environment dimension are about the factors that can

influence the environment in which the decision is made. Examples of these factors could be the level of superior pressure and accountant's independence. Factors which affect the professional dimension are related to the technical knowledge, experience and personality. Technical knowledge is acquired through formal training and professional experience, while experience is achieved through engagement and employment. Personality factors are related to the professional as a person and human being.

By concentrating on the personality factor from the third dimension of the professional reasoning and judgement, we can divide the professional related factors into two groups. One group includes the non-personality factors such as professional knowledge and experience, and the other group includes the personality factors. While accountants can acquire professional knowledge and experience by attending schools and participating in the workplace, the personality factors are intrinsic values that either they are part of the person's personality or developed by being surrounded by certain societal environment (Chatzisarantis and Hagge, 2008). Personality factors found to be associated with individuals' judgement and choice of action (Davis, et al., 2007). Therefore, even if accountants receive the finest education for their technical knowledge and acquire relevant experiences, their judgment and choice of action could still be affected by their personality factors.

Machiavellianism as a Personality Attribute

Machiavellianism has been discussed often as an effective personality factor that could influence the individuals' judgment and choice of action. Machiavellianism, which is based on the ideas of Niccolo Machiavelli in *The Prince* (1513), is defined as the aggressive, manipulative, explosive, and devious pursuit of individual and organizational goals without regard for the feelings, rights, and needs of other people (Calhoun, 1969; Paal and Bereczkei 2007). Christie and Geis (1970) argue that Machiavellian behavior includes negativism, lack of conventional morality, and emotional detachment. High Machiavellians differ from low Machiavellians in situations requiring a rapid response or involving interaction with others, and they also have unflattering and cynical opinions of others. Business studies mainly found an inverse relationship between the high degree of Machiavellianism and choice of ethical actions. For instance, Cyriac and Dharmaraj (1994) found that managers' choices of actions were influenced by their degree of Machiavellianism. Verbeke, Ouwerkerk, and Peelen (1996) found that salespersons with high Machiavellianism scores chose unethical actions more frequently. Dalton and Radtke (2013) found Machiavellianism to be negatively related to whistleblowing, and Pratama (2017) found a significant impact of Machiavellianism on students' perception about tax avoidance.

While the effect of Machiavellianism on the choice of ethical actions has been documented by prior studies, it is unclear whether individuals with different Machiavellianism scores react differently when they are exposed to ethical awareness. Ethical awareness can create a sense of responsibility and ethical behavior, but its extent could vary and the success of introducing ethical awareness could be conditional and limited to the individuals' Machiavellianism scores. Christie and Geis (1970) argue that high Machiavellians are more manipulative, successful, and persuasive but less open to persuasion. If this claim is accurate, then the effect of ethical awareness should be insignificant on high Machiavellians. Therefore, Machiavellianism can potentially influence the effectiveness of accountants' ethical awareness. High Machiavellian individuals adhere to the situation faster and seek their personal interest more often. Therefore, they may not incorporate the ethical principles in their decision-making even when they are exposed to ethical awareness.

On the other hand, low Machiavellians may be more receptive of ethical concepts and react positively to ethical awareness. Accordingly, the first two hypotheses are defined as the following:

H1: The effect of ethical awareness is significant on accountants' judgment and choice of action.

H2: The Machiavellianism score interact with the effect of ethical awareness on the accountants' judgment and choice of action.

Gender Interaction

Gender has been frequently discussed as an effective factor in connection with the choice of ethical action. Eagly and Kite (1987) explained the sex differences and similarities in social behavior and present support for the idea that males and females behave according to their socially expected stereotypes. Loo (2003) argued that males are more socialized in principles and relate to competitive success and personal growth, thus being more susceptible to unethical behavior than females. Pant and Sharp (1998) investigated the effect of gender on ethical evaluations, intentions, and orientation and found that women consistently differed from men for the choice of ethical action. Shuab (1994) also found female students demonstrated higher levels of moral reasoning. Similar findings were reported by Stylianou et al. 2013, Dalton & Ortegren 2011, and Weber and Urick 2016. On the other hand, there are studies that found no gender effect on ethical decision-making (Geiger and O'Connell 1999; Hay et al. 2001, Thorne 2001) or found that the difference is quite small (Thoma, 1986).

Prior studies' findings on the interaction effect of Machiavellianism and gender is often contradictory. Jonason and Davis (2018) found that Machiavellianism is lower in females. However, Hunt and Chonko (1984) reported that women had higher Machiavellian scores than men. Yet, Webster and Harmon (2002) and Burns, Tackett and Wolf (2015) found no significant difference in Machiavellian scores between women and men. The interaction effect of Machiavellianism and gender may affect the individuals' reaction to ethical awareness issues.

In the context of ethical awareness, there few studies that address the gender issue directly and they reported contradictory results. Wang and Calvano (2015) found males were more responsive to business ethics education than women. However, Ritter (2006) concluded that women showed significantly improved moral awareness after taking an ethics training. Collectively, the results of the studies about the effect of Machiavellianism on ethical decision making provide support for the argument that gender effect may interact with the effect of Machiavellianism on the success of ethical awareness. This notion is consistent with the argument by Cragg (1997) that ethics education is effective on individuals who are already primed to consider ethical strategies and moral values. Accordingly, one may expect to see the effect of introducing ethical awareness be higher on individual with lower Machiavellianism score, females, or both. Additional support is provided by Szabo and Jones (2019) that found that high Machiavellian males, but not high Machiavellian females, were non-impulsive and high in planning. Thus, low Machiavellians and females may react more positively to ethical awareness. Therefore, the third hypothesis is defined as the following:

H3: The interaction effect of Machiavellianism and gender influences the effectiveness of ethical awareness on accountants' judgment and choice of action.

METHODOLOGY

Sample Selection

The research subjects were senior accounting students at an AACSB-accredited university in the North-eastern of the U.S.A. Subjects had not previously taken any business ethics courses or received any formal ethics training as a part of their classes (the approval for the experiment was received from the university). Students were told that participation was voluntary, and they could leave the room if they were unwilling to participate. Students were assured of anonymity for their answers. Total of 165 students participated. We found four questionnaires incomplete and unusable. Thus, we had 161 usable questionnaires that we used in this study. There were 79 students in the experiment group who watched an ethics vignette and discussed its contents (38 males and 41 females). There were 82 students in the control group who did not watch the vignette and had no ethics discussion (42 males and 40 females). The average age of the students was 22 years old. Participants were full-time students and did not have work experience.

Experiment Approach

Participants were divided into two groups of experiment group and control group. Subjects in the experiment group viewed an ethics vignette and discussed the ethical issues involved, as introducing ethical awareness. They were asked to prepare their answers using a six-step approach based on the methodology employed by Alvin et al. (2002): "1. Obtain and list the relevant facts; 2. Name the ethical issues from the facts; 3. Identify the individual or individuals affected by the outcome of the dilemma and state how each person or group is affected; 4. Specify the alternative actions available to each person who must resolve the dilemma; 5. Identify the most likely outcome of each alternative; 6. Pick the appropriate action." This approach is consistent with the four-component model of Rest (1986), whereby a moral individual must: a) recognize a moral issue; b) make a moral judgment; c) resolve to place moral concerns ahead of other concerns; and d) act on the moral concerns.

Students formed self-selected groups of three or four members to discuss the issues in the vignette. The instructor called on each group to present to the class the issues raised in their discussion and what decision they had reached to resolve the dilemmas. The entire class participated in deciding which group followed the best procedures and had the optimal solutions for the dilemmas. Students were given enough time to discuss and prepare their answers. The discussion was intended to create ethical awareness and educate students about the ethics-related issues. After viewing and discussing the ethical dilemmas, students were given details of a hypothetical scenario and asked to make a decision. Then, they were given a second questionnaire intended to capture Machiavellian scores. Subjects in the control group received the same hypothetical scenario and were asked to make a decision without first watching the vignette and discussing the ethical issues it raised.

The experiment case focused on the deductibility of an expense for business income tax purposes. Participants were asked to assume the role of an income tax return preparer. All participating students had taken the relevant tax courses (covering preparation of corporate and personal tax at basic levels). However, the experiment case was very simple and did not require any sophisticated tax knowledge. Tax preparation case was selected because research shows that

pressure is exercised on accountants to perform unethical acts in income tax return preparation (Cruz, William, and Strawser, 2000; Merz and Groebner, 1982 and Finn, Chonko, and Hunt, 1988). The experiment case is presented in the Appendix A. Subjects read the experiment case and circled the number representing their likelihood of taking a specified action on a Likert scale from zero (definitely not taking this action) to ten (definitely taking this action). Machiavellianism was measured by the 20-item instrument formulated by Christie and Geis (1970). This instrument is widely used to measure Machiavellian scores (e.g., Corzine, Buntzman, & Busch 1999; Dion & Banting 1988; Siu & Tam 1995). The Machiavellianism instrument is presented in Appendix B.

Testing the Realism of the Experiment Case

Participants rated the realism of the experiment case on a Likert scale, from zero (not realistic) to ten (very realistic). Ethical content was also evaluated on a scale from zero (this case does not include any ethical issue) to ten (this case definitely includes ethical issues). These two questions were added at the end of the experiment case. For realism, the experiment group mean was 6.78 (S.D. = 1.98) and the control group mean was 7.12 (S.D. = 1.73). For ethics content, the experiment group mean was 8.45 (S.D. = 1.75) and the control group mean was 8.12 (S.D. = 2.11). These scores indicate that most respondents considered the experiment case to be realistic and include ethical issues.

Hypotheses Testing

We use a 3-way between-subject ANOVA to test the hypotheses. If there is a significant interaction between Machiavellianism, gender, and ethical awareness, then at least one combination of these factors affects the professional judgment and ethical choice of action. In addition, we conduct a mean difference analysis using the Tukey-Kramer (Tukey's W). If the mean difference is significant, it presents support for the hypothesis that the choice of the action is significantly different between the groups. We tested the model's homogeneity of variance using Leven's test, and it was supported ($P < .567$).

RESULTS

Descriptive Statistics

Table 1 presents descriptive statistics. For the control group, the mean for individuals with low Machiavellian score was 4.60 (Standard deviation = 1.67), while the mean for individuals with high Machiavellian score was 5.73 (Standard deviation = 1.86). For the experiment group, the mean for individuals with low Machiavellian scores was 3.5 (Standard deviation = 2.04) and the mean for the individuals with high Machiavellian score was 5.49 (Standard deviation = 2.14). The overall mean is 4.50 (Standard deviation = 2.09) for the experiment group who had watched and discussed the vignette, and 5.17 (Standard deviation = 1.77) for the control group that did not.

Insert Table 1 about here

Test Results

Table 2 presents the test results for the hypotheses.

Insert Table 2 about here

Test Results for Hypothesis 1:

Table 2 shows that the F test for the ethical awareness variable is 5.913 and it is significant at $p < .016$ level. These results present support for hypothesis 1 that the introducing ethical awareness has a significant effect on the subjects' judgment and choice of action.

Test Results for Hypothesis 2:

Table 2 shows that the F test for the interaction of Machiavellianism score and ethical awareness is 6.235 and it is significant at $p < .014$. These results support hypothesis 2 and indicate that Machiavellianism score interact significantly with the effect of introducing ethical awareness on accountants' judgment and choice of action.

Test Results for Hypothesis 3:

Table 2 shows that the F test for the variable that represents the interaction of Machiavellianism score, gender and ethical awareness is 10.892 and it is significant at $p < .001$ level. These results present support for hypothesis 3 and indicate that the interaction of Machiavellianism score and gender influences the effect of ethical awareness on accountants' judgement and choice of action.

In addition, we tested the group mean differences to provide further results about how Machiavellianism score and gender influence the effectiveness of ethical awareness. Results are presented in Table 3.

Insert Table 3 about here

The mean difference for high Machiavellian males who were introduced to ethical awareness and high Machiavellian males who were not was .57 and it is not significant. These results show that introducing ethical awareness had no significant effect on high Machiavellian males for their judgement and choice of ethical action. The mean difference for high Machiavellian females who were introduced to ethical awareness and high Machiavellian females who were not was .45 and it is not significant. These results indicate that introducing ethical awareness has no effect on high Machiavellian females for their choice of ethical action. The table shows that the mean difference for low Machiavellian male who were introduced to ethical awareness and low Machiavellian males who were not was 1.79 and it is significant at $p < .001$ level. These results show that ethical awareness had significant effect of low Machiavellian males. Table 3 shows the results for the mean difference for low Machiavellian females who were introduced to ethical awareness and low Machiavellian females who were not. The mean different for these two groups was .38 and it is not significant.

DISCUSSION AND CONCLUSION

The current study presents support for the argument that ethical awareness has positive effect of the accountants' judgment and choice of action, but such effect is moderated by individuals' Machiavellian score and gender. Specifically, our test results show the choice of ethical action by both high Machiavellian males and females were the same regardless of being or not being exposed to ethics awareness. However, the results were different for individuals with low Machiavellian scores. Male with low Machiavellian scores chose ethical actions more often when they were exposed to ethics awareness. But we do not see such an effect on females with low Machiavellian scores, and their choices of ethical actions were the same regardless of being or not being exposed to ethical awareness.

Our results may be helpful in understanding the inconsistency in prior studies' findings about the effect of ethical awareness. Also, our findings could benefit accounting ethics educators to consider the interaction effect of gender and Machiavellianism to increase the effectiveness of their ethics classes and education. Our results improve prior studies' findings in this area by providing evidence about the interaction effect of Machiavellianism and gender on the effectiveness of ethical awareness.

This study is limited in some respects. Viewing and discussing an ethics vignette may not be representative of a longer ethics education program, although some studies have found that short ethical awareness programs are effective in creating a positive attitude toward ethical actions. Replication of the current study using more comprehensive ethics training could be beneficial. Since the questionnaire was administered during class time, absent students (although a very small number) did not have chance to respond. No attempts were made to administer the questionnaires to them later. This fact may bias the results, albeit infinitesimally. The scenario used in this study is artificial and responses may not coincide with behavior in actual situations. In future studies, ethical dilemmas in other non-tax-related situations could also be used. If data collection challenges can be overcome, a longitudinal study may enhance the generalizability of the findings. This study's participants had received no prior ethics training; therefore, whether the findings are applicable to students who have prior ethics training is unclear. These limitations present research opportunities for future studies in this area.

APPENDIX A

Description of the Case in the Questionnaire

In this study, you are asked to assume the role of a tax practitioner in an accounting firm called Green Line Accounting Services. You report to the vice president who, in turn, reports to the president. The president is the sole shareholder of the firm. To the best of your knowledge, the firm has always prepared and filed income tax returns accurately.

When you are preparing a business tax return, based on the tax regulations, you should not deduct any personal expenses related to the business owners or officers. Deducting a personal expense in a business tax return reduces the business' taxable income and results in lower tax liability to the government. You have several years of experience in preparing and signing income tax returns, but you are currently in your first year at this accounting firm.

You are working on the tax return of a client called Blue-Sky Manufacturing, Inc. Everything looks fine for this business tax return except one expense item that you are concerned about. The expense is for the President of Blue-Sky Manufacturing trip to a vacation resort. You are concerned whether this expense item is a personal expense and then you should not deduct it from the business tax return. You know it is possible for businesses, but not always, to hold business meetings in hotels and vacation resorts. You asked your boss about the nature of this trip and the expenses related to that. Your boss told you that it is business related expense. However, despite requesting it, neither your boss nor the client provided you with any document to support that the trip was a business-related trip. Thus, you are uncertain whether the trip was really a business, or it was a personal vacation trip. Blue Sky Manufacturing is a large and important client to Green Line Accounting Services. As a professional accountant, you must decide whether this expense item should be deducted from the tax return. Please indicate the probability that you will deduct this expense from the Blue-Sky Manufacturing tax return on the scale of 0 to 10 below.

Action to be taken:	<p>I will deduct the travel expense.</p> <p>Circle a number:</p> <p>0 1 2 3 4 5 6 7 8 9 10</p> <p>0 means you definitely will not deduct the expense from the Blue-Sky Manufacturing tax return, and 10 means you will definitely will deduct it.</p>
----------------------------	---

Answer the following two questions about the case that you just read:

<p>How realistic is the situation just described compared with situations you believe could be encountered in practice?</p> <p>0 1 2 3 4 5 6 7 8 9 10</p> <p>(0 means this case seems not realistic to me at all, 10 means it seems highly realistic.)</p>
<p>Does the deduction of this undocumented travel expense involve any ethical issues?</p> <p>0 1 2 3 4 5 6 7 8 9 10</p> <p>(0 means this case doesn't seem to include any ethical issue, and 10 means this case definitely include ethical issues.)</p>

Appendix B

Instrument for Machiavellian Score Measurement

You will find a series of general statements listed below. Each represents a commonly held opinion, and there are no right or wrong answers. You will probably disagree with some items and agree with others. What is needed for this research is the extent to which you agree or disagree with such matters of opinion.

Please read each statement carefully. Then indicate the extent to which you agree or disagree by circling the appropriate number to the right of the statement corresponding to your feelings. The response range is as follows:

1 – Strongly disagree; 2 - Disagree; 3 – Slightly disagree; 4 – Neither disagree nor agree;

5 – Slightly agree; 6 – Agree; 7 – Strongly agree.

1.	The best way to handle people is to tell them what they want to hear.	1	2	3	4	5	6	7
2.	When you ask someone to do something for you, it is best to give the real reasons for wanting it, rather than giving reasons which might carry more weight.	1	2	3	4	5	6	7
3.	Anyone who completely trusts anyone else is asking for trouble.	1	2	3	4	5	6	7
4.	It is hard to get ahead without cutting corners here and there.	1	2	3	4	5	6	7
5.	Honesty is the best policy in all cases.	1	2	3	4	5	6	7
6.	It is safest to assume that all people have a vicious streak, and it will come out when they are given a chance.	1	2	3	4	5	6	7
7.	Never tell anyone the real reason when you did something unless it is useful to do so.	1	2	3	4	5	6	7
8.	One should take action only when it is morally right.	1	2	3	4	5	6	7
9.	It is wise to flatter important people.	1	2	3	4	5	6	7

10.	All in all, it is better to be humble and honest than important and dishonest.	1	2	3	4	5	6	7
11.	Barnum was wrong when he said there's a sucker born every minute.	1	2	3	4	5	6	7
12.	People suffering from incurable diseases should have the choice of being put painlessly to death.	1	2	3	4	5	6	7
13.	It is possible to be good in all respects.	1	2	3	4	5	6	7
14.	Most people are basically good and kind.	1	2	3	4	5	6	7
15.	There is no excuse for lying to someone else.	1	2	3	4	5	6	7
16.	Most men forget more easily the death of their parents than loss of their property.	1	2	3	4	5	6	7
17.	Most people who get ahead in the world have lead clean, moral lives.	1	2	3	4	5	6	7
18.	Generally speaking, men won't work hard unless they are forced to do so.	1	2	3	4	5	6	7
19.	The biggest difference between most criminals and other people is the criminals are stupid enough to get caught.	1	2	3	4	5	6	7
20.	Most men are brave.	1	2	3	4	5	6	7

REFERENCES

- Alvin A. A., Elder, R. J., Beasley, M. S. and Hogan, C. E. Auditing and Assurance Services: An Integrated Approach, 9th Edition, Pearson Education, Inc. 2016.
- Bonner, S. E. (1999). Judgment and decision-making research in accounting. Accounting Horizons, 13(4), 385-398.

- Burns, D. J., Tackett, J. A., Wolf, F. 2015. The Effectiveness of Instruction in Accounting Ethics Education: Another Look, in (ed.) Research on Professional Responsibility and Ethics in Accounting (Research on Professional Responsibility and Ethics in Accounting 19: 149-180.
- Calhoun, R. P. 1969. Niccolo Machiavelli and the Twentieth Century Administrator. *Academy of Management Journal* 12: 205–280.
- Cameron, R. A. and O'Leary, C. 2015. Improving Ethical Attitudes or Simply Teaching Ethical Codes? The Reality of Accounting Ethics Education. *Accounting Education* 24 (4): 275-290.
- Chatzisarantis, N. D. and Hagger, M. S. (2008). Influences of personality traits and continuation intentions on physical activity participation within the theory of planned behavior, *Psychology & Health*, 23:3, 347-367.
- Christie, R., & Geis, F. I. 1970. *Studies in Machiavellianism*. New York: Academic Press.
- Corzine, J. B., Buntzman, G. F., & Busch, E. T. 1999. Machiavellianism in U.S. Bankers. *International Journal of Organizational Analysis* 7: 72–81.
- Cragg, W. 1997. Teaching Business Ethics: The Role of Ethics in Business and in Business Education. *Journal of Business Ethics* 16: 231–245.
- Cruz, A. C., William, E. S., & Strawser, J. R. 2000. A Multidimensional Analysis of Tax Practitioners' Ethical Judgments. *Journal of Business Ethics* 24: 223–244.
- Cyriac, K. & Dharmaraj, R. J. 1994. Machiavellianism in Indian Management. *Journal of Business Ethics* 13 (4): 281-286.
- Dalton, D., & Ortegren, M. 2011. Gender Differences in Ethics Research: The Importance of Controlling for the Social Desirability Response Bias. *Journal of Business Ethics* 103: 73–93.
- Dalton, D., & Radtke, R. R. 2013. The Joint Effects of Machiavellianism and Ethical Environment on Whistle-Blowing. *Journal of Business Ethics* 117: 153–172.
- Davis, C, Oatte, K., Tweed, S. and Curtis, C. (2007). Personality Traits Associated with Decision-Making Deficits. *Personality and Individual Differences*, 42 (2): 279-280.
- Dalton, D. and Ortegren, M. 2011. Gender Differences in Ethics Research: The Importance of Controlling for the Social Desirability Response Bias. *Journal of Business Ethics* 103 (1): 73-93.
- Dion, P. A., & Banting, P. M. 1988. Industrial supplier-buyer negotiations. *Industrial Marketing Management* 17: 43–47.
- Eagly, A. H., & Kite, M. E. 1987. Are Stereotypes of Nationalities Applied to Both Women and Men? *Journal of Personality and Social Psychology* 53: 451.
- Finn, D. W., Chonko, L. B., & Hunt, S. D. (1988). Ethical Problems in Public Accounting: The View from the Top. *Journal of Business Ethics* 7: 605–615.
- Geiger, M. Z., and O'Connell, B. T. 1999. Accounting student (sic) ethical perceptions: an analysis of training and gender effects. *Teaching Business Ethics* 2: 371–388.
- Hay, D., P. M., Larres, Oyelere P., and Fisher, A. 2001. The ethical perception of undergraduate students in computer related situations: an analysis of the effects of culture, gender and prior education. *Teaching Business Ethics* 5: 331–336.
- Hunt, S. D., & Chonko, L. B. 1984. Marketing and Machiavellianism. *Journal of Marketing* 48: 30–42.
- Jonason, P. K. and Davis, M. D. 2018. A gender role view of the Dark Triad traits. *Personality and Individual Differences* 125 (15): 102-105.
- Loo, R. 2003. Are women more ethical than men? Findings from three independent studies. *Women in Management Review* 18 (4): 169–181.

- Merz, C. M., & D. F. Groebner. 1982. Toward a Code of Ethics for Management Accountants. *Strategic Finance* 64: 44–49.
- Paal, T., and T. Bereczkei. 2007. Adult theory of mind, cooperation, Machiavellianism: the effect of mindreading on social relations. *Personality and Individual Differences* 43: 541–551.
- Pratama, A. 2017. Machiavellianism, tax knowledge, and ethical perceptions of tax avoidance: survey of undergraduate students in West Java, Indonesia. *International Journal of Trade and Global Markets* 10 (1): 83-90.
- Rest, J. R. 1986. *Moral Development: Advances in Research and Theory*. New York: Praeger.
- Ritter, B. A. 2006. Can Business ethics be trained? A study of the Ethical decision-making process in business students. *Journal of Business Ethics* 68: 153-164.
- Szabó, E. and Jones, D. N. 2019. Gender differences moderate Machiavellianism and impulsivity: Implications for Dark Triad research. *Personality and Individual Differences* 141 (15): 160-165.
- Shuab, M. K. 1994. An analysis of the association of traditional demographic variables with the moral reasoning of auditing students and auditors. *Journal of Accounting Education* 12 (1): 1.
- Shawver, T. J., and Miller, W. F. 2017. Moral Intensity Revisited: Measuring the Benefit of Accounting Ethics Interventions. *Journal of Business Ethics* 141: 587–603.
- Siu, W., & Tam, K. 1995. Machiavellianism and Chinese banking executives in Hong Kong. *International Journal of Banking Marketing* 13: 15–21.
- Stylianou, A.C., S. Winter, Y. Niu, R.A. Giacalone, and M. Campbell. 2013. Understanding the behavioral intention to report unethical information technology practices: the role of Machiavellianism, gender, and computer expertise. *Journal of Business Ethics* 117 (2): 333–343.
- Thoma, G. A. 1986. The Structure of Trade Between the U.S. and China. *Business Economics* 21: 32–36.
- Thorne, L. 2001. Refocusing ethics education in accounting: an examination of accounting students' tendency to use their cognitive moral capability. *Journal of Accounting Education* 19 (2): 102-117.
- Verbeke, W., Ouwerkerk, C., & Peelen, E. 1996. Exploring the Contextual and Individual Factors on Ethical Decision Making of Salespeople. *Journal of Business Ethics* 15: 1175–1187.
- Wang, L. C. and Calvano, L. 2015. Is Business Ethics Education Effective? An Analysis of Gender, Personal Ethical Perspectives, and Moral Judgment. *Journal of Business Ethics* 126 (4): 591–602.
- Watts, L. L., Medeiros, K. E., Mulhearn, T. J., Steele, L. M., Connelly, S. and Mumford, M. D. 2017. Are Ethics Training Programs Improving? A Meta-Analytic. *Review of Past and Present Ethics Instruction in the Sciences. Ethics & Behavior* 27 (5): 351-384.
- Weber, J. and Urlick, M. J. 2016. Examining the Millennials' Ethical Profile: Assessing Demographic Variations in Their Personal Value Orientations. *Business and Society Review* 122 (4): 469-506.
- Webster, R. L., & Harmon, H. A. 2002. Comparing Levels of Machiavellianism of Today's College Students with College Students of the 1960s. *Teaching Business Ethics* 6: 435–445.

Table 1**Descriptive Statistics**

		High Mach.	Low Mach.	Overall
Control Group (Without Ethics Discussion)	Male	6.35 (1.78)	5.11 (1.67)	5.73 (1.72)
	Female	5.11 (1.95)	4.10 (1.68)	4.60 (1.81)
	Total	5.73 (1.86)	4.60 (1.67)	5.17 (1.77)
Experiment Group (with Ethics Discussion)	Male	5.98 (2.23)	3.06 (1.98)	4.52 (2.10)
	Female	5.01 (2.05)	3.95 (2.11)	4.48 (2.08)
	Total	5.49 (2.14)	3.50 (2.04)	4.50 (2.09)

Numbers in the table indicate means. Standard deviations are presented in parentheses.

Table 2
3-Way ANOVA Results for the Analysis of the Interactive Effect of Ethics Discussion, Machiavellianism score and Gender.

Source	Type III Sum Squares	d.f	Mean Square	F	P ^b
E.A.	45.139842	1	45.139842	5.913	0.016
Mach ^a	31.49025	1	31.49025	4.125	0.044
Gender	14.290848	1	14.290848	1.872	0.173
E.A. * Mach	47.59799	1	47.59799	6.235	0.014
E.A. * Gender	25.405952	1	25.405952	3.328	0.070
Mach * Gender	26.45181	1	26.45181	3.465	0.065
E.A. * Mach * Gender	83.149528	1	83.149528	10.892	0.001
Error	1168.002	153	7.634		

^a E.A. and Mach stands for Ethical Awareness and Machiavellianism respectively.

^b One-tailed test.

Table 3
Test Results for the Mean Differences

Combinations	<u>Mean Difference</u>	Sig.
High Mach Male, Exposed to Ethics Discussion versus High Mach Male, Not Exposed to Ethics Discussion	.57	.554
High Mach Female, Exposed to Ethics Discussion versus High Mach Female, Exposed to Ethics Discussion	.45	.876
Low Mach Male, Exposed to Ethics Discussion versus Low Mach Male, Not Exposed to Ethics Discussion	1.79	.001*
Low Mach Female, Exposed to Ethics Discussion versus Low Mach Female, Not Exposed to Ethics Discussion	.38	.801

*Significant at $P < .05$. Tukey-Kramer (Tukey's W) test.

Good Communication Methods & Stakeholder Engagement Resulting In Effective Project Management

Thakaramahlaha Lehohla

Enrolled with: Faculty of Engineering and the Built Environment, University of Witwatersrand
Department of Industrial Engineering
Johannesburg, South Africa
South Africa
thakalehohla@yahoo.com

ABSTRACT

Communication acts as an effective tool resulting in positive project results, hence project manager spent 90% of the time communicating in an ideal project environment leading to benefits of strategic importance. Communication management in essence entail the process ensuring that information needs of and stakeholders expectations are met via the development of templates and implementation of activities designed to achieve effective information exchange. Project communication classified in two areas, the initial part entails developing a strategy to ensure communication is effective for stakeholders the second area entails carrying out the activities necessary to implement the communication strategy. In the South African scenario, the paper highlights on two project scenarios and the application of project communication management; these are the Gautrain and the E-TOLL project. This exercise will draw a comparison between effective communications and stakeholder management resulting in positive results versus poor communication and stakeholder engagement leading poor project results.

INTRODUCTION

Both the Gautrain and E-tolls project are national projects in South Africa, the Gautrain project was incepted in response to the heavy traffic congestion in the Gauteng national roads and the response to this was the diversification of the National Transport system to a high-tech and efficient rail system connecting Tshwane, Ekurhulani and Johannesburg key metros. This was further enhanced by an integrated bus service for residential and business neighbourhoods in close proximity to the Gautrain rail stations. This projects beauty is the intense communication and stakeholder project management it encountered; whereby from the project initiation phase

the stakeholder identification was intense as required by the project knowledge area and process group. As this was a construction project and would involve requesting residents of certain areas to relocate and be compensated accordingly as the designed route encompassed many residential areas.

This was very crucial as the initiation phase of the project entails the highest risk in terms of reaching consensus with internal and external stakeholders. The medium and frequency of the communication is also essential in terms of the project. This outstanding stakeholder engagement and communications methods reflected in the Gautrain project showed positive results in the end project and product; immediately upon completion of the project in 2010 it managed to integrate well into the major 2010 world cup in terms of transportation of world cup spectators from the O.R Tambo to their various destinations upon arrival, during and post-world cup. At a macro-economic level, the Gautrain has been a major success in terms of an efficient transport system between metros of Tshwane, Johannesburg and Ekurhuleni in terms of occupancy over the years there has been an increase of to date, as a result there is a need to expand the network to greater Johannesburg and Tshwane neighbouring areas.

Whilst the Gautrain scored great victory in terms of good communications methods and stakeholder engagement across the board leading to a successful project and final product the E-tolls were a contrast this can be related to mainly poor communication and stakeholder engagement. To date there is a continuous debate on whether the payment of the tolls are going to be enforced or it will be included in the fuel levy, there has not been a final solution as the majority of citizens are not paying , e-tolls were erected as to enable the maintenance of roads by the users and is required for users to pay based on the utilisation of the roads, there was a tag system in place whereby road users were required to tag to ensure the capturing and recording data on road usage, this has not been a success with mixed feelings on the project and alleged corruption. Clearly when looking at the merits of the project there should be a maintenance of national roads and requires contribution by users but in this instance there was no adequate communication and stakeholder engagement hence the deficiency in the project and product. In ensuring project success communication method accompanied by stakeholder engagement strategy is key.

COMMUNICATION METHODS

The communication methods could comprise of proper design of messages to be communicated, identification of stakeholder through a stakeholder map, a table identifying the

information and audiences, the planned communication including dates or frequencies to explain how the audiences may be engaged, this can all be collated through the RACI (Responsible, Accountable, Communicate and Inform) matrix. Below is a perfect example of a communication method that can be utilised towards gearing the public in accommodating the fuel levy or other forms of contribution towards the roads infrastructure; now that the e-toll has proven not to be a success. The model is a communication method comprising of messages directed to the public in relation to the sustenance and maintenance of road infrastructure and what is required of a responsible citizen in achieving the mentioned.

<p>AUDIENCES</p> <p>MESSAGES</p>	<p>The general public and motorists utilising roads, Pretoria, Ekurhulani, Johannesburg</p>	<p>Business associations, public interest groups</p>	<p>The general public and motorists utilising roads, Pretoria, Ekurhulani, Johannesburg</p>	<p>Business associations, public interest groups</p>
<p>The maintenance and sustenance of South African roads infrastructure</p>	<p>Methods:</p> <p>Project team to communicate message through electronic media, print media and social media</p> <p>In the form of five-minute advert on TV stations and social media</p> <p>Posting of placards and banners on strategic key areas on national roads and major high-ways</p>	<p>Methods</p> <p>Workshops and conferences to be hosted by SANRAL and invitation to key strategic stakeholders with influence on communities</p>		
<p>The benefits of maintaining good road infrastructure for public, business and economy</p>			<p>Methods</p> <p>Project team to communicate message through electronic media, print media and social media</p> <p>In the form of five-minute advert on TV stations and social media</p> <p>Posting of placards and banners on strategic key areas on national roads and major high-ways</p>	<p>Methods</p> <p>Workshops and conferences to be hosted by SANRAL and invitation to key strategic stakeholders with influence on communities</p>
<p>The approach to collect the road levy for usage of roads by public:</p> <ul style="list-style-type: none"> - Integrate within fuel levy - E-toll payment 			<p>Methods</p> <p>Project team to communicate message through electronic media,</p>	<p>Methods</p> <p>Workshops and conferences to be hosted by SANRAL and invitation to key</p>

			<p>print media and social media</p> <p>In the form of five-minute advert on TV stations and social media</p> <p>Posting of placards and banners on strategic key areas on national roads and major high-ways</p>	<p>strategic stakeholders with influence on communities</p>
--	--	--	--	---

Table 1. Communication methods (Source: Adapted from Robert, 2020).

Plan communication management approach

A consensual approach is recommended when a project will involve stakeholder communities, it requires actively lobbying their views, as a means of limiting resistance, such as that of the e-toll cooperation from South African motorists. The purpose is to ensure community allegiance to the vision. In the South African e-toll project context this requires sharing of messages among stakeholders(motorists), the sense of sharing the vision of erection of e-tolls for the maintenance and sustenance of the roads infrastructure. Whereas currently there is various interpretation on the matter as top political leadership did not include the public view on e-tolls; the post erecting of the tolls and dearth of participation from the public is leading to alternatives and discussions being held. Table 1 is a detailed communications management approach ought to have been implemented during project inception phase.

According to Burke (2021), plan communication management is the process of gathering and analysing project stakeholder information and requirements in order to develop an appropriate communications approach for the project. Inactivating this process, the communication plan is utilised as a tool.

1. Identify stakeholder requirements	This entails packaging communication along the maintenance and substance of the road infrastructure and required participation from the responsible citizenship in ensuring roads maintenance, sustenance and improvement. This would be along the lines of safety and security along the national roads, as well as ensuring that vehicles are protected from being damaged by potholes and other hazardous factors when road infrastructure is not maintained.
2. Scope of communication	The aim is to communicate sufficient information pertaining to the importance of a well-maintained road infrastructure the benefits as well as the role of the responsible and true citizen
3. Communicate with whom	This requires formal or informal lines of communication in line with the organisational chart, clearly outlining the project manager position and also who

(lines of communication)	implying the levels of responsibility , authority and whom reports to whom; then linking stakeholders (clients, vendors , contractors and general public comprising of motorists).
4. Format and content	Develop reporting format and content to be discussed with stakeholders. The information presented should be in an easy to understand format so that the motorist and users of the road may assimilate with the road maintenance infrastructure in relation to the paying of the road levy or e-toll.
5. Method of communication	The most appropriate method of communication should be agreed upon with the motorists and other stakeholders affected or impacted by the e-toll project. Communication may be in the form of internet or mobile phones.
6. Timing	The frequency of reports and turnaround time for responses should be discussed and agreed. The reporting cycle should capture progress on a particular day, then processed upon an immediate agreed upon date, then sequentially reported on the next consecutive day which is a weekly progress meeting. The communication plan should produce a schedule of the key communication activities and integrate them into project plan
7. Document control	The communication plan should outline the level of document control.
8. Administration	The administration of the communication process would normally be a project management office function; which entails information gathering, processing, transmitting, filing , storing and retrieval.
9. Resources	The communication plan should outline the resources and budgets required to set up and manage the communication system.

Table 2. Communication plan (Source: Adapted from Steyn, 2021).

In relation to the communication lines the following is a schematic representation on how the lines of communication should flow from within the organisation to the impacted or affected stakeholders. This acts as a guide for the e-toll company regarding the development of communication lines between the entity and stakeholders. This acts as an enabler for communication reinforcements between the affected or impacted parties. Communication and the related lines are an essential as a principle as collaboratively they enforce the standardisation of vocabulary and structure of information to be exchanged between the E-toll company and the affected stakeholders, Alleman (2014).

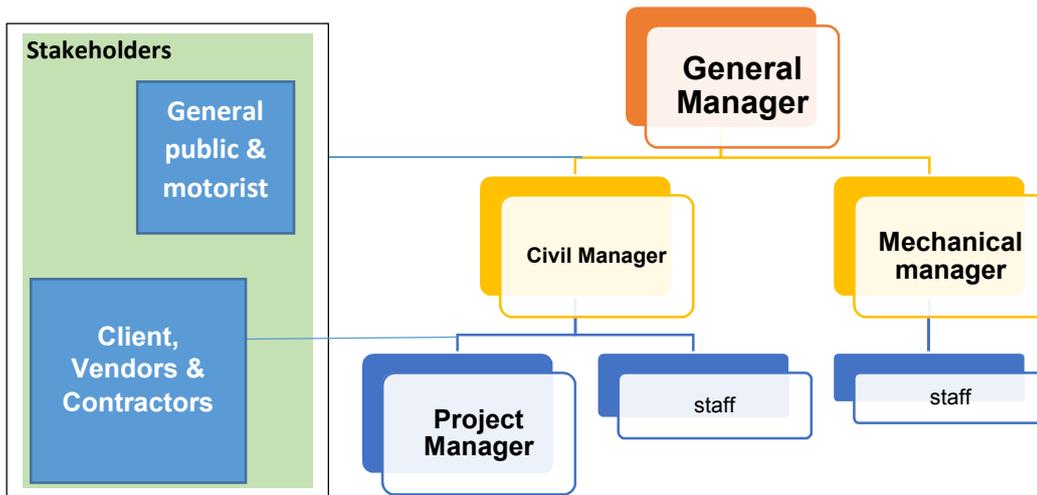


Figure 1. Lines of Communication (Source: Adapted from Steyn, 2021).

Having highlighted on the importance of the role of communication management precipitating to successful project management, the next phase of the paper is to highlight on the importance of project stakeholder management in delivering projects successfully.

Project Stakeholder Management

Stakeholder management is defined as the systematic identification, analysis, planning and implementation of actions designed to engage with stakeholders, Burke (2021). In relation to the e-toll project this is an ideal concept as motorists and other affected public by the project require a critical analysis and planning to be developed based on their interest being taken into cognisance; the current apathy by the motorist in terms of paying the e-tolls provides a clear indication of a dearth of stakeholder management that was performed at the project planning phase.

To reiterate on the above assertion Newton (2013), highlights that stakeholder management is a critical skill that project teams must possess, as the level of effort required can be quite immense. It is recommended that there be continuous interactions with the stakeholder community as to ensure project success. Key objectives in this regards entail ensuring productive relationship with stakeholder community and finally making stakeholders proactive resource for the project.

In light of this it is crucial that the e-toll management applies the following stakeholder management approach in future projects as to rectify and avoid the impasse that occurred in the previous project.

1. Identify stakeholders	In this particular instance of the E-toll project the concerned stakeholders are the motorists utilising the highways and major routes for conducting day to day business or leisure; in furthermore classifying their composition in terms of their ability or willingness to part-take in the project further research, interviews, and checklist to be conducted.
2. Assess their interest and influence	This looks at the classification of stakeholders and their impact on the project, this is represented in a form of a matrix that estimates interest and influence on a simple scale (low/medium/high) . This takes into account that for the e-toll system to work you require 90% of users to pay for the system. The maximum ever paid was 40% of motorists , and currently only 20% is paying.
3. Develop communication plans	This requires for the lines of communication to be established between the e-toll company's project team , vendors, contractors and motorists in terms of addressing the information required , when it is required and how it should be communicated
4. Engage and influence stakeholders	This process entails the e-toll company's ability to communicate and work with motorists, contractors and vendors in order to meet their needs and expectations , ability to address issues as they occur and finally the ability to foster appropriate stakeholder engagement in the project decisions and activities.

Table 3. Project Stakeholder management steps (Source: Adapted from Burke, 2021).

The above tool in its form has a greater propensity in countering the following hindrances contributing to poor stakeholder engagement and result being project failure. As highlighted by Newton (2013), the common bottlenecks comprises:

Treatment of stakeholder as a low-priority activity, resulting in insufficient effort in stakeholder management. Inter alia, No planning of stakeholder management, subsequently leading to inefficient resources and time, lack of clarity of responsibility for stakeholder management. Finally; inability to see things from the stakeholder perspective. These pointers are applicable to the E-toll project as there was no engagements and final agreement with the public to undertake this project and the result is that there is poor collection of the e-toll service, as the public is not cooperating. The above tabular representation addresses, these deficiencies through the stakeholder identification process, their interests and influence; finally communication and continuous engagement process.

CONCLUSION

In order for project success to prevail in project driven environments project communication and stakeholder management are paramount, as failure to acknowledge them and apply them at the critical phases of the project lifecycle will lead to project failure in the end. In relation to the project stakeholder management it becomes very essential in the initiation phase of the project and can certainly make or break a project, the e-toll project is a typical example of this whereby only 20% of the road users are contributing towards the e-toll, whereas for the success of the project is it required for 90% of the motorist to contribute. Project communication management occurs throughout the life cycle of the project approximately 90% of the time the project manager is on continuous communication, in the case of the e-toll project communication with the motorist stakeholder was most certainly not emphasised. The lesson learned from this experience makes it vital that future projects should integrate the appropriate project stakeholder and communication management processes and systems into the entire project value chain as a measure to prevent adversities at any phase of the project life-cycle.

BIBLIOGRAPHY

Burke, R. 2021. *Project Management Techniques Artificial Intelligence*. 4th edition. Hong Kong: Burke Publishing.

Alleman, G. 2014. *Performance Based Project Management. Increasing the probability of Project Success*. United States of America. AMACOM.

Newton, R. 2013. *The Project Management Book*. United Kingdom: Pearson.

Roberts, P. 2020. *Change and Project Management getting it right and achieving lasting benefit*. 3rd edition. London. The Economist.

Steyn, P.2021. *Project Management: Leading, Creating, Implementing & Improving*.

South Africa: Cranefield.

Biographies

Thakaramahlaha Lehohla is in pursuit for a Phd in Industrial Engineering, the research topic is Embracing Total Quality Management Systems in Manufacturing Value Chains for global competitiveness in the post covid economy. Currently holds a PhD in Operations & Quality Management, Master's Degree, Post-Graduate Diploma, Advanced Diploma and Advanced Certificate in Project & Programme Management all obtained from Cranefield College of Project &

Programme Management. Furthermore, holds a Bachelor of Commerce from the University of Pretoria. The Master's degree dissertation was based on the environment within the Manufacturing Competitiveness Enhancement Programme; which was a flagship programme at the Department of Trade and Industry meant to protect and develop industrialisation during turbulent economic period, as well as advance global competitiveness, skills development, export promotion and employment in the manufacturing industry. The topic of dissertation was "Applying Project and Programme Management Principles to the Manufacturing Competitiveness Enhancement Programme. The researcher has been employed by the department for the past eight years, serving as a deputy director. The Ph.D qualification enrolment was with the School of Engineering and Built Environment, under the faculty of Quality and Operations Management. The researcher's topic for thesis was "Developing Project, Programme and Operations Management Methodologies for Sustainable Industrialisation In South Africa". His interests include Industrial and Economic development research, with focus of integrating the project and operations management body of knowledge.

Developing a Document of Good Practice to Support SEND Professionals in Their Role of Eliciting Pupil Voice and Achieving Meaningful Participation

Pooja Sharma

Abstract— Eliciting the views, wishes and feelings of children and young people with Special Educational Needs and Disabilities became the primary element of a new law within England, as part of the Government's commitment to improving services for vulnerable children and their families, known as the Children and Families Act (2014). Despite the consistent rhetoric on the importance of eliciting the voice of children and young people and ensuring person-centered approaches to assessment, professionals encounter a range of barriers that inhibit their ability for meaningful engagement. This study worked with a Local Authority and two specialist schools within England to implement a 'document of good practice' for professional use when eliciting voice. Through participatory methodology and iterative cycles of refinement, the end document serves as a toolkit to support professionals in their duty to elicit voice and achieve meaningful participation of children and young people with special educational needs.

Keywords—inclusion, special educational needs and disabilities, pupil participation, pupil voice.

Coping for Academic Women Departmental Heads during COVID-19: A Capabilities Approach Perspective

Juliet Ramohai

Abstract— This paper explores how women departmental heads in higher education experience leadership in a time of the COVID-19 crises. The focus is mostly on their care and coping as they work in virtual spaces. Most scholars have looked at the effects and challenges that different employees face while working from home during a lockdown. However, very few take a dedicated focus on women in leadership and the coping mechanisms and resources that they use for effective leadership during this difficult time. The paper draws on two aspects of Sen's Capabilities approach, functionings, and agency, to cast a closer understanding of the institutional and individual coping mechanisms that might be at these women's disposal. The qualitative approach used for this paper and a feminist lens provides a critical and in-depth understanding of the real-life stories of the women and how they make sense of their virtual leadership. Data for this paper was collected through semi-structured interviews with 10 women in the positions of head of departments and analysed thematically using capabilities approach concepts as an analytical tool. The findings in this paper indicate that functionings and freedoms are tightly linked to institutional ethnographies. These ethnographies might support or hamper coping for women leaders, especially during times of crisis.

Keywords— capability approach, women leaders, higher education, COVID-19.

Educational Video Games Aimed at Enhancing Academic Motivation and Learning Among African American Males

Kenneth Philip Jones Sr.

Kenneth P Jones Ken is with the University of Illinois-Urbana, United States (e-mail: drjonesk67@gmail.com).

Abstract

Games are the most elevated form of investigation. - Albert Einstein.

This dissertation investigates the potential of developing educational-based video games to motivate and engage African American males. The study employed a qualitative methodological approach by investigating African American males who are avid video game players and are currently enrolled at a college or university. The participants were individually and collectively video and audio recorded during the interviews and observations. Situated Learning theory analyzed how motivation and engagement can transfer from a video game to an educational context. The research aims to address the disparities in our educational systems when it comes to providing a culture, climate, and atmosphere that will enable the academic development of African American males. The primary objective of the findings is based on the participants' responses and the data collected to provide recommendations to educators and scholars on how to address the issues that have demoralized African American males in education and provide a platform that will allow for equality in educational development and advancement.

Keywords: video games, cognition, motivation, behavioral, emotional, learning transfer.

Acknowledgments

First, I would like to thank God, without whom I would not have been able to make such accomplishments.

First, I would like to thank my mom for giving me the strength and courage to accomplish this goal. Though you are no longer with me in the physical state, I feel your spirit and presence every day. I remember when I first started college, you told me how proud you were of me. Those words have resonated in my heart and soul, and I can honestly say they were the driving force on days when I thought about giving up. I hope that I have made you proud, and I still have more work to do. Next, I would like to thank my sisters and sister-in-law. I will be forever grateful for your love

and support. There is a lot more that I would say, but you all know how I am and that I can go on for days, but you know how I feel about you.

A big thank you goes to Dr. Denice Hood for answering my call and being there for me when I needed someone to talk to at that time. I was afraid and nervous before our phone conversation held on a Sunday afternoon. However, you restored my faith and belief that genuine people care, which means more to me than you can imagine. I am forever grateful and appreciative of you.

I can hardly say how much you, Dr. Yoon Pak, have done for me. The first time I spoke to you, I honestly did not believe that our conversation would be favorable for me, but after a couple of minutes of talking to you, I knew that I had made the right decision. You made me feel important as a human being, which stood out to me more than anything. There are so many things that I want to say, but I will keep it short. I hope you will be my mentor, and I look forward to working with you and starting projects that will benefit many people's lives.

I am saving the best for last. To my wife, Valerie, who has always believed in me and stuck by my side, and for that, I will always be grateful. You respected my space and provided me with the time and place to execute the required reading and writing to propel me to my current position. You have always been my biggest supporter and a true inspiration. Now, it is my turn to do the same and take care of you on your journey. Again, I am keeping it short because if I say everything that I want to, it will take another ten-twenty page, and I do not wish to look at the faces of Dr. Pak and my committee chairs after reading a novel.

Finally, I would like to thank my children, Kela, Kinnady, and Kenneth Jr. One aspect of this journey was to be an example for you; no matter what happens, you can accomplish anything in this world, and no one can stop you from achieving success. Therefore, I hope you follow my path and obtain your higher education degrees. My wish is to start a family legacy of high academic achievements for further generations.

CHAPTER I

Introduction

“Dear Young Black Males, it's okay to be different. Don't be afraid to be yourself. Have courage! Follow your dreams, no matter how BIG your dreams may seem. Attitude is everything! Make sure that you keep a positive one, in spite of any obstacles that may come your way. Don't be so quick to give up, and please remember that self-discipline is your friend. Be strong, persevere, and most importantly, BELIEVE in yourself. Don't listen to anybody's negativity. Move forward knowing that you CAN and you WILL. Be unstoppable! - Stephanie Lahart

It was once a hope and promise that schools would be that entity that would change the downward trajectory of the African American male, but that hope has not come to fruition. It is well-publicized that highly effective school experiences play a vital role in determining the quality of life for many African American males. However, developing an educational environment that

accepts the culture and responsibility of educating Black people, specifically, Black males remains a problem for schools in the United States of America. African American males are familiar with the perception that they are viewed as a “problem.” On the other hand, school administrators, teachers, and society emphasize the issues associated with African American males but do not recognize their tremendous potential. As a result, this has led to a misconception that most Black males will fail at school, which would abolish a platform for developing solutions.

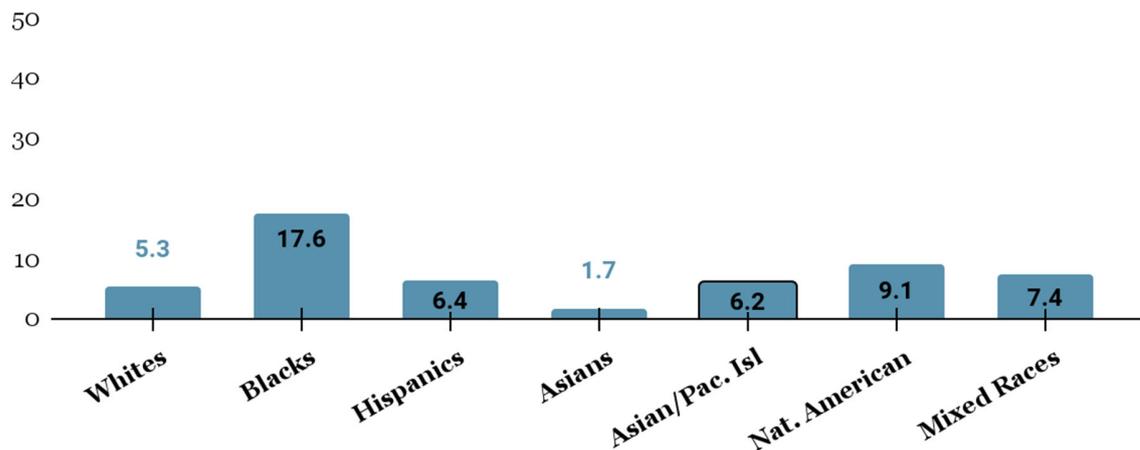
Up Against The Wall

Educational initiatives such as No Child Left Behind, Race To The Top, and Common Core State Standards, designed to offer equal opportunities, have not produced positive results for many African Americans and specifically Black males. Numerous studies have consistently shown that African American males have the lowest statistical academic grades, test scores, and the highest number of high school dropouts than any other race or gender in America. In addition, African American males are still one of the most academically and socially marginalized groups of students in the United States of America (E. Anderson, 2008; Fergus & Noguera, 2010; Gill, 1992; Noguera, 2008; Terry, 2010; Toldson, 2008).

A considerable amount of research has presented significant obstacles for African American males in obtaining an equal and quality education. Pedro Noguera’s (2008) study on *The Troubles With Black Boys And Other Reflections on Race, Equity, And The Future Of Public Education* stated that African American males have the highest number of homicides as the victims and perpetrators in the United States. Black men have been contracting HIV and AIDS at a higher rate than any other race or gender. African American males have the highest number of incarcerations and criminal convictions than any demographic. In the United States, Black men are the only group experiencing a decline in life expectancy. Black males are the least likely to be hired for jobs regardless of their skills or educational level. In the United States, Black men continue to have the lowest employment rates than any race or gender.

In K-12 schools, African American males suffer the most when it comes to receiving a quality education. Black males have the highest number of suspensions and expulsions than any other group in America (Harry & Anderson, 1995). In addition, factors such as socioeconomic status, low academic achievements, underachievements, residences in undesirable urban communities place Black males, in particular, at high risks for suspensions. Furthermore, the commonality among Black males and low family incomes enhances schools’ probability of exclusionary discipline practices. Many young Black men are classified with a mental illness or learning disability and assigned to special education classrooms as early as kindergarten. Also, African American males placed in remedial courses receive the highest number of visits to the principal’s office (Wu et al., 1982).

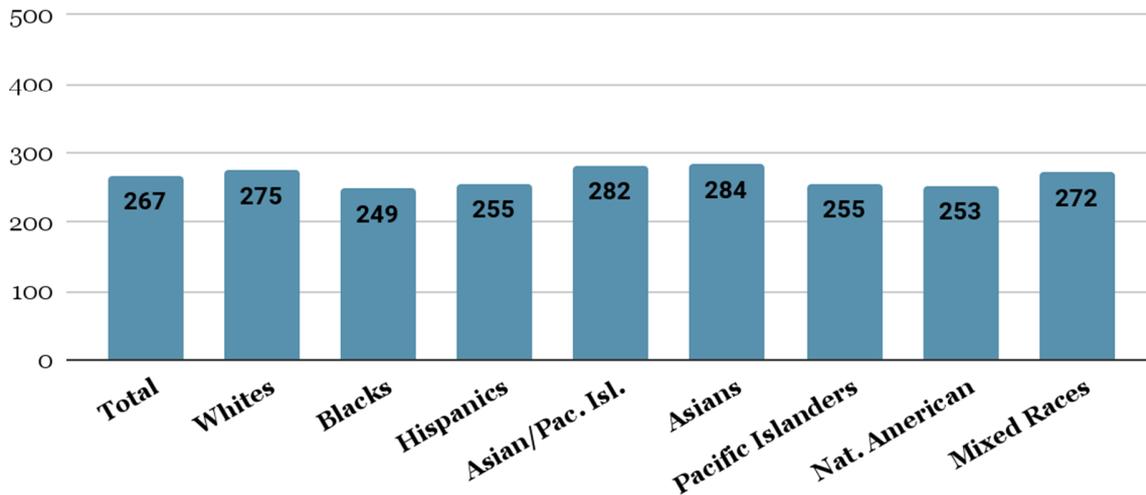
Percentage of male students who received out-of-school suspensions, by race/ethnicity in 2013-14



Source: Digest of Education Statistics 2017, table 233.28

The statistical data from this table demonstrates the disparaging number of African American males who received out-of-school suspensions. In addition, black males nearly doubled Native Americans and nearly or slightly tripled the total number of males for other racial categories. As a result, African American males continuously suffer the most with below-average grades and low test scores. Also, missing time in the classroom damages the ability to establish social interactions amongst peers, which research has stated is critical in providing a healthy learning environment for any student and particularly for African American males.

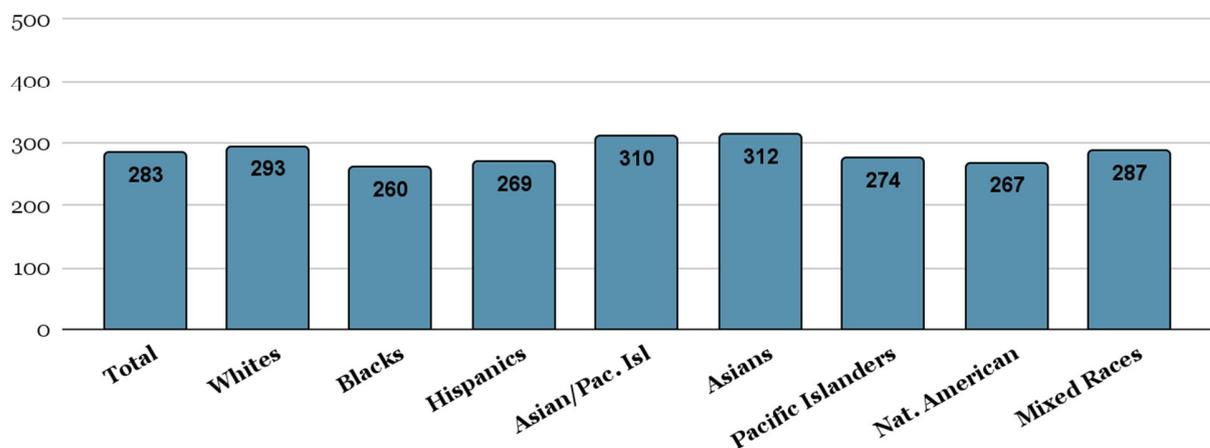
National Assessment Educational reading scale scores of 8th-grade students, by race/ethnicity: 2017



Source: Digest of Education Statistics 2017, table. 221.10

The reading scale scores are a composite of both males and females. African Americans continuously rank at the bottom in reading proficiencies than any other race. Research has stated that African American females perform significantly higher than males, which could have accounted for the overall score. Hispanics, Pacific Islanders, and Native Americans performed slightly above African Americans. The statistical numbers show that more students of color are reading below their peers nationally.

National Assessment Educational mathematics scale scores of 8th-grade students, by race/ethnicity, 2017



Source: Digest of Education Statistics, 2017, table 222.10

African Americans again ranked at the bottom in mathematics with a slightly wider gap between Hispanics and Native Americans. All of the racial groups demonstrated a higher scale score in mathematics. The statistical graph shows a broader gap between mathematics and reading. In reading, Asian Americans (the highest performing group) had a 35 point differentiation over African Americans. In mathematics, Asian Americans had a 52 point difference over the lowest-performing group (African Americans).

After-school activities, such as debating, writing for the school newspaper, or participating in a science-based activity, are not customarily recommended for African American males. Instead, younger Black males are urged to join sports activities and music, specifically basketball, hip-hop, or rap. As a result, Black men are more than likely to be absent from advanced placement and honors classes. In addition, middle-class Black men have been found to have a lower grade point average and standardized test scores than their white peers (Noguera, 2008).

Most disturbing about these statistics centers on the lack of response from community stakeholders. Unfortunately, many educators in the United States of America have become so accustomed to seeing young Black males fail at school that this mentality has become normalized and accepted as commonplace for Black men. Coupled with the educational problems that have continuously affected Black males, there is no outcry, anger, or concern when presented with critical statistical information.

The objective of this qualitative study will focus on examining the potential of utilizing video games as a new method to academic learning through motivation and engagement, particularly for African American males. In addition, my position as a doctoral candidate and African American educator will allow me to gain deep insight and perspective into the emotions and feelings of the

participants through our shared cultural relevance and experiences. The video and audio sessions from the participants were recorded during the late spring to early summer of 2021.

The two general research questions will address the potential impact of educational video games on motivating and engaging African American males to succeed in academic institutions, and they are as follows:

1: What are the mechanisms through which video games can enhance learning among African American males?

2: How can video games transfer learning concepts into educational contexts to promote academic achievements for African American males?

African American Males: Beloved and Hated

African American males were selected for this study due to the vast disparities in learning among all other races and genders. Black males have faced continuous disappointments and more hardships regarding education, economic achievements, and imprisonment. African American males are continuously categorized as a nemesis to societal order, who strikes fear in the minds of many, and at some point, will commit an act of violence and be labeled as undesirable by society. However, the perception of Black males continues to remain a mystery in the more significant portion of society. Black males have played a significant role in designing and creating some of the most significant inventions of modern popular culture. Some of these creations have generated multibillion-dollar products, and major corporations have profited dramatically off of the talents, intellect, and creative genius of Black males (Hill, 2009; Kitwana, 2002).

In many aspects, the world has created a love-hate relationship with Black males. However, when Black men are viewed as economic producers such as athletes and entertainers, the world loves them. Conversely, if they are seen as not conforming, hostile, strong-willed, independent, and outspoken, they are viewed as a problem (Howard, 2014). These attitudes and perceptions of Black males are viewed within our society and, in particular, the public and private school system. Ladson-Billings (2011) explains in her words the issues of how society looks at Black males. “We see Black males as “problems” that our society must find ways to eradicate. We regularly determine them to be the root cause of most problems in school and society. We seem to hate their dress, their language, and their effect. We hate that they challenge authority and command so much social power. While the society apparently loves them in narrow niches and specific slots-music, basketball, football, track-we seem less comfortable with them than in places like the National Honor Society, the debate team, or the computer club.” (p 8)

Today, the images of a Black male are not that extreme. Nevertheless, they are still negatively viewed in many instances in the news and social media. The rise of the Black male athlete has presented a more positive image than in the past. Nevertheless, it has not erased the traditional negative stereotypes that still exist for many African American males. Society has portrayed Black men as derelicts, neglectful fathers, gang members, and drug dealers. Also, they view their significance to life based on the sexual conquests of women and the number of children born.

Ralph Ellison's 1952 novel *Invisible Man* starts with the narrator, a Black man describing himself as a man of substance, fiber, flesh and bone, liquids, and a mind but who is invisible to the rest of the world because people refuse to see him (Ellison, 1952). Black males are often portrayed as criminals, slicksters, and buffoons in literature and movies throughout American history. According to Apel (2004), the image of the Black man has often been symbolized with some form of violence. Some members of society have regarded Black males as untamed and violent men with unrefined masculinity. Gray's (2020) research on *Intersectional Tech – Black Users in Digital Gaming* depicted the masculine identity of young African American males who had entrenched themselves in the gaming world as an escape from the heightened violence occurring across America. One of the participants interviewed by Gray spoke about why he had immersed himself in the world of gaming. "Nah, I ain't never left gaming. Well, not unless I couldn't afford Live! [Laughing] This is my place, though. Social media cool. It just makes me mad. Well, Twitter cool, but Facebook for the birds. I like following people like BlackLivesMatter people. And the hashtag that kinda show us in a different way. That one had me thinking real hard about how they show us. That was so powerful. It put on Front Street exactly how media do us. Seeing that. All the stuff from Black Lives Matter. Learning about Cointelpro. Hoping White folks don't call the police on me. And other stuff. It is just hard. It is a lot. I am glad I got the game" (p. 62). A significant number of these gamers have used the digital communities and the world of gaming as a seclusive retreat from the consistent threat of violence and death.

Stereotyping in Video Games

Over the past few decades, a significant amount of research has begun to examine how people of color are depicted in video games. A study by Glaubke et al. (2001) looked at over 70 video games. Moreover, the research determined that specific minorities (e.g., Latina women and Native American men) had no presence. All video games developed under the format for kids only included white characters who were not human. Latino avatars' had a strong presence in sports games. Asian characters were predominantly present in games that were associated with wrestling and fighting. However, video games have often presented Black male characters as untouched by violence and any form of physical suffering.

Further investigations by Dill et al. (2005) presented findings showing that African American males, in particular, are seldom seen as the protagonist in a video game and secondary characters. According to a study on racial stereotypes (Dill & Burgess, 2011), white students depicted Black men as aggressive criminals according to their video games' images. A significant amount of research has looked at the damage caused by the negative portrayals of minorities and particularly Black males in video games.

One of the most egregious depictions comes from the video game *Grand Theft Auto*. In Boston.com, Haitian and Cuban descent characters are criminals who incite hate crimes and violence. The seriousness of this matter reached political representatives of the Haitian and Cuban governments, who denounced these portrayals in the game after a couple of characters stated: "Kill the Haitians" and "Kill the Cubans" (Diaz & Medina, 2003). As a result, the producer of *Grand Theft Auto* had the lines removed from the game's storyline.

Aggression, hostility, and criminality have continuously perpetuated African American males by devaluing their social identities. According to Berkowitz (2008), characters being stereotyped has become almost commonplace for African American males in video games. He also states that "those who are associated with aggression generally and with gratifications for aggression in

particular or who are associated with negative affect are likely to be the victims of aggression” (p.128).

The Henry J. Kaiser Family Foundation (2002) presented a study that discovered African American males have a higher probability of playing video games than any other demographic. Moreover, research supported this finding by stating that between the ages of 8 and 13, African American males will spend approximately one hour each day playing video games (Glaubke et al., 2000; Koiko et al., 2003).

Glaubke et al. (2000) stated that the lack of African American male representations of video game characters is an indictment of the ongoing issues in the industry. Furthermore, Black representations in video games were primarily seen as characters from sports and hip-hop music on a larger scale. It is significant to note that Black males’ negative images as video game characters exist strongly, albeit the people are not real. For example, in the much-acclaimed video game *Spider-Man: Miles Morales*, the lead protagonist is a half-African American and half-Puerto Rican male living in a high-crime area of Harlem, NY. The community is ethnically diverse and has been victimized and oppressed by rich white men in Manhattan, NY. According to Slusher and Anderson (1987), people respond to visual imageries and events in games in the same manner as they do in real life, and for them, it supports the stereotypes. For instance, suppose this particular viewpoint is correct; when an individual witnesses another fantasy or imaginary figure of a video game character, they should be treated the same. Thus, there will be no difference between a Black male criminal in a video game and a Black male criminal in real life. Both instances confirm the societal-cultural perspective of the stereotype.

A substantial amount of research has proven that stereotypes play a significant role in how a specific ethnicity and gender assesses people from other ethnic groups. These determinations have led to the generalizations, marginalizing, and fear of minorities, particularly African American males. Amodio and Devine (2006) studied the repetitive exposure to a specific examination of particular members of a race or gender impacts how that particular culture is revered. For example, a violent Black man will teach other ethnicities that it is logical to be apprehensive to all Black men no matter their appearance or how they present themselves in any formal or informal setting. These thoughts and behaviors are then catalysts to discrimination and hate.

Explanations Of Chapters

The literature reviews for chapter two will examine the tenets of games and how video games can transfer knowledge into an educational learning context. First, the chapter will explore the learning capabilities in games and demonstrate how these components can be transferred to educational environments. Next, the literature will present the different types of learning transfers, engagements, motivational factors, and the impacts of dreams when a person sleeps. Finally, the chapter will conclude by looking at the importance of assessments in an educational-based video game.

Chapter three of this study will present the methods used to observe and interview the four African American male college students. Situated Learning will serve as the framework to examine key factors guiding the data and analyzing video game learning principles. The researched method section will chronicle the process used to select the participants for this study. Highlights of this section will include the method used to recruit the participants and the approach employed to gather the data for analysis. The chapter will also detail the process used to analyze the

observations and interviews. Finally, it will provide empirical evidence that learning can be derived from video games and transferred into an educational learning platform.

Chapter four will provide the data analysis on the findings from the participants in the study. First, the chapter will provide an outlet for the voices of young Black males who aim to dispel the traditional stereotypes that African American men are the least likely to achieve a degree in higher education. Second, the chapter will conclude with the findings demonstrating that video games have components that can produce motivation and engagement.

Chapter five will conclude with the potential of designing and creating educational video games that will change the negative trajectory of African American males and low academic achievements. I will provide educators and game designers with recommendations to collaborate and design entertaining and educational video games. Finally, I will guide present and future scholarships that can alleviate the number of academic and social injustices that have kept African American males at the bottom of most categories in-school performances. My aim was to apply my dissertation findings and use the data to create a new learning method that utilizes video games as a cornerstone. Nevertheless, extreme measures must be considered to change African American males' continuous downward plight in our educational and social system.

Identifying Family Needs, Support, and Barriers for More Effective Involvement in Early Intervention Services

Sadeem A. Alolayan

University of South Florida

Author Note

Sadeem ALolayan a PhD candidate at University of South Florida — Special education Department -

College of Education (COEDU) salolayan@mail.usf.edu

Abstract

The purpose of Early Intervention (EI) programs and services is to minimize the impact of disability on children ages 0-5 and to reduce future special education cost. This literature review identifies the status of families of children with special needs. Four major themes emerged from this literature review. The first was the family's needs and the expressed desire for services to be obtained or outcomes to be achieved. The second was family support, meaning any information or skills needed to facilitate parents' role as professionals in order to enable them to train and provide their child with the best quality of life. The third theme, barriers, was defined as parents' actions or life circumstances that hindered families in obtaining appropriate EI services. The conclusions derived from the recommendations are that effective parent participation involves careful planning, establishing and maintaining a trusted rapport between parents, and EI providers that understand parents' individual needs and interests, thus motivating effective parent involvement in early intervention programs.

Keywords— Early Intervention (EI), Individuals with Disabilities Education Act (IDEA) , parents, Recommendations.

I. Introduction

Over the decades, there has been a significant development in the provision of services for early intervention for children from the age of birth to three years. Children with disabilities or at risk of developmental disabilities are qualified to receive a Free Appropriate Public Education (FAPE) under the Individuals with Disabilities Education Act (IDEA). Part C of IDEA is Early Intervention (EI), which is a program that each state administers to guarantee children a FAPE [1]EI is a system that provides different services to children who have been diagnosed with a disability or who are

at risk of developmental and behavioral disorders. Early Intervention is an integrated program with services that support children physically, socially, and academically provided by different types of specialists. EI is a crucial part of all activities that occur naturally and routinely for the family and the child. The IDEA Part C EI program aims to enhance the development of infants and toddlers with disabilities, to increase the capacity of families to meet their child's needs and to minimize the need for special education or services later in the child's life. Lack of effective family participation in these programs may result in medical and academic complications [2]

Parents significantly contribute to the child's development and lay the foundations for advocating better serves on behalf of their children. The family system has the most important social impact on child development. Therefore, family support, training, and guidance have become a top priority. This led to developing the individualized family service plan (IFSP). This plan focuses on the family and on therapies that can help children with developmental delays. An IFSP is developed with input from the child's entire family, and it includes features that are designed to support the entire family [3]

Supportive services are available for families at all levels of planning, decision-making, implementation, and evaluation of services. Understanding this concept is vital to providing appropriate EI services to families. However, there are needs that prevent or limit family participation in EI services. In this review of literature, the focus is on identifying these needs and appropriate support for these needs, as well as understanding the barriers that stand in the way of parental participation in EI services. It is worthwhile to define the themes mentioned in this paper:

- Family Support Services are the specific services for families that enhance their child's development provided by agency professionals.
- Family needs refers to the life skills or knowledge that need to be present to enable family members to function optimally.
- Barriers are any issues standing in the way of accessing or participate in EI services.

This literature has documented the needs of parents of children with special needs. Moreover, the majority of the evidence indicated that more support is required for parents. The literature identified potential barriers to positive participation in EI services. Included in the literature are recommendations to enhance coaching parents and accessing and delivering services. In this review, the aim is to focus on knowing the needs of parents of young children with disabilities. What support do parents need to become involved in the services of early intervention? What are the barriers to family involvement in EI services?

Knowing the needs, support, and barriers to EI would help to find solutions to reduce great cost to governments. Of all the potential benefits of intervention focusing on children's early development, it is language and communication skills, and parenting support that are where the most significant impact could be made. As stated in the Early Childhood Technical Assistance Center report (2018) [4], 70 percent of children with disabilities, aged 0 to 66 months achieved similar functioning performance to their normal peers of the same age because of EI services. Clearly, what families need to successfully engage with EI services is an important topic to understand.

Family Needs

Emotions

Bailey and Blasco [5] defined family need as the expressed desire for services to be obtained or outcomes to be achieved. A child's disability entails miscellaneous and crucial challenges, it affects all family members, especially the parents. Indeed, during the time of the diagnosis process, the parent might go through mixed emotions such as denial or grief like those experienced when grieving the loss of a loved one [6]. Having an exceptional child in a family could cause many issues, which may have a significant impact on family adaptation. Maintaining the psychological and mental health of the family is essential. The most prominent emotional problems have been mentioned among parents with different social statuses whether married or singles, employed or not. In other words, these emotional challenges can happen to any parent. Emotional upheaval can cause marital crises and increased aggression and depression, and feelings of guilt, anxiety, and stress [7]

Stress. Over the past years, research has shown that parents of children with disabilities experienced a disproportionately greater level of stress related to their children than did those of children without disabilities [8] Furthermore, many parents experienced periods of stress as they adjusted to the demands of parenting a disabled child. In one study [9] mothers were reported to have lower levels of coping compared to fathers, as well as greater levels of psychological stress. High levels of parenting stress counteracted the effectiveness of early teaching interventions [10] However, stress sources were found to be varied. For instance, the level and intensity of the disability strongly impacted the parents psychologically [11]; [12]; [13]; [14]. Helping siblings understand the problems associated with disabilities and decision-making and getting the child into the right program seem to be other factors that played a significant role in raising the parents' stress level [15]. Parents' level of stress is affected by ensuring their children's physical safety. Dealing with undesirable behaviors and providing constant care constitute a psychological burden that increases the level of stress among families. [16] Lack of attending social activities, a lack of outside or personal social activities, and the inability to have private time for parents would mean the well-being of the family in its entirety could be threatened, potentially leading to impairments in family functioning [17]

Depressions. Parents of children with disabilities appear to be more likely to experience elevated levels of depression as a result of a decrease in quality of life and other issues [18] Generally, evidence indicates that mothers of children with disabilities often report problems with depressive symptoms. Meirsschaut and his colleagues in 2010, found that mothers of children with Autism Spectrum Disorder (ASD) reported elevated stress levels and were at an increased risk for depression [19]. According to Resch and his colleagues in 2012, the research results indicated that no significant correlations were found between depression and parents' age, parent education level, annual household income, employment status, or any of the child's disability variables between the two depression risk status groups [18]. These results are consistent with what Greenberg, Seltzer, Krauss, Chou and Hong found in 2004. The results indicate that there are no differences in the overall levels of optimism, depression, wellbeing or health among mothers of adult children with Autism and Down syndrome [20]. Furthermore, depressed parents had shown several other symptoms associated with hopelessness and depression, such as less problem-solving ability, lower family satisfaction, and lower physical health compared to parents who were not depressed. However, Brotherson's and colleagues in 2010 intended to explore the dynamics of emotional support in home visits. Researchers found that one of the most important emotions that affect parent participation rates and working with EI specialists is grief, depression, and guilt. Parents

had anxiety and sorrow due to their feeling and perception that their child was different and felt a sense of urgency to have concrete ideas and actions to make things better. Thus, several professionals expressed their frustrations because of the family's delay in cooperation [21]

Social Support

Benefits for families can vary due to the various aspects of support received from their social and cultural networks. A lack of or insufficient access to key sources of social support and advice on dealing with daily living skills with an exceptional child such as peer support programs, friends, family, and other social support networks was a significant factor in a family's need for social support [17] One important coping resource is social support, including practical help, information, and emotional support [17] Identifying the types of support that families of exceptional children need is an urgent matter. Formal and informal social support were frequently noted as an integral part of the coping experiences of parents of children with disability [7]. Disability is confusing for families, especially if the families are immigrants. In this regard, Jabra (2016) aimed to study the experience of Spanish immigrant mothers who had an autistic child, by attempting to understand the social, environmental impact of participation in their children's education. It was found that social isolation and stigmatization was clearly present while raising a child with autism in their social environments [22] In addition, there is need for more formal support of mothers of African American children with Autism, because mothers need to directly discuss their concerns with a social services person [23]. For example, one mother named Tameka said, "It would be nice if upon discovery there was more than just some print out that they were handing you—if there was an actual person that they could hand you. Like the social services person—have that person you can call" (as cited in [23]p.28). Interviewing African American parents of young children with Autism to explore reasons that facilitate and impede early diagnosis and accessing EI services revealed that family members and faith were a great source of social support for the mothers [23]. In short, more effective formal and informal support is in demand. Support of the family, friends, and community, especially, emotional support of neighbors, and co-workers would be an endless source of support. The value of this support is the feeling that there are individuals who stand by parents and can support them in situations of distress.

Family Member Participation

Although EI is designed to meet both child and family needs, research indicates that families request more attention to their needs in order to effectively be involved in the practices of EI [24]. Besides the fact that both parents need information, it is valuable to mention that fathers, in particular, need an advanced level of information. Despite being able to have a positive impact on their children with disabilities, fathers are almost entirely unavailable to participate in EI (birth to 3) services [25]; [26]; [27]. Understanding the reasons for their absence is important to solve the problem. Despite the emphasis on fathers' participation in the progress of their child's EI program, unfortunately, mothers continue to be the primary caregiver for the child [25]. Firstly, it should be pointed out that research on the needs of fathers in intervention programs is rarely found. This is likely due to the negative stereotype of non-involvement of fathers. For instance, research has indicated that the traditional view of fathering by Latino fathers is correct, which held that Latino fathers were often distant and behaved according to stereotyped roles of aggressive masculinity [28] Race conceptualization contributes to the neglect of broader social and ecological factors and of variability that exists among parents of certain minorities' fathers, which has led to lack of

generalizable results. This shed additional light on the unique challenges and perspectives faced by fathers of children with disabilities.

The laws of deportation are among the reasons that led to the absence of the father's involvement in EI services. Ijalba's study (2014) conducted on Hispanic immigrant mothers sought to understand the experiences of raising an autistic child. The wives indicated that the absence of the husbands due to deportation caused them to feel sadness. Therefore, single mothers expressed their desire to have a husband to help them in decision-making and support them financially. Losing fathers was a substantial factor in the mother's life [22]. Single African American moms shared the same feeling of sadness that the child's father was not involved. Lack of emotional support when mothers are going through the stages of grief has a huge impact on their involvement in EI [23]. The circumstances of divorce or separation between parents may affect the child's life. Furthermore, the move between two houses or two states impact the child's services and were also some of the needs that mothers expressed. Parents need more stability for their children so the family can settle down and receive services in a satisfactory manner [21].

Transitions and the Importance of Communication

it is essential for children with disabilities who are leaving Part C for Part B of IDEA to move smoothly between those parts, that can be achieved through seamless transitions plans. Parents need to move smoothly from one program or system to another, or among services within a program or system. Studies identified that several parents expressed challenges, frustrations, and barriers they faced when seeking access to transition services for their children. Lack of effective communication with professionals and a deep understanding of placement had the potential to negatively influence the parents' ability to cope with the transition process [29] [14]. Transition planning aims to reduce the stress and worry parents may be going through. Many issues worthy of consideration in the transition process include increasing the connection between the EI program and kindergarten teachers [17]. This connection could be enhanced by maximizing the continuity of services and supporting existing gaps caused by changing services and locations [30] [31]. Miscommunications between EI service providers and families can be a challenge for some parents [30]. Additionally, the school staff in preschool has a more positive impact on families than their counterparts in elementary school [31].

Immigrant parents and transitions. Studies are rare that have focused on diverse or immigrant families during the critical transition process, despite the increase in recent arrivals. These rare studies indicated that parents' language difficulties were associated factors with miscommunication [32] [31]. Discussions between teachers and parents with limited English skills were described as "battles", which clearly reflects the lack of effective communication. However, parents need to express their feelings, and discuss their child's development plan in a natural way. Non-English-speaking parents express the stress and frustration of their incapacity to advocate for their children or communicate with the school personnel regarding their child's needs [31]. To phrase things in a more professional way was a difficulty that arose when parents communicated with staff [31]. Furthermore, Stuart, Flis, and Rinaldi (2006) found that parents felt the most important aspects of a transition program were school climate and open communication [33]. Teachers' lack of knowledge and experience about other families' cultures, values, and beliefs had a strong impact on the rising sense of tension during conversations [32]. In addition, Karoly and Gonzalez (2011) have provided a number of obstacles that represent unique issues faced by immigrants, such as language barriers, cultural sensitivities, and miscommunication [34].

Interestingly, these challenges could be potentially addressed. EI programs should consider using liaisons for effective interagency communication in order to support and guide families during transitions between care systems [29].

Improving communication process. Training the providers themselves with extensive communication skills seems to be an ideal solution for improving communication. A number of researchers have identified methods and strategies that can be useful for building and maintaining effective communication between educators and family members [35] [36]. For example, research has shown active listening strategies as effective for information sharing [35]. In addition, providing families with relevant information during the transition process in a written form translated into the parents' language facilitates the process [29]. However, Rous and his colleagues in a 2007 study titled *Strategies for Supporting Transitions of Young Children with Special Needs and Their Families* found a compelling result. They found that attending and participating in a comprehensive meeting involving all parties was an efficient strategy to help support more positive relationships and communication among families, agency staff, IFSP, and the Individual Education Plan (IEP) team [36].

Overall, it is worthy to point out that in general, the majority of parents expressed the need and desire to build trusting relationships between the kindergarten personnel even before the actual transition transpired [14] [31]. Careful planning and consideration of an array of transition practices appropriate for the individual needs of the child, family, school, and community is essential. Likewise, EI providers should give considerable effort to the professional language they use with parents. This will allow for successful two-way communication, an essential characteristic of effective interaction [36]. To sum up this section on the family's needs, a deep understanding of parental needs would greatly help to identify the forms of support needed.

Family Support

Family support services are any services or supports provided by an EI provider to help families as they learn about their child's delays and learn how to care for a child with a disability. In most cases, parents are seeking to gain the appropriate skills and information that would facilitate their mission of caring for children with special needs. Many countries have established laws and regulations to support children with disabilities and their parents. Thus, EI providers are responsible to carry on delivering the message of the importance of EI [15]. The providers should make sure the parents understand their roles as partners in successful EI programs as they coach their children. Therefore, active participation and partnership between families and EI program providers leads parents to become advocates on their children's behalf [2].

Information as a Form of Support.

It is important to support parents in building a strong information base that can be used with any of their child's caregivers [31]. Parents must also learn about their child's daily routine and developmental activities at home [37]. The EI providers should be aware of the best way to give instruction to parents based on the family's ability to learn, and since EI providers are the major conduit for providing information and resources, they should identify the options the parents have [37] [25] [31].

Family's education background. Referring to what Edward (2016) and his colleagues recommended in a study conducted in Australia to explore the parental perspectives on the EI

message, the researchers found that much must be learned about how EI is communicated to parents of children with special needs [15]. In addition, it is essential to enhance EI providers' knowledge about the family's education background, and how to discuss information. Family-centered care and the involvement of families as partners in their child's care are enhanced when accurate and succinct information is provided in the right time and in the right way [2]. Parents' greatest obstacles relating to EI services were lack of information and depending on untrustworthy internet sources to get information [38]. It is worthwhile to mention that EI providers giving insufficient information about the disability and the referral system was a common complaint among parents of children with ASD in Saudi Arabia [38]. Furthermore, there was a discrepancy when researchers wanted to measure family satisfaction with EI services, because parents reported differently based on their level of education [14] [39]. One study indicated that the higher the level of parent's education the greater their satisfaction [14]. Leite and Pereira (2013) indicated to the contrary: educated parents have higher expectations of professionals and services, and EI providers should take the differences in education levels into consideration. Parents of higher education level require more attention from EI professionals to meet their needs. EI providers must take care with delivery information methods, the quality and quantity of services provided to support the family in acquiring self-confidence and in learning new concepts [14].

Parent's learning style. Systematic review and recommendations by Flippin and Crais in 2011 highlighted that attention should be given to the father's style of learning. Men tend to learn by practice and by receiving comments from their peers rather than from the EI providers. Men also prefer to learn independently in a competitive learning environment [25]. Flippin and Crais (2011) provided many recommendations for teaching fathers. For example, give father appropriate learning strategies such as peer feedback. In other words, agencies should provide information in an amicable manner that contributes significantly to involving fathers in interacting with their children [25]. Further research would be beneficial that focuses on how the EI providers should communicate to parents and explore strategies and evidence-based practices to support the family need for information. Parents need support from EI providers in striking a balance between reality and hope [15].

Understanding Legal Aspects

EI services require family involvement and agreement on the child's plan. In many cases, parents played the role of advocates for their child for multiple issues [30]. Parents may need a full understanding of the legal system or demonstrate a greater understanding of federal and state statutory law, as well as their eligibility for IDEA, IFSP, and an IEP to support their readiness to participate in decision-making for their child's future [1]. EI professionals who provide services to the child and the family must inform parents about the child's needs, caregiving issues, and intervention goals along with special education laws, legislation, and regulations relating to the child's rights [40] [41].

Moreover, Zhenga, Maude, Brotherson, and Merritts (2016) found an interesting result in a qualitative study conducted in China to explore the experiences of families of children with disabilities receiving EI services. Researchers indicated highly significant negligence by EI specialists to raise awareness of policies and implementation. Parents used phrases such as "We didn't know ... we never knew that," which indicated unawareness of policies and resources related to children with disabilities that resulted in a delay in joining EI services [42]. Many parents in China advocated for the government to develop EI services. Rather than receiving

rehabilitation services, parents wanted trained providers to work with young children [42]. Similarly, there are clear indicators among Saudi parents that EI programs have been progressing toward comprehensive involvement [38]. Yet parents are still demanding as advocates of their children to have better services, especially regarding issues related to specialists and legal systems. Thus, close attention should be taken to improve professional development for EI providers in order to develop EI service systems [38].

Effective strategies to be advocate. Bruns and LaRocco (2017) recommended some practices for EI providers to help them eventually enable families to assume the role of advocate. Listening actively, asking open-ended and clarifying questions to better understand a family's concerns and showing empathy and demonstrating respect are effective approaches [2]. Other approaches include developing parental skills in identifying and prioritizing needs, problem solving, and co-planning. These are strategies to promote families to become better advocates for their children [2]. In this regard, Kim and colleagues in 2014 add that the role of service providers is to educate parents and provide them with adequate information in order to prepare them for the next stage in their life, in particular, the transition period. Knowledge of special education law terms such as least restrictive environment would help parents to claim the most beneficial placement for their child [17]. Interestingly, ignorance of certain laws or regulations may lead to losing opportunities for the child. For example, Mattern (2015) conducted a study in Pennsylvania to determine parental knowledge of evidence-based practices. Many parents indicated that little was known about the possibility of bringing an advocate to IEP meetings, and that information was not directly communicated to them by EI service providers, so that made them feel uncertain [1]. Parents want to have the information they need to be advocates for their child. Generally, parents lack the information, resources, and support that is necessary for raising children and addressing their developmental concerns [17].

Coaching

A large number of research studies have included evidence-based practices that have proven effective for child development [43] [25] [44]. The child from birth to preschool spends considerable time with parents and because of the unique needs of children with exceptional abilities, parents of these children have an increasing need to acquire these practices. Hence, supporting parents to become coaches and transfer what they have learned to their child at home is crucial [45]. Parents can be the best teacher for their children. Thus, training parents in some strategies that they can implement in the home environment would be beneficial for child development. A case study in Turkey aimed to investigate the opinions of mothers who have a child with Autism [46]. The study indicated that the greatest source of educational support during the diagnoses and arrangement of special education services was by other families with Autistic children. The researchers pointed out that parents preferred to follow the same path as the parents who were in the same situation. From the mothers' perspective, they value the expertise of other parents as a critical resource for learning about their children [46].

Mentoring methods. Related research suggests some methods have been evaluated as effective for educated parents [37] [45] [17]. For example, didactic sessions, observation, applied behavior analysis, role-play and rehearsal, video modeling performance feedback, evaluation and reflection are all effective [37] [45] [17]. Furthermore, training home visits were clearly helpful and meaningful due to meeting the parents' needs in better understanding and acting upon the child's condition [35]. Besides home visits, Responsive Teaching (RT) is another method of coaching

parents, which is a fully manualized relationship-based intervention. The purpose of RT is to teach parents to use incidental behavioral teaching strategies to help children learn targeted developmental behaviors and skills. RT strategies have proved as effective practices to support the families in expanding interactions and to have more family daily routine activities [43]. To sum up, parent-coaching strategies can help practitioners utilize family-centered practices, which is a very important principle of EI service delivery.

Home Visits to Support Family Members

EI personnel address the needs of the child, and they are responsible for supporting family structure [30]. Home visits provide a great solution for mothers with more than one child. Raising other siblings and life circumstances may hinder the mother's involvement in centers or clinics. In addition, home visits seem to be a remedy for that dilemma [30]. Home visits have been described as a convenient and supportive process for the family so that the family is not required to arrange any meetings or make any preparations. More importantly, it offers a wonderful opportunity to include all the family members [35]. Canada's parents in 2014 shared stories about their perspective about the effectiveness of EI services. Pighin and colleagues found that parents greatly appreciated the EI providers' effort in involving the sibling in the process of educating and interaction with their brother/sister with special needs. EI providers during the home visit participated in training all members of the family to encourage the family activities. Parents felt that including sibling in the EI sessions contributes to further strengthening of family relations [35]. Siblings' involvement was also mentioned in a study by Cummings, Hardin and Meadan (2017) conducted in rural communities. The majority of parents expressed content with their children having mainly sibling playmates. As a result, this emphasizes the importance of siblinghood in the social and linguistic interaction of the child with special needs. Thus, many EI play activities should involve siblings [40] [47].

The Barriers for Family's Participation in EI

There are a number of factors that influence the degree to which parents can be involved in their child's development. Parents seeking to enroll their children in EI programs may face numerous issues [48]. These obstacles may be due to the family structure, the economic status, or may be due to external factors such as geographical location, migration, etc. [48]. These barriers may need to be immediately addressed when the family tries to help their child but families often face a shortage of specialists that results in a lack of access to EI services [48].

Economic Status

Poverty. Raising an exceptional child may affect family well-being. Consequently, poverty in childhood is more common among children with disabilities than among children without disabilities. Fujijura and Yamaki (2000) reported that 28% of U.S. children with disabilities lived below the Federal Poverty Level (FPL), as contrasted with 16% of children without disabilities [49]. A more recent study by Addy, Engelhardt, and Skinner (2013) reported that from 2006 to 2011, the overall percentage of infants and toddlers living in low-income households in the United States increased from 44% to 49%. Parental involvement in children's EI programs may vary due to economic levels. Existing research points out that family problems associated with financial barriers are widely common [50] [51] [52]. According to Afolabi (2014), poverty and illiteracy hinder parental involvement in the process of educating their child alongside with other educators [51]. Excessive sensitivity due to indigence and ignorance of the educational terms used by

educators may lead parents to feel that their input is not welcome resulting in the reluctance to participate in their children's education [51]. With regard to the impact of poverty on the quality of the parents caring for their child, the results of a study conducted by Longtin and Principe (2016) with urban African American parents of children with ASD indicate that there was a relationship between poverty and the parents' level of knowledge. Parents below the poverty level have less knowledge and awareness of evidence-based interventions and their child's program of service compared to parents above the poverty level [52]. One of the useful strategies that appeared in Corr, Santos, and Fowler's (2015) review of literature to support poor families for coping with their autistic child was to use daily routines [53]. Furthermore, the researchers also pointed out some of the obstacles for using this routine. Poor families face many barriers to achieving such routines; for example, they have to move between shelters, experience a consistent low income, and face problems accessing food [53]. According to Swafford, Wingate, Zagumny, and Richey (2015), a qualitative study was conducted on 17 families in Tennessee who live near or below the federal poverty level, to study families' experiences with EI services. The results indicate that two-parent families were satisfied with the services offered, while single mothers described particular services, such as referral processes, as barriers that affected their children's outcomes [54]. The complication of referral processes in Part C was also mentioned in the Corr and colleagues study [53].

Health insurance. Lack of medical insurance or not getting a full insurance plan that covers all the medical and therapeutic needs of the child was found to be a barrier for parents. For instance, focus study groups investigated the barriers of using therapy services for infants, toddlers, and preschoolers with disabilities and found that families might have a lack of financial resources [55]. One of the reasons for not using a therapy service was the high cost of the therapy sessions [55]. This is in line with what the research by Alotaibi and Almalki (2016) and Zhenga and colleagues (2016) suggested. It is clear that Saudi and Chinese parents need more financial support from governments to cover the cost of obtaining the best EI services for their children with special needs [38] [42]. Pearson and Meadan (2018) also pointed this out in Midwestern state. They found that mothers were more satisfied when their children were covered under state plans as opposed to private insurance. The mothers also made it clear that the negotiations with insurances companies wasted time due to waiting for an approval. In addition, the beginning of the year often lead to paying at their own expense instead [23].

Employment. There are many factors that have an influence on a family's economic status. These factors include the difficulty in obtaining or keeping a job, or even working for long hours to provide the child's needs. Indeed, compared with parents who have children without disabilities, parents who have children with disabilities have lower rates of employment [56]. These difficulties were expressed by parents in many studies, and described as obstacles to their participation in EI programs [28] [23] [56]. Along the same lines, Campos's (2008) found that family involvement in early childhood programs among fathers in immigrant Latino families, indicated that logistical barriers, such as multiple jobs and long work hours were snags in the father's way for more involvement compared to white families' participation [28]. Likewise, working African American mothers expressed the lack of time to implement EI strategies except on the weekend [23]. Mothers were dissatisfied with the fact that the jobs take a long time and effort by the mother, which hinders her from performing her primary duty of taking care of her child [23]. There is another matter to consider: working for a certain number of hours per week to meet and achieve the state or federal conditions in order to receive the federal assistance negatively affects parents' lives [53]. In other

words, working for long hours to improve the family's finances in order to maintain federal assistance have consequences. Long job's hours result in difficulty in scheduling appointments for the child to receive EI services [53]. An additional hindrance was the long distances to reach work among working parents who live in rural communities as the travel time took time away from their involvement with their children [40]. Working at a faraway job had a negative influence on the kinds of activities the rural parent participated in with their children [40].

Geographical Location

One of the dilemmas related to parental involvement in EI service was reaching EI service center locations, especially for parents who live in villages and remote communities [38] [40] [42]. In a study exploring whether ecocultural features were enhanced or prevented parents in rural communities in North Carolina from involvement in EI services, the results were differentiated participation [40]. Some parents indicated that they had to travel far distances to access specialty services and that traveling faraway was generally a constraint. Conversely, some parents pointed out that access to services took 20 or 30 minutes and did not negatively affect their participation [40]. Again, Chinese and Saudi parents have full agreement about their suffering with limitation of services at limited locations. Some families from rural areas needed to travel long distances to get services and supports for their children; subsequently, parents demanded their governments to increase centers in the countryside [38] [42]. The difficulty of access to services is not only related to rural communities, but also the EI services' location within cities for poor parents may also be a challenge [52]. Families who live in poverty often live in poor neighborhoods with few resources, and due to transportation costs that might affect their participation in EI services [38] [53] [34]. Finally, Mattern (2015) stated that families moving from one state to another might become a barrier due to the differences in systems and policies.

Shortage of Specialists

Several studies reported that EI services did not well meet the parents' needs and that was because of lack of competencies [38] [55]. Staffing issues and shortages of occupational therapists, pediatric therapists, physical therapists, and speech-language pathologists were documented in multiple studies [38] [55]. According to Anthony (2014), "There is a national shortage of teachers of students with visual impairments and O&M specialists in the United States, and not all of the existing personnel have the knowledge and skills they need to adequately meet the needs of birth-through-age-5 learners and their families" (p. 516) [57]. Furthermore, EI providers may have little-specialized training. For example, in a study by Little, Kamholz Corwin, Barrero-Castillero, and Wang (2015), the researchers identified gaps in the EI system. However, the results indicated that parents were displeased with the delay in obtaining services after referral. In addition, EI providers expressed their uncertainty due to lack of experience as they were young clinicians who had newly graduated [58]. Similarly, Brotherson and his colleagues (2016) pointed out that the providers expressed fatigue due to the increase in workload and the number of work hours, which hindered them from achieving their best performance when working with families [21].

Language Barrier

Speaking a language other than the EI provider's language seems to be a substantial complexity surrounding accessing appropriate services for non-English speaking families [28] [34] [31] [48]. For instance, although the Spanish language is the second language used in the United States, it tends to be a factor barring Spanish families from reaching EI services. In 2013, Williams and

colleagues investigated barriers to accessing services for children under the age of three years presenting with language delays and behavioral difficulties. Phone calls were made in the Spanish language, to seek an appointment for a 2-year-old child, with 30 agencies who provide early intervention services for children. The findings indicated that less than half of the callers were able to set up appointments in order to discuss the child's condition, resulting in considering language as a barrier to accessing services [48]. As previously mentioned, communication between parents and agencies is the key for successful EI programs. Thus, immigrant families face additional barriers to accessing and engaging in developmental services for their child, including language, culture, and varied perceptions [34]. Lack of English-speaking skills, especially among immigrant parents, was a clear and prominent issue [28] [34] [31]). In addition, there was a consensus among immigrant parents regarding the difficulties in sharing information with EI providers in an academic and professional manner, which led to extreme frustration due to their inability to do so [30]. There was a notable result in the Alotaibi and Almalki (2016) study showing that two-way communication was not the only limitation for accessing services. Parents in Saudi Arabia referred to language as a hindrance to obtaining information about their autistic child's development. Most of the sources of information on the internet are in English. Again, the language was a barrier to parents benefiting from this information [38].

Cultural Background

Many families face challenges in obtaining adequate services for their children. Additionally, obstacles related to cultural differences frequently appear among parents [34] [31]. Cultural misunderstandings may lead to losing many developmental opportunities for children who need EI services. The main potential barrier to enrollment in EI programs is lack of understanding and cultural sensitivity among EI providers, which leads to ineffective response and support of the family's needs [34] [31]. Researchers suggest that vulnerable families, who may benefit most from EI, sometimes decline services because of social concerns, stressors, or wariness about home visits [58].

Stereotypes. There is continued concerns about the academic achievement of a child under supervision of a migrant family, there remains a persistent assumption that parent's lack education may reflected in neglecting the importance of education, so cannot be involved in their children educational development [23] [59]. According to one study finding demonstrate that parents whose cultural background is different from the majority may exposed to discrimination in the school system [31]. For example, in Starr et al. 's (2016) work on ethnically diverse parents' experience during the transition, wearing a hijab was a reason for the teachers to neglect the mother based on the assumption that she did not know what was best for her child [31]. According to Pearson and Meadan (2018), African American parents might be treated differently and have fewer chances to schedule an appointment with specialists because of stereotyping, discrimination, and racism.

Social norms. Some hindrances to parents for enrolling in EI services were due to incorrect beliefs about disabilities, stigmatization associated with EI programs, and fear of jeopardizing immigration status, especially among those who are undocumented [34]. The parents' perspective and reluctance to identify disability in many cultures due to associated stigma, often leads to a delay in enrolling in EI services [22]. To what extent the practices are appropriate for other environments should be taken into consideration. Studies highlight that some strategies may not be compatible with the prevailing parental methods in some cultures, and thus impede the implementation of some practices used in EI programs [43] [28]. For example, in the Arab and

Latino cultures, parents tend to be strong regulating authorities restricting their children's autonomy and rarely using affection or praise. This is contrary to the principles of reinforcement and responsive interaction with the child [43] [28]. Furthermore, cultural patterns may play a major role as barriers to providing the best services for the child. For example, Starr (2016) pointed out that in some cultures, the mother is not allowed to communicate directly with the EI specialists, so the father plays the role of mediator, which results in the omission of some essential information that may significantly affect certain decisions regarding the child.

Beliefs. Misconceptions about disability are common in some cultures. Studies indicate that cultural, religious, and folk beliefs might have an influence on parents' decisions. These beliefs are attributed to extrinsic factors, such as God's will and spirituality. In addition, some parents might ask for advice from others such as grandmothers, friends, and those who do not have the appropriate experience for diagnosis or who cannot provide specialized assistance [22] [34] [23]. Steeley and Lukacs (2015) highly recommended further studies to more deeply investigate cultural barriers.

Recommendations

This section of the review will cover some recommendations for EI providers and parents of children with disabilities in order to implement high-quality EI services. Researchers have made suggestions that build on findings, such as what needs parents have, support types that are meeting the needs of diverse families, and barriers or challenges that limit families' access to services for at-risk or disabled children. Ultimately, this would enable decision-makers to understand barriers and thus seek to develop appropriate solutions. The following specific recommendations related to these themes suggest promoting positive relationships between families and agencies that position parents and EI providers as collaborators and key influencers in facilitating each child's development.

Recommendations at the local and state level for EI providers:

- Activate strategies that aim to reduce high rates of stress and depression among parents, such as social support groups to facilitate interaction among parents who are facing similar challenges.
- Developing effective communication skills is one of the most important elements of the cooperation process. All parties should strive for this kind of cooperation and be able to express thoughts in an encouraging, respectful, and friendly environment.
- More attention should be given to considering the unique characteristics, circumstances, and background of families with children with additional needs.
- Train professionals to promote more acceptance and understanding of cultural differences and avoid prejudice so that the providers can become better qualified to apply EI practices in a greatly an effective way.
- For diverse families and those who do not speak English, offer low-cost/free bilingual services in the home in addition to providing information in writing or visual form. Having a translator would help parents better understand information, which could increase their involvement. Home visits should occur regularly and continually, especially for families living in

smaller urban or in rural locations, due to the many benefits including solutions for working parents and training siblings on EI techniques.

- Educate parents on the IFSP and IEP process and other legal aspects through holding mini-seminars or workshops that use factual language and well-explained terminology.

Recommendations at the federal level for policy makers:

- The policy-makers should make concerted efforts to support parents' financial resources that help overcome the effects of poverty, thus enhancing EI services.
- To solve the problem of the shortage of specialists, increased funding for EI would allow agencies to offer salaries that are more competitive and hire experienced providers.
- Increase public funding of awareness activities to inform the public of EI services. Networks and nongovernmental organizations could hold informational support campaigns on the legal rights of special needs children and their families.

Recommendations for future research:

- Future research should focus on facilitating the transition plan and looking at the relationship between parental involvement and understanding parents' role as members who have a voice in the IEP and can help identify the least restrictive environment.
- Researchers should focus on addressing immigrant issues, especially thoughts related to stigmatization associated with EI programs, because that would encourage parents to seek out options, services, and possibilities. Weigh the options and make decisions with trusted professionals.

Summary

Parents' knowledge of normative child development is a multidimensional construct comprised of parents' knowledge about developmentally appropriate practices, conventional child development, and maintenance and promotion of child health. Overarching themes emerged from the literature review that documents the needs of parents as the expressed desire for services to be obtained or outcomes to be achieved. Next identifying barriers that should be addressed, a number of barriers for parents attempting to obtain services for young children with disabilities or developmental delays were explained. Finally, this author made suggestions for researchers, EI providers, and policymakers to improve enrolment in EI services as well as developmental outcomes for young children at risk for or with developmental delays. Parents of children with disabilities face a range of social and psychological problems. Given that, a cornerstone of the Part C policy of the IDEA is a focus on the entire family so they will have access to appropriate services that support their needs. In order to reach a high level of family involvement in a child's IEP, many steps should be applied. For example, listen carefully to the parents' questions and respond to them in a realistic and non-exaggerated manner. Ultimately, effective communication is the key for successful EI. Furthermore, EI providers should be aware of the benefits of the partnership approach and learn about how to best support parents to help their children. Finally, home visits allow therapists to gain an accurate picture of the child's delays, behaviors, progress, and at the same time help parents and siblings to apply EI practices in more natural way. Parental

involvement in EI programs has a significant impact on children's lives. Understanding and accommodating parents' needs will support them in this vital aspect of their child's development.

References

- [1] J. A. Mattern, "A mixed-methods study of early intervention implementation in the Commonwealth of Pennsylvania: Supports, services, and policies for young children with developmental delays and disabilities," *Early Childhood Education Journal*, Vols. (1), 57-67, no. 43, 2015.
- [2] D. A. Bruns, . D. J. LaRocco, O. L. Sharp and K. M. Sopko, "Leadership competencies in US early intervention/early childhood special education service systems.," *Infants & Young Children*,, vol. 30, no. 4, pp. 304-319, 2017.
- [3] J. Brooks-Gunn, F. R. Liaw and P. K. Klebanov, "Effects of early intervention on cognitive function of low birth weight preterm infants.," *The Journal of pediatrics*,, vol. 120, no. 3, pp. 350-359., 1992.
- [4] "Part C State Performance Plan/Annual Performance Report 2018 Indicator Analyses," *U.S. Department of Education*, 2018.
- [5] D. B ,Bailey Jr, and P. M. Blasco, , "Parents' perspectives on a written survey of family needs.," *Journal of Early Intervention*, vol. 14, no. 3, pp. 196-203, 1990.
- [6] J. Kandel, J Merrick,, "The birth of a child with disability. Coping by parents and siblings.," *TheScientificWorldJOURNAL*, 3,, no. 3, pp. 741-750, 2003.
- [7] B. Poon, and A. Zaidman-Zait, , "Social support for parents of deaf children: Moving toward contextualized understanding.," *Journal of Deaf Studies and Deaf Education*,, vol. 2, no. 19, pp. 176-188.2014 ,.
- [8] L. L. . Dyson, " Fathers and mothers of school-age children with developmental disabilities: Parental stress, family functioning, and social support.," *American journal on mental retardation*, , vol. 102, no. 3, pp. 267-279., 1997.
- [9] R. Tway, P. Connolly, and J. M. Novak, , "Coping strategies used by parents of children with autism.," *Journal of the American Academy of Nurse Practitioners*, vol. 19, no. 5, pp. 251-260., 2007.
- [10] L. Osborne, , L. McHugh, , J. Saunders, and P. Reed, , "Parenting stress reduces the effectiveness of early teaching interventions for autistic spectrum disorders," *Journal of autism and developmental disorders*,, vol. 38, no. 6, pp. 1092-1103., 2008.
- [11] B. Ingersoll, and D. Z. Hambrick, , "The relationship between the broader autism phenotype, child severity, and stress and depression in parents of children with autism spectrum disorders.," *Research in Autism Spectrum Disorders*,, vol. 5, no. 1, pp. 337-344., 2011.
- [12] D. Mandell, and M. S. Salzer, , "Who joins support groups among parents of children with autism?.,," *Autism*, vol. 11, no. 2, pp. 111-122., 2007.

- [13] M. Meirsschaut, , H. Roeyers, and P. Warreyn, , "Parenting in families with a child with autism spectrum disorder and a typically developing child: Mothers' experiences and cognitions.," *Research in Autism Spectrum Disorders*,, vol. 4, no. 4, pp. 661-669., 2010.
- [14] Y. Pang, "Facilitating family involvement in early intervention to preschool transition.," *School Community Journal*, vol. 20, no. 2, pp. 183-198., 2010.
- [15] A. Edwards, , C. Brebner, P. McCormack, and C. MacDougall "The early intervention message: Perspectives of parents of children with autism spectrum disorder.," *Child: Care, Health and Development*,, vol. 43, no. 2, pp. 202-210., 2017.
- [16] N. Miodrag, and . R. M. Hodapp "Chronic stress and health among parents of children with intellectual and developmental disabilities.," *Current opinion in psychiatry*,, vol. 23, no. 5, pp. 407-411., 2010.
- [17] J. Kim,, N. Cavaretta, and K. Fertig "Supporting Preschool Children with Autism Spectrum Disorders (ASD) and Their Families " *Journal of the American Academy of Special Education Professionals*, Vols. 87,, no. 93., 2014.
- [18] J. Resch, T. Elliott, and M. R. Benz "Depression among parents of children with disabilities," *Families, Systems, & Health*,, vol. 30, no. 4, p. 291, 2012.
- [19] M. Meirsschaut,, H. Roeyers, and P. Warreyn "Parenting in families with a child with autism spectrum disorder and a typically developing child: Mothers' experiences and cognitions.," *Research in Autism Spectrum Disorders*,, vol. 4, no. 4, pp. 661-669., 2010.
- [20] J. Greenberg, J. S., . M. M. Seltzer,, M. W. Krauss, , R. J. A. Chou, and J. Hong "The effect of quality of the relationship between mothers and adult children with schizophrenia, autism, or Down syndrome on maternal well-being: The mediating role of optimism," *American Journal of Orthopsychiatry*, vol. 74, no. 1, pp. 14-25, 2004.
- [21] M. Brotherson, , J. Summers,, L. Naig, , K. Kyzar, , A. Friend, , P. Epley, and A. Turnbull "Partnership patterns: Addressing emotional needs in early intervention.," *Topics in Early Childhood Special Education*, vol. 30, no. 1, pp. 32-45., 2010.
- [22] E. Ijalba, "Hispanic immigrant mothers of young children with autism spectrum disorders: How do they understand and cope with autism?," *American Journal of Speech-Language Pathology*, , vol. 25, no. 2, pp. 200-13., 2016.
- [23] J. Pearson ,. H. Meadan "African American parents' perceptions of diagnosis and services for children with autism.," *Education and Training in Autism and Developmental Disabilities*, vol. 53, no. 1, pp. 17-32., 2018.
- [24] K. Decker and ,. C. D. Vallotton "Early intervention for children with hearing loss: Information parents receive about supporting children's language.," *Journal of Early Intervention*, vol. 38, no. 3, pp. 151-169., 2016.

- [25] M. Flippin, and E. Crais "The need for more effective father involvement in early autism intervention: A systematic review and recommendations.," *Journal of Early Intervention*,, vol. 33, no. 1, pp. 24-50., 2011.
- [26] B. McBride, , S. Curtiss, , K. Uchima, , and . D. Laxman,, "Father involvement in early intervention: Exploring the gap between service providers' perceptions and practices.,," *Journal of Early Intervention*, vol. 39, no. 2, pp. 71-8, 2017.
- [27] T. G. Mueller, and . P. C. Buckley,, "Fathers' experiences with the special education system: The overlooked voice.," *Research and Practice for Persons with Severe Disabilities*,, vol. 39, no. 2, pp. 119-135., 2014.
- [28] R. Campos, " Considerations for studying father involvement in early childhood among Latino families.," *Hispanic Journal of Behavioral Sciences*,, vol. 30, no. 2, pp. 133-160., 2008.
- [29] L. Fox, , G. Dunlap, and L. Cushing, , "Early intervention, positive behavior support, and transition to school.," *Journal of emotional and behavioral disorders*, vol. 10, no. 3, pp. 149-157., 2002.
- [30] , J. Marshall, , A. Adelman, , . S. M. Kesten,, R. A. Natale, and B. Elbaum, , "Parents' experiences navigating intervention systems for young children with mild language delays.," *Journal of Early Intervention*,, vol. 39, no. 3, pp. 180-198., 2017.
- [31] E. M. Starr,, . T. S. Martini, and . B. C. Kuo,, "Transition to kindergarten for children with autism spectrum disorder: A focus group study with ethnically diverse parents, teachers, and early intervention service providers.," *Focus on Autism and Other Developmental Disabilities*,, vol. 31, no. 2, pp. 115-128, 2016.
- [32] L. Fontil and H. Petrakos , "Transition to school: The experiences of Canadian and immigrant families of children with autism spectrum disorders.," *PsycholSch.*, vol. 52, p. 773–788, 2015.
- [33] S. K. Stuart,, L. D. Flis, and . C. Rinaldi,, "Connecting with families: Parents speak up about preschool services for their children with autism spectrum disorders.," *Teaching Exceptional Children*,, vol. 39, p. 46–51, 2006.
- [34] L. Karoly, and G. Gonzalez, , "Early care and education for children in immigrant families.," *Future of Children*, vol. 21, no. 1, pp. 71-101., 2011.
- [35] M. Pighini,, H. Goelman, , . M. Buchanan, and R. Schonert, "Learning from parents' stories about what works in early intervention. International Journal of Psychology:," *Journal International De Psychologie*,, vol. 49, no. 4, pp. 263-70, 2014.
- [36] B. Rous,, . C. T. Meyers, and S. B. Stricklin,, "Strategies for supporting transitions of young children with special needs and their families.," *Journal of Early Intervention*,, vol. 30, p. 1–18, 2007.
- [37] Y. Akamoglu, and L. Dinnebeil "Coaching parents to use naturalistic language and communication strategies.," *Young Exceptional Children*, vol. 20, no. 1, pp. 41-50., 2017.

- [38] F. Alotaibi, N. Almalki "Parents' perceptions of early interventions and related services for children with autism spectrum disorder in Saudi Arabia.," *International Education Studies*, vol. 9, no. 10, pp. 128-140., 2016.
- [39] C. S. Correia Leite, and . A. P. Da Silva Pereira,, "Early intervention in Portugal: family support and benefits.," *Support for Learning*,, vol. 28, no. 4, pp. 146-153., 2013.
- [40] K. Cummings, , B. Hardin, and . H. Meadan, , "Understanding parental engagement in early learning opportunities for families in rural communities.," *Rural Special Education Quarterly*, vol. 36, no. 3, pp. 116-127., 2017.
- [41] G. Mahoney, and . J. Filer,, "How responsive is early intervention to the priorities and needs of families?," *Topics in Early Childhood Special Education*,, vol. 16, no. 4, pp. 437-57, 1996.
- [42] Y. Zheng, , S. Maude, , M. Brotherson and A. Merritts, "Early childhood intervention in china from the families' perspective.," *International Journal of Disability, Development and Education*,, vol. 63, no. 4, pp. 431-449., 2016.
- [43] T. Aquraini, , . A. Al-Odaib,, . H. Al-Dhalaan,, H. Merza, and G. Mahoney, , "Relationship-based intervention with young children with autism in Saudi Arabia: Impediments and consequences of parenting stress and depression.," *International Journal of Disability, Development and Education* , vol. 66, no. 3, pp. 233-24, 2016.
- [44] P. A. Gallagher,, C. A. Rhodes, and S. M. Darling, , " Parents as professionals in early intervention: A parent educator model.," *Topics in early childhood special education*,, vol. 24, no. 1, pp. 5-13, 2004.
- [45] W. Machalicek, , R. Lang, and T. Raulston, , "Training parents of children with intellectual disabilities: Trends, issues, and future directions.," *Current Developmental Disorders Reports*,, vol. 2, no. 2, pp. 110-118., 2015.
- [46] N. Kayhan, and L. Ozaydin, , "The quality of mainstreaming in preschool: The views of parents of children with Down syndrome.," *Journal of Education and Learning*,, vol. 7, no. 2, pp. 157-173., 2018.
- [47] J. Marshall, , . J. Goldbart, and . J. Phillips,, "Parents' and speech and language therapists' explanatory models of language development, language delay and intervention.," *International Journal of Language and Communication Disorders*, vol. 42, no. 5, pp. 533-555, 2007.
- [48] M. E. Williams, . J. L. Perrigo,, T. Y. Banda, , T. Matic, and ,. F. D. Goldfarb, "Barriers to accessing services for young childre," *Journal of Early Intervention*, vol. 35, no. 1, pp. 61-74., 2013.
- [49] G. Fujiura, and K. Yamaki, , "Trends in demography of childhood poverty and disability.," *Exceptional Children*,, vol. 66, no. 2, pp. 187-99., 2000.
- [50] S. Addy,, . W. Engelhardt, and C. Skinner, , "Basic facts about low-income children: Children under 18 years.," National Center for Children in Poverty, New York, NY:, 2013.

- [51] O. Afolabi, " Parents' involvement and psycho-educational development of learners with special educational needs (SENS): An empirical review.," *International Journal of Early Childhood Special Education*, , vol. 6, no. 2, pp. 177-203., (2014).
- [52] S. E. Longtin, and . G. M. Principe,, "The relationship between poverty level and urban African American parents' awareness of evidence-based interventions for children with autism spectrum disorders: Preliminary data. Focus on Autism and Other Development," *Focus on Autism and Other Developmental Disabilities*, vol. 31, no. 2, pp. 83-91, 2016.
- [53] C. Corr, , R. Santos, and . S. Fowler,, "The components of early intervention services for families living in poverty: A review of the literature.," *Topics in Early Childhood Special Education*,, vol. 36, no. 1, pp. 55-64., (2016)..
- [54] M. D. Swafford,, . K. O. ,Wingate,, L. Zagumny, and . D. Richey,, "Families living in poverty: Perceptions of family-centered practices.," *Journal of Early Intervention*, 37(2),, vol. 37, no. 2, pp. 138-154., 2015.
- [55] R. A. McWilliam, , H. J. Young, and K. Harville, , "Therapy services in early intervention: Current status, barriers, and recommendations.," *Topics in Early Childhood Special Education*, vol. 16, no. 3, pp. 348-374., (1996)..
- [56] M. M. Seltzer, , . J. S. Greenberg,, , F. J. Floyd, Y. Pettee, and J. Hong, , "Life course impacts of parenting a child with a disability.," *Journal Information*,, vol. 106, p. 265–286., 2001.
- [57] T. Anthony, "Family support and early intervention services for the youngest children with visual impairments.," *Journal of Visual Impairment and Blindness*, , vol. 108, no. 6, pp. 514-519., 2014.
- [58] A. Little, , . K. Kamholz,, B. Corwin, , . A. Barrero-Casti and C. J. Wang, , "Understanding barriers to early intervention services for preterm infants: Lessons from two states.," *Academic Pediatrics*,, vol. 15, no. 4, pp. 430-8., 2015.
- [59] S. Steeley, and K. Lukacs, , "Cultural and linguistic diversity and special education: A case study of one mother's experiences.," *International Journal of Special Education*,, vol. 30, no. 2, pp. 20-31., (2015)..

Principal Creative Leadership for Teacher Learning and School Culture

Yashi Ye

Abstract—Principles play vital roles in shaping the school culture and promote teacher professional learning by exerting their leadership. In the changing time of the 21st century, the creative leadership of school leaders is increasingly important in cultivating the professional learning communities of teachers for eventually improving student performance in every continent. This study examines under what conditions and how principal creative leadership contributes to teachers' professional learning and school culture. Data collected from 632 teachers in 30 primary and middle schools in cities of Chengdu and Chongqing of mainland China are analyzed using structural equation modeling and bootstrapping tests. A moderated mediation model of principle creative leadership effects is used to analyze professional teacher learning and school culture in which the mediator will be school culture and the moderator will be power distance orientation. The results indicate that principal creative leadership has significant direct and indirect effects on teacher professional learning. The positive correlation among principal creative leadership, teacher professional learning and school culture is observed. Further model testing found that teacher power distance orientation moderated the significant effect of principal creative leadership on school culture. When teachers perceived higher power distance in teacher-principal relations, the effects of principal creative leadership were stronger than for those who perceived low power distance. The results indicate the "culture change" in young generation of teachers in China, and further implications to understand culture context in the field of educational leadership are discussed.

Keywords—Power distance orientation, principal creative leadership, school culture, teacher professional learning.

I. INTRODUCTION

THE global trend of educational reform in educational policy, practice, and systems (e.g., leadership, curriculum, teaching, and school environment) has pushed scholars to build better theories to understand how school reform policies and practices have evolved over the past 30 years to refine school systems[1]. To improve school outcomes by enhancing teacher learning, shaping school culture, and eventually promoting student outcomes, school nowadays communities of teacher professional learning are established widely in most schools in every country [2],[4]. Teacher professional learning, despite being largely linked with the concept of "professional learning community (PLC)", has been adapted from the Western context to mainland China with cultural differences but a similar essence [2],[3]. Based on a wide range of previous research, the leadership quality of principals is found to have a direct impact on constructing the internal activities and teacher relations inside the PLCs[1],[5],[6], the construction of which are

positively related to student performance[7]. However, the different leadership types of principals in constructing a school culture of cooperation and building an effective cooperative relationship between teachers may vary because of cultural contexts, which can have an evident effect on teachers perceiving principals and school leaders as positive resources, especially in terms of the concept "power distance orientation". In previous researches, there was a paradox about teachers perceiving power distance in school settings in China. But more researches showed a viewpoint that China is a society with significant power distance, and teachers would perceive more inequity in their relationship with principals. According to Huffman and Jacobson, principal transformational leadership is more effective than other leaderships in promoting teacher professional learning through their influence, even though they are not in the position of teachers' direct leaders [8]. The core of transformational leadership is to lead change [9] consciously. According to the researchers, the 'change' in the definition context of transformational leadership means a large transformation in levels of system structure, organization culture, and leader reputation [10]. But in fact, the application of transformational leadership may be unpractical as it is difficult to "catch the wind" of sharply change through routine principal training [11], and there have been many overlapping but contradictory studies on transformational leadership. On the other hand, Creative leadership, a developing concept after transformational leadership in school settings, has a similar context about its essence to "change" for organizations in this era, can be practiced and acquired [12] by school leaders, and enable the reconstruction of more collaborative school culture and development of teacher learning[13]. When considering the culture of high power distance in Asian countries, creative leadership is prevalent as it guides principals to build a strong school culture of collaboration and trust in teachers' shared vision, values, and attitudes without constraining their motivation and self-efficacy when they perceive principals having too much control over themselves in schools[14]-[16]. Contrary to creative leadership, principals usually directly intervene in teaching activities and dominate the PLCs[17]. The two evidences about the influence of principal instructional leadership on PLCs in China both indicate that the teacher learning can be hindered by the instructional leadership style of principals on teacher self-efficacy in the high-power-distance context of mainland China[1],[18].

Yashi Ye was with Johns Hopkins University, School of Education, MD 21218 USA. She is now with Learning Cheetah Educational Consulting & Research Company, Chengdu, SC 610041 China (e-mail: yeyashi@learning-cheetah.org)

The proposed ideas are organized in figure 1. And the author presents three hypotheses to explore the relationships among four variables with the teaching age as a control variable. The three hypotheses are:

1. What effects of principal creative leadership on school culture and teacher professional learning?
2. Dose the level of power distance in teacher-principal relationships influence the effect of principal creative leadership?
3. Does the level of teacher power distance orientation in teacher-principal relationships moderate the effects of principal creative leadership on teacher professional learning through school culture?

In this study, data were obtained from 637 teachers at 30 elementary and secondary schools in the Chinese cities of Chengdu and Chongqing. The "conditioned indirect effects" (moderated mediation) of principal creative leadership on teacher professional learning were examined using structural equation modeling. Meanwhile, school culture is a mediator, and power distance is a moderator.

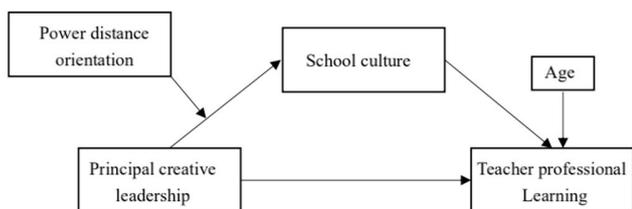


Fig.1 Proposed moderated mediation model connecting principal creative leadership to school culture and teacher professional learning in China.

This study makes three additions to a rising body of researches that investigated how principals' creative leadership encourages teachers in schools. First, the author describes how school culture operates through the effect of principal creative leadership on teacher perception, which fosters teacher learning in the workplace. Second, the study confirms the findings of prior researches that evaluated teacher power distance orientation as a significant moderating factor when principals exercise creative leadership [1], [18], [19]. Thirdly, the focus on power distance at the individual level of teachers illustrates how cultural orientation differences among teachers within China society might moderate the effectiveness of leadership exercises.

II. METHOD

A. Participants and Procedures

The author gathered information from teachers in 30 primary and middle schools in the Chinese cities of Chengdu and Chongqing. Participation in the study was entirely voluntary. Teachers submitted anonymous online questionnaires to the researcher via Questionnaire Star's survey link. In total, 632 of 637 valid online questionnaires were gathered, resulting in a response percentage of 127.4% when compared to the expected number of 500 questionnaires issued at the start. 52.53 % of the respondents are male, while 47.47% are female,

which is consistent with the general population of instructors in Chengdu and Chongqing (see Table 1). The average number of years spent teaching was 9.75 (not listed). 44.78% of the sample held a bachelor's degree, while 55.22% held a graduate degree. Lastly, 17.56% of respondents are senior professionals, 31.17% are first-grade professionals, and 51.27% are second/third-grade professionals or have no professional rank.

TABLE I
CHARACTERISTICS OF TEACHER SAMPLE AND POPULATION

	n	%
Teacher gender		
male	332	52.53%
female	300	47.47%
Teaching experience age		
less than 4	139	21.99%
4~7	217	34.34%
8~15	162	25.63%
16~23	79	12.50%
more than 23	35	5.54%
Teacher education		
Bachelor's degree	283	44.78%
Graduate degree	349	55.22%
Professional rank		
Second/third grade or no rank	324	51.27%
First grade	197	31.17%
Senior	111	17.56%

B. Measures

The variables involved in this study were all individual-level constructs. Each was rated on a Likert scale of 1-5. Almost Never to Almost Always was used to rate teacher professional development and school culture. 17 questions of principal creative leadership were answered Very Disagree to Very Disagree, and the remaining 12 questions were answered Almost Never to Almost Always. Two bilingual (English-Chinese) speakers back-translated all scale items into English to get similar meaning [19].

Mediating variable: school culture. The author created a 28-item school culture scale. The items were developed using Gruenert's school culture survey design principles[20] and Luo and Wang's original questions[21]. The full-scale reliability was $\alpha = 0.948$.

Control variable: teacher age. As a categorical control variable, the age (1=<3, 2=4-7, 3=8-15, 4=16-23, 5>24) was included. Teacher age was found to have a negative association with teacher professional learning in earlier studies in China, with younger teachers demonstrating stronger professional learning engagement than more senior teachers [1].

Moderator variable: power distance orientation. In the school settings, power distance orientation is described as teachers' acceptance that principals (and their supervisors) have authority to command their behavior and that power can be distributed unequally in teacher-principal relationships [22]. An nineteen-item scale was produced in this study using eight items from a

previous questionnaire created by Earley and Erez in 1997 ($\alpha=0.71$) and amended by Kirkmanto, Chen, Farh, Chen, and Lowe in 2009 ($\alpha=0.87$)[23],[24]. The author investigated teachers' perceptions of their principals' power distance orientation in this section of the questionnaire. The higher the score, the greater the power distance.

Independent variable: principal creative leadership. There are limited number of empirical researches about creative leadership strategies of K-12 school principals to promote teacher professional learning [25]. The author created a 27-item measure to collect data on a principal's innovative leadership. The full-scale reliability was $\alpha=0.949$ of the scale, which was derived from Obsborn's concept of Creative Problem Solving (CPS), which Parker elaborated on in 1971[27].

Dependent variable: teacher professional learning. The concept of teacher professional learning is derived from learning organization. Through collective learning, cooperation and interaction, teachers acquire new idea, skills and strategies of teaching. The author used 11 items from a scale developed by Kwakman and Schechter, and Qadach [1] and 8 items adjusted from the scale for teacher professional learning (TPL) created by Liu, Hallinger and Feng[28] ($\alpha=0.940$).

C. Analytical strategy

The author ran confirmatory factor analysis (CFA) on the suggested model prior to testing to establish construct distinctiveness. And the Statistical Package for the Social Sciences (SPSS) version 28.0 and Mplus 8.7 were used to conduct all analyses in this study[29].

To begin with, the measuring model was evaluated for its reliability and validity, followed by a structural model to determine the significance of the correlations between the four variables. SPSS 28.0 employed Pearson's correlation coefficients (r) to determine the relationship between principal creative leadership, teacher professional learning, school culture, and power distance orientation. Pearson's correlation coefficients of 0.68 to 1.0 were considered to be significant [30]. A correlation coefficient analysis was made to determine the strength of the correlations between the four constructs. Correlation coefficient analysis was used to determine the strength and direction of a linear relationship between two variable. The degree of correlation shows the strength of the association [31].

Additionally, structural equation modeling was employed to evaluate the research model and test the hypotheses. Mplus 8.7 was used to conduct the analysis. The method supports concurrent factor and path analysis. Using the results of factor analysis, the validity of the measurement model was examined. The structural model's fit was evaluated using t-tests and bootstrapping (generating 5000 sub-samples and applying individual sign changes).

III. RESULTS

A. Preliminary analyses

Correlation analysis was used to determine whether two variables are related. Correlations between individuals are

employed in this study, and the correlation matrix is provided in Table II.

TABLE II
MEANS, STANDARD DEVIATION, AND CORRELATIONS AMONG STUDY VARIABLES

Variable	TA	PCL	TPL	SC	PDO
Teaching age	--				
Principal creative leadership	0.055	($\alpha=0.949$)			
Teacher professional learning	0.031	0.952**	($\alpha=0.940$)		
School culture	0.042	0.951**	0.950**	($\alpha=0.948$)	
Power distance orientation	-0.031	-0.873**	-0.858**	-0.867**	($\alpha=0.895$)
M	2.450	3.598	3.645	3.605	2.876
SD	1.128	0.565	0.592	0.559	0.548

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed)

Table II contains descriptive statistics, bivariate correlations, and Cronbach's alphas for every variable. High levels of principal creative leadership (mean = 3.598), teacher professional learning (mean = 3.645), and school culture (mean = 3.605) are indicated by the mean scores. Notably, the mean scores indicate a lower level of inclination toward power distance (mean = 2.876). China was recognized as 'high power distance society' in earlier research by Hofstede[22]. The result in this study were consistent with the research conducted by Liu and Hallinger later in 2020[1]. In the broader society, China has undergone ideological change on greater tolerance towards society difference during the last two decades [32], and the evidences of teacher perceiving lower power distance in mainland China nowadays may be supported by China governmental policy[33] and the other research conducted by He and Zhang[34].

The variables' correlations were in the predicted direction. Notably, a correlation existed between principle creative leadership (PCL), teacher professional learning (TPL), and school culture (SC). Furthermore, power distance orientation (PDO) is inversely related to PLC, TPL, and SC.

All variables show an adequate level of internal consistency, with alphas greater than 0.8, indicating that the items representing the four components have a high level of internal consistency overall. Furthermore, CFA(Table III) demonstrated that the model fit the data well, indicating that the scales have strong construct validity.

TABLE III
MEANS, STANDARD DEVIATION, AND CORRELATIONS AMONG STUDY VARIABLES

Variable	χ^2	df	χ^2/df	RMSEA	CFI	TLI	SRMR
School culture	420.421	347	1.21	0.018	0.99	0.99	0.021
Principal creative leadership	376.110	324	1.16	0.016	0.99	0.99	0.020
Teacher professional learning	187.898	152	1.23	0.019	0.99	0.99	0.021
Power distance orientation	197.688	170	1.16	0.016	0.99	0.99	0.023

B. Model testing

The research questions were addressed in three interconnected steps. The author initially examined the partial mediation model to evaluate if school culture worked as a substantial mediator of the impacts of creative leadership on teacher professional development. Next, we examined whether power distance orientation affected the link between creative

leadership and school culture among principals. Finally, we examined whether power distance orientation affected the strength of the association between principal creative leadership and teacher professional learning, as mediated by school culture.

1. Testing The Mediated Effects Model.

As a mediator, school culture was used to test the hypothesized partial mediation model of principal creative leadership effects on teacher professional learning. $2/df = 0.667$, $p = 0.414$, RMSEA 0.001, CFI > 0.999, TLI > 0.999, SRMR = 0.004 were all confirmed by SEM. This shows the proposed model is statistically sound.

The next step was to quantify the indirect benefits of principle creative leadership on teacher professional development via school culture (see Figure 2).

The results show a statistically significant and large “direct effect” of principle creative leadership on teacher professional learning ($b = 0.530$, 95 percent CI [0.450, 0.610]). This suggests that administrators' innovative leadership had a moderate direct effect on teachers' professional learning practices.

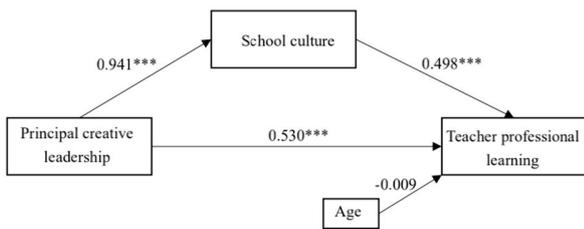


Fig.2 Unstandardized estimates of the path coefficients of the mediation model

The bootstrap analyses also demonstrated a statistically significant indirect effect of principal creative leadership on teacher professional development through school culture ($b = 0.469$, 95 percent CI [0.390, 0.544], not crossing zero). Determining whether school culture is a meaningful modulator of the impacts of leadership is the first study topic.

2. Testing Moderation Of Principal Creative The Mediated Effects Model.

The second research question explored if and how the power-distance orientation of teachers influenced the effect of innovative school leadership by principals. The purpose of this study was to determine, in a practical sense, whether a large or low power distance in social interactions between teachers and principals influenced the good effects of creative leadership on school culture.

Teachers' individual power distance orientation significantly moderated the positive connection between principal creative leadership and school culture ($b = 0.025$, 95% CI [0.005, 0.045], not crossing zero; $SE = 0.103$, $p = 0.017 < 0.05$) (see Table IV).

Table IV

UNSTANDARDIZED COEFFICIENTS FOR TESTING MAIN EFFECTS AND MEDIATING EFFECTS.

	School culture	Teacher professional learning
Principal creative leadership	0.805***[0.756, 0.854]	0.506***[0.437, 0.574]

Power distance orientation	-0.153***[-0.202, -0.105]	
Interaction: PCL * PDO	0.025*[0.005, 0.045]	
School culture		0.470***[0.402, 0.538]

95% bias-corrected CIs reported as: [Lower limit CI, Upper limit CI]. $n = 1194$, sample bootstrap = 5000, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

To facilitate interpretation of these interactions, the author further examined the relationships among the three variables included in this analysis: power distance orientation, principal creative leadership, and School culture.

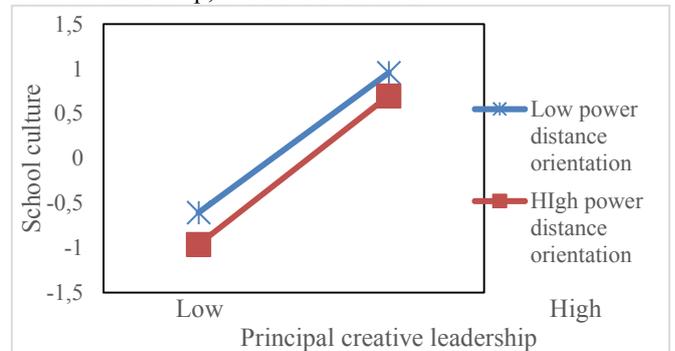


Figure 3. Interaction between principal creative leadership and power distance orientation with regard to school culture.

This analysis demonstrates(Fig. 3), that principal creative leadership has a stronger positive association with school culture when teachers perceive a greater power distance orientation. This indicates that when teachers experience a higher power distance in their relationships with principals, the principal's creative leadership has a more favorable effect on the school culture.

B.3 Testing for moderation of the partial mediation model.

The researcher picked a study using a similar methodology[1] as a guide to see if and how teachers' perceptions of power distance influenced or moderated the partial mediation model evaluated in the first process (see Fig.3). In addition, bootstrapping was utilized to assess the moderating effects of teacher power distance orientation on the mediation model at three degrees of power distance (M - SD, mean, and M + SD).

Higher levels of power distance in relationships between teachers and principals were associated with stronger indirect effects of principal creative leadership on teacher professional learning through school culture ($b = 0.390$, $SE = 0.040$, 95 percent CI [0.313, 0.468], not crossing zero). The bootstrapping results can be seen in Table V. When power distance orientation in teacher–principal relationships is lower, the indirect impacts of instructional leadership reduced ($b = 0.367$, $SE = 0.035$, 95% CI [0.296, 0.437], not crossing zero). The author concludes, accordingly, that the power distance orientation of teachers modifies the mediation model. Second, the results seem to indicate that a greater power distance between principals and teachers has a positive impact on school culture and teacher professional learning.

Table V.

CONDITIONAL INDIRECT EFFECTS OF PRINCIPAL CREATIVE LEADERSHIP ON TEACHER PROFESSIONAL LEARNING AT VARIOUS LEVELS OF MODERATOR POWER DISTANCE ORIENTATION.

	Effect	BootSE	BootLLCI	BootULCI
M - SD	0.367	0.035	0.296	0.437

M	0.378	0.037	0.305	0.450
M + SD	0.390	0.040	0.313	0.468

Note: 95% bias-corrected CIs reported as: [lower limit CI (LLCI), upper limit CI (ULCI)]

IV. DISCUSSION

With the ongoing researches digging into the field of educational leadership and creativity in a world with rapid change, school leaders, educational policymakers, and teachers have reached a consensus on the importance of the cultural context of schools in shaping the school culture and school leadership [35],[36]. Moreover, this study conducted in mainland China indicates that principal creative leadership have a strong indirect effect on shaping school culture and a medium direct effect on teacher professional learning. Moreover, the author examined the crucial cultural dimension of China context and found the teacher perception of power distance can moderate the effects of principal creative leadership on school culture and professional learning. In this final section, the researcher reviews the study's limitations and gives a interpretation of the findings.

A. Limitations of the study

First, the researcher developed some items of the previous scales. But the lack of fully developed scales in all dimensions linking to the four variables may produce bias results.

Secondly, as there is no maturity scale for principal creative leadership and one suitable for the cultural context of China, the online questionnaire was designed with 95 Likert-scale questions integrated with several scales from previous research. The large number of questions increases the reliability but may decrease the validity of the results.

Finally, only a single mediator, school culture is examined. Other mediators (e.g. organizational commitment, teacher trust, and teacher collective-efficacy) also need to be studied which can mediate the principal leaderships[38]-[40]. At the same time, other dependent variables could also be examined in order to guarantee the effectiveness of the study.

B. Interpretation and implications of the findings

Based on previous studies, the researcher found few empirical studies about creative leadership in the educational field. [38]. The author started this investigation with a school leadership mediation model whose usefulness has been proven in previous studies. The findings supported the validity of a partial model in which principle creative leadership had moderate direct and indirect effects on teacher professional learning via school culture.

What is noteworthy about the contribution of this study is that the author found the relationship between principal creative leadership, developed after transformational leadership [41], and power distance orientation to be a positive correlation. The creative leadership of principal may fit into the constructing of PLCs to benefit teacher learning. Finally, the empirical research indicates that young generation of teachers in China perceive

lower power distance in teacher-principal relationship. The result from this study is in conformity to that in the research about relationship among principal instructional leadership, teacher professional learning and power distance orientation as moderator.

Moreover, this study has three implications of the findings. Firstly, the study emphasizes that principals as leaders rather than managers may focus on influencing teachers rather than directly intervening in their teaching and professional learning. Secondly, China is a large country with diversified culture and differentiated levels of economic development in different cities and during a period of change. Future studies may not conclude similar results even using the exact measurement and data analysis techniques. Therefore, the author suggests that scholars keep a dynamic perspective on the relationship between culture and school leadership. In general, the findings of this research indicate that creative principals are the leaders who encourage teachers to involve in school decision-making, creatively solve problems, and work collaboratively with their colleagues.

ACKNOWLEDGMENT

This Research was funded by Xichuanhui Jindu K12 Private School and supported by Xichuanhui Education Group. I would like to thank Mrs. Yu Xian to give mentally support and professional suggestion in this work. I would like to appreciate all 637 teachers who participated in the questionnaire.

REFERENCES

- [1] Shengnan, L., & Hallinger, P. (2020). Unpacking the effects of culture on school leadership and teacher learning in China. *Educational Management Administration & Leadership*, 49(2), 214-233. <https://doi.org/10.1177/1741143219896042>
- [2] Wang, X. (2015). A review of the research on principal leadership type in western teacher professional learning community. *Modern Educational Management*, 10, 103-108. <https://doi.org/DOI:10.16697/j.cnki.xdjygl.2015.10.018>[In Chinese]
- [3] Hairon, S., Dimmock, C.. Singapore schools and professional learning communities: Teacher professional development and school leadership in an Asian hierarchical system[J]. *Educational Review*, 2012, 64(4): 405-424.
- [4] Qian H, Walker A and Li X (2017) The west wind vs the east wind: Instructional leadership model in China. *Journal of Educational Administration* 55(2): 186–206.
- [5] Zheng, X., & Zhang, J. (2015). Differences in Professional learning Communities between Chinese and Western teachers: a Cross-cultural perspective. *Studies In Foreign Education*, 42, 83-95. Retrieved 2 May 2022, from. [In Chinese]
- [6] Schechter, C. The Professional Learning Community as Perceived by Israeli School Superintendents, Principals and Teachers [J]. *International Review Education*, 2012, (58): 717-734. E. H. Miller, "A note on reflector arrays (Periodical style—Accepted for publication)," *WASET Trans. Antennas Propagat.*, to be published.
- [7] Kedzior, M., & Fifield, S. (2004). Teacher professional development.
- [8] Huffman, J., & Jacobson, A. (2003). Perceptions of professional learning communities. *International Journal Of Leadership In Education*, 6(3), 239-250. <https://doi.org/10.1080/1360312022000017480>
- [9] Yammarino, F., Spangler, W., & Bass, B. (1993). Transformational leadership and performance: A longitudinal investigation. *The Leadership Quarterly*, 4(1), 81-102. [https://doi.org/10.1016/1048-9843\(93\)90005-e](https://doi.org/10.1016/1048-9843(93)90005-e)
- [10] Puccio, G., Mance, M., & Murdock, M. (2012). *Creative Leadership: Skills That Drive Change* (2nd ed., pp. 12). Szechwan People Publishing House. [In Chinese]

- [11] Leithwood, K., Jantzi, D. A Review of Transformational School Leadership Research 1996-2005 [J]. *Leadership and Policy in Schools*, 2005, 4 (3): 177-199.
- [12] Branch, J. (2015). *Creative Educational Leadership: A Practical Guide to Leadership as Creativity*. Jacquie Turnbull. New York, NY: Bloomsbury. 2012. 192 pp. \$39.95 (paperback). *Educational Studies*, 51(3), 263-266. <https://doi.org/10.1080/00131946.2015.1033520>
- [13] Martin, A. (2007) *What's next: the 2007 changing nature of leadership survey*. Center for Creative Leadership.
- [14] Pinquart, M. (2021). School Environment. *Encyclopedia Of Evolutionary Psychological Science*, 6858-6868. https://doi.org/10.1007/978-3-319-19650-3_2434
- [15] Kalkan, Ü., Altınay Aksal, F., Altınay Gazi, Z., Atasoy, R., & Dağlı, G. (2020). The Relationship Between School Administrators' Leadership Styles, School Culture, and Organizational Image. *SAGE Open*, 10(1), 215824402090208. <https://doi.org/10.1177/2158244020902081>
- [16] Engels, N., Hotton, G., Devos, G., Bouckennooghe, D., & Aelterman, A. (2008). Principals in schools with a positive school culture. *Educational Studies*, 34(3), 159-174. <https://doi.org/10.1080/03055690701811263>
- [17] Zhang, Q., Siribanpitak, P., & Charoenkul, N. (2022). Creative leadership strategies for primary school principals to promote teachers' creativity in Guangxi, China. *Kakseatsseatsat Rjtoujourunranlaolfosfoscoiacliaslesiecniecnecse*, 41, 275-281. Retrieved 4 May 2022, from.
- [18] Harris, A. (2009). Creative leadership. *Management In Education*, 23(1), 9-11. <https://doi.org/10.1177/0892020608099076>
- [19] Klar, H. W. Fostering Department Chair Instructional Leadership Capacity: Laying the Groundwork for Distributed Instructional Leadership [J]. *International Journal of Leadership in Education*, 2012, 15 (2): 175-197.
- [20] Brislin RW (1980) Translation and content analysis of oral and written materials. *Methodology* 389-444.
- [21] Gruenert, S. W. (1998). *Development of a school culture survey*. University of Missouri-Columbia.
- [22] Luo Ruguo, & Wang Shanshan. (2011). A survey on school culture satisfaction of primary and secondary school teachers. *Shanghai Education research institute* (7), 4. [Chinese version]
- [23] Hofstede, G., Hofstede, G. J., Minkov, M. Cultures and organizations: Software of the mind[M]. New York: McGraw-Hill, 2010:6.
- [24] Earley P. C., Erez M. (1997). *The transplanted executive: Why you need to understand how workers in other countries see the world differently*. New York: Oxford University Press.
- [25] Kirkman, B., Chen, G., Farh, J., Chen, Z., & Lowe, K. (2009). Individual Power Distance Orientation and Follower Reactions to Transformational Leaders: A Cross-Level, Cross-Cultural Examination. *Academy Of Management Journal*, 52(4), 744-764. <https://doi.org/10.5465/amj.2009.43669971>
- [26] Stoll, L., & Temperley, J. (2009). Creative leadership: A challenge of our times. *School Leadership & Management*, 29(1), 65-78.
- [27] Parker, J. (1978). The Process: Cps in 5 Steps. *G/C/T*, 1(2), 60-63. <https://doi.org/10.1177/107621757800100219>
- [28] Liu, S., Hallinger, P., & Feng, D. (2016). Supporting the professional learning of teachers in China: Does principal leadership make a difference?. *Teaching And Teacher Education*, 59, 79-91. <https://doi.org/10.1016/j.tate.2016.05.023>
- [29] Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin*, 103(3), 411.
- [30] Taylor, R. (1990). Interpretation of the Correlation Coefficient: A Basic Review. *Journal of Diagnostic Medical Sonography*, 6(1), 35-39. <https://doi.org/10.1177/875647939000600106>
- [31] Schober, P., Boer, C., & Schwarte, L. (2018). Correlation Coefficients. *Anesthesia & Analgesia*, 126(5), 1763-1768. <https://doi.org/10.1213/ane.0000000000002864>
- [32] Zhang, Q., Siribanpitak, P., & Charoenkul, N. (2022). Creative leadership strategies for primary school principals to promote teachers' creativity in Guangxi, China. *Kakseatsseatsat Rjtoujourunranlaolfosfoscoiacliaslesiecniecnecse*, 41, 275-281. Retrieved 4 May 2022, from.
- [33] Tian, y. (2007). A study on the classroom environment and the cultivation of students' creativity in primary and secondary schools. *Tianjin Normal University: Basic education*, 8(2), 5e10. [in Chinese]
- [34] He J and Zhang Y (2017) The “trickle-down” model of ethical leadership on employees' CWB in China: Power distance orientation as moderator. *Management Review* 29(8): 179-186.
- [35] Dimmock C and Walker A (2005) *Educational Leadership: Culture and Diversity*. London: SAGE.
- [36] Stoll, L., & Temperley, J. (2009). Creative leadership: a challenge of our times. *School Leadership & Management*, 29(1), 65-78. <https://doi.org/10.1080/13632430802646404>
- [37] Hofstede, G. (2009). Geert Hofstede cultural dimensions.
- [38] Karacabey, M. F., Bellibaş, M. Ş., & Adams, D. (2022). Principal leadership and teacher professional learning in Turkish schools: Examining the mediating effects of collective teacher efficacy and teacher trust. *Educational studies*, 48(2), 253-272.
- [39] Berkovich, I., & Bogler, R. (2021). Conceptualising the mediating paths linking effective school leadership to teachers' organisational commitment. *Educational Management Administration & Leadership*, 49(3), 410-429.
- [40] Meyer, A., Richter, D., & Hartung-Beck, V. (2020). The relationship between principal leadership and teacher collaboration: Investigating the mediating effect of teachers' collective efficacy. *Educational management administration & leadership*, 1741143220945698.
- [41] Sternberg, R.J., Kaufman, J. C., & Pretz, J.E. (2004). A propulsion model of creative leadership. *Creativity and Innovation Management*, 13, 145-153.

Biographies



Yashi Ye (M'22)

Yashi was born in October 1992 in Chengdu, China. From 2013 to 2016, the author received her BACHELOR's degree in finance from the University of North Carolina at Greensboro. From 2020 to 2022, the author received her MASTER'S degree in educational innovation and global leadership at Johns Hopkins University in Baltimore, USA.

Since 2017, she has been working for Sichuan Xichuan Hui Education Group and Sichuan Xichuanhui Jindu K12 Private School in Chengdu, China. From a teacher to a school administrator, she has been supported by her employer during her studies and research at the Johns Hopkins School of Education. Since the beginning of her master's program, her academic research interests have focused on principal leadership, teacher professional learning and school management.

Ms. Ye are also a member of American Educational Research Association (AERA) and International Academic Forum (IAFOR). Accepted abstract title is Principal Creative Leadership for School Culture and Teacher Professional Learning.

Using Game-Based Learning Platforms in the Student Experience: Inclusion for All

Ellen Spender

Abstract— Technology has meant that students are no longer mainly reliant on textbooks and passive attendance at lectures to gain knowledge. This session explores the use of Kahoot! in curriculum design to enhance quality, diversity and inclusion for all. It is a real-life study aimed to individualise students to help break down barriers students face today. The focus will be on the role Kahoot! played before, and during, the move to remote learning as a result of the pandemic.

Keywords— attainment, engagement, inclusion, kahoot.

Cultural and Social Conflicts in Former British Settler Colonies (the US, Canada, Australia and New Zealand): The Fate of Indigenous Communities and Rights

Sophie Croisy

Abstract— The web of globalization has been spinning out for ages, with an acceleration of the process during the age of adventurers and discoverers who relied on key technological developments to open commercial routes and conquer resources and people for the benefit of the great kingdoms and queendoms of the world. These adventurers were followed by settlers looking for promised lands of new economic and cultural opportunities. And during this era of colonization, the relationship between these global settlers and the indigenous populations inhabiting these promised lands was characterized by complex patterns of conflict with their geopolitical specificities. For the purpose of this presentation, I have chosen to ground my analysis of the stakes of cultural diversity in the historical and contemporary relationship between the former settler colonies of the British Empire (the US, Canada, Australia and New Zealand) and the indigenous communities of these lands. As I rely on this particular context and as I defend the idea that preserving cultural diversity is essential, my goal is to really insist on the defense of cultural diversity as a human right that has been historically flouted by settler colonies and then settler nations, and as a condition of, not an impediment to, democratic progress in settler nations which need to confront their foundational narratives to alleviate the nation's intercultural tensions. My point is not to deny the importance of indigenous group rights which are central to their cultural survival, rights acquired by treaty or other legal texts, but to encompass the preservation of these group rights in a broader discussion on the preservation of cultural diversity as a human right and as a condition of democracy (power to all people).

Keywords— human rights, indigenous rights, cultural conflicts, democracy.

Sheriffs and Localized Racial Injustice in the American Jim Crow South, 1880-1965

Megan Qingxi Li

Abstract— Sheriffs have had an outsized role in the history of law enforcement in America, something that often goes unnoticed by the wider public and scholars. As elected officials, and the only law enforcement in rural areas, they control full counties with little to no accountability. This essay argues that although the Office of Sheriff has been one of the most overlooked and least understood positions in law enforcement, holders of the office dramatically shaped the experiences of localized racial injustice and exacerbated the failures of the legal system for historically marginalized populations, specifically African Americans in the American South, during the early 20th century. To prove this, three Supreme Court cases that demonstrate their abuse of power will be examined; *United States v. Shipp* in 1906, *Mack Claude Screws v. United States* in 1943, and *United States v. Cecil Price, et al* in 1964. A half-century after the end of the period examined in this paper, problems with sheriffs have not disappeared but instead have taken on additional forms, such as mass incarceration in jails, especially of Hispanic and Black people.

Keywords— sheriffs, racial injustice, Jim Crow South, Supreme Court.

Diversity and Intersectionality Through a Disability Lens

Leigha Shoup

Abstract— Presenters will pull back the curtain on diversity and intersectionality through the disability lens. This is a new perspective that is rarely trained on. It is crucial for service providers, educators, medical and mental health professionals, and others to understand intersectionality and how it impacts the disability community. This presentation will facilitate a more in-depth look into this topic by building a foundation of knowledge, reviewing the historical perspective, and discussing individuals' intersections from the disability community. Participants will receive the information in a myriad of forms, including lectures, facilitated group discussion, videos, and other activities.

Keywords— diversity, disability, mental health, intersectionality.

Arousal, Encoding, and Intrusive Memories

Hannah Gutmann, Rick Richardson, Richard Bryant

Abstract— Intrusive memories following a traumatic event are not uncommon. However, in some individuals, these memories become maladaptive and lead to prolonged stress reactions. A seminal model of PTSD explains that aberrant processing during trauma may lead to prolonged stress reactions and intrusive memories. This model explains that elevated arousal at the time of the trauma promotes data driven processing, leading to fragmented and intrusive memories. This study investigated the role of elevated arousal on the development of intrusive memories. We measured salivary markers of arousal and investigated what impact this had on data driven processing, memory fragmentation, and subsequently, the development of intrusive memories. We assessed 100 healthy participants to understand their processing style, arousal, and experience of intrusive memories. Participants were randomised to a control or experimental condition, the latter of which was designed to increase their arousal. Based on current theory, participants in the experimental condition were expected to engage in more data driven processing and experience more intrusive memories than participants in the control condition. This research aims to shed light on the mechanisms underlying the development of intrusive memories to illustrate ways in which therapeutic approaches for PTSD may be augmented for greater efficacy.

Keywords— stress, cortisol, SAA, PTSD, intrusive memories.

Indigenous Healers and Indigenous Trauma: Healing at the Intersections of Colonial, Intergenerational, and Individual Trauma for Indigenous Peoples in Canada

Suzanne L. Stewart, Mikaela D. Gabriel

Abstract— Background: Indigenous People face multiple barriers to successful life transitions, including housing, employment, education, and health. Current statistical trends paint devastating life transitions for Indigenous Peoples, but colonization and its intergenerational impacts are typically lacking as the crucial context in which these trends occur. This presentation will illustrate the massive impact of colonization on Indigenous Peoples; its intergenerational transmission, and how it impacts Indigenous clients seeking mental health treatment today. Methods: A qualitative, narrative inquiry methodology was used to honour Indigenous storytelling and knowledge transmission. Indigenous Elders, outreach workers, and homeless clients were interviewed and narratively analyzed for in-depth trends and themes. Impact: This research provides a wealth of in-depth information as to the life transition needs of Indigenous clients, identify the systemic impacts of colonization to the health and wellbeing of Indigenous People, and strategies for mental health treatment.

Keywords— indigenous trauma, indigenous peoples of canada, intergenerational trauma, colonial trauma and treatment.

Preventative Programs for At-Risk Families of Child Maltreatment: Using Home Visiting and Intergenerational Relationships

K. Gordon

Abstract— One in three children in the United States is a victim of a maltreatment investigation, and about one in nine children has a substantiated investigation. Home visiting is one of several preventative strategies rooted in an early childhood approach that fosters maternal, infant, and early childhood health, protection, and growth. In the United States, 88% of states report administering home visiting programs or state-designed models. The purpose of this study was to conduct a systematic review on home visiting programs in the United States focused on the prevention of child abuse and neglect. This systematic review included 17 articles which found that most of the studies reported optimistic results. Common across studies was program content related to (1) typical child development, (2) parenting education, and (3) child physical health. Although several factors common to home visiting and parenting interventions have been identified, no research has examined the common components of manualized home visiting programs to prevent child maltreatment. Child maltreatment can be addressed with home visiting programs with evidence-based components and cultural adaptations that increase prevention by assisting families in tackling the risk factors they face. An innovative approach to child maltreatment prevention is bringing together at-risk families with the aging community. This innovative approach was prompted due to existing home visitation programs only focusing on improving skillsets and providing temporary relationships. This innovative approach can provide the opportunity for families to build a relationship with an aging individual who can share their wisdom, skills, compassion, love, and guidance, to support families in their well-being and decrease child maltreatment occurrence. Families would be identified if they experience any of the risk factors including parental substance abuse, parental mental illness, domestic violence, and poverty. Families would also be identified as at risk if they lack supportive relationships such as grandparents or relatives. Families would be referred by local agencies such as medical clinics, hospitals, schools, etc. that have interactions with families regularly. The aging community would be recruited at local housing communities and community centers. An aging individual would be identified by the elderly community when there is a need or interest in a relationship by or for the individual. Cultural considerations would be made when assessing for compatibility between the families and aging individuals. The pilot program will consist of a small group of participants to allow manageable results to evaluate the efficacy of the program. The pilot will include pre-and post-surveys to evaluate the impact of the program. From the results, data would be created to determine the efficacy as well as the sufficiency of the details of the pilot. The pilot would also be evaluated on whether families were referred to Child Protective Services during the pilot as it relates to the goal of decreasing child maltreatment. The ideal findings will display a decrease in child maltreatment and an increase in family well-being for participants.

Keywords— abuse, child maltreatment, home visiting, neglect, preventative

K. G. Author is with California Baptist University, Riverside, CA 92504 USA (phone: 951-542-3860; e-mail: kgordon@calbaptist.edu).

Sceletium Tortuosum: A Review On Its Phytochemistry, Pharmacokinetics, Biological and Clinical Activities

TL Olatunji^a, F Siebert^a, AE Adetunji^b BH Harvey^{c,d,e}, J Gericke^{c,e}, JH Hamman^c, F Van der Kooy^{c*}

^aUnit for Environmental Sciences and Management, North-West University, Potchefstroom, South Africa.

^bSchool of Life Sciences, University of KwaZulu-Natal, Durban 4001, South Africa.

^cCentre of Excellence for Pharmaceutical Sciences, North-West University, Private Bag X6001, Potchefstroom, 2520, South Africa.

^dSAMRC Unit on Risk and Resilience in Mental Disorders, Department of Psychiatry and Mental Health, University of Cape Town, South Africa.

^eDepartment of Pharmacology, School of Pharmacy, North-West University, Potchefstroom Campus, South Africa 2520

*Corresponding author.

Tel: +27 (0)18 299 2236

E-mail address: frank.vanderkooy@nwu.ac.za (F. van der Kooy)

Abstract

Ethnopharmacological relevance: *Sceletium tortuosum* (L.) N.E.Br., the most sought after and widely researched species in the genus *Sceletium* is a succulent forb endemic to South Africa. Traditionally, this medicinal plant is mainly masticated or smoked and used for the relief of toothache, abdominal pain, and as a mood-elevator, analgesic, hypnotic, anxiolytic, thirst and hunger suppressant, and for its intoxicating/euphoric effects. *Sceletium tortuosum* is currently of widespread scientific interest due to its clinical potential in treating anxiety and depression, relieving stress in healthy individuals, and enhancing cognitive functions. These pharmacological actions are attributed to its phytochemical constituents referred to as mesembrine-type alkaloids.

Aim of the review: The aim of this review was to comprehensively summarize and critically evaluate recent research advances on the phytochemistry, pharmacokinetics, biological and clinical activities of the medicinal plant *S. tortuosum*. Additionally, current ongoing research and future perspectives are also discussed.

Methods: All relevant scientific articles, books, MSc and Ph.D. dissertations on botany, behavioral pharmacology, traditional uses, and phytochemistry of *S. tortuosum* were retrieved from different databases (including Science Direct, PubMed, Google Scholar, Scopus and Web of Science). For pharmacokinetics and pharmacological effects of *S. tortuosum*, the focus fell on relevant publications published between 2009 and 2021.

Results: Twenty-five alkaloids belonging to four structural classes viz: mesembrine, Sceletium A4, joubertiamine, and tortuosamine, have been identified from *S. tortuosum*, of which the mesembrine class is predominant. The crude extracts and commercially available standardized extracts of *S. tortuosum* have displayed a wide spectrum of biological activities (e.g. antimalarial, anti-oxidant, immunomodulatory, anti-HIV, neuroprotection, enhancement of cognitive function) in *in vitro* or *in vivo* studies. This plant has not yet been studied in a clinical population, but has potential for enhancing cognitive function, and managing anxiety and depression.

Conclusion: As an important South African medicinal plant, *S. tortuosum* has garnered many research advances on its phytochemistry and biological activities over the last decade. These scientific studies have shown that *S. tortuosum* has various bioactivities. The findings have further established the link between the phytochemistry and pharmacological application, and support the traditional use of *S. tortuosum* in the indigenous medicine of South Africa.

Keywords: Aizoaceae, Mesembrine, Serotonin, *Sceletium tortuosum*, Zembrin®, psychoactive, antidepressant.

Micromechanical Compatibility Between Cells and Scaffold Mediates the Efficacy of Regenerative Medicine

Li Yang¹, Yang Song^{1,2}, Martin Y.M. Chiang²,

1. School of Bioengineering, Chongqing University, Chongqing 400044, China

2. Material Measurement Laboratory, National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899, USA

Abstract

Objective: To experimentally substantiate the micromechanical compatibility between cell and scaffold, in the regenerative medicine approach for restoring bone volume, is essential for phenotypic transitions

Methods: Through nanotechnology and electrospinning process, nanofibrous scaffolds were fabricated to host dental follicle stem cells (DFSCs). Blends (50:50) of polycaprolactone (PCL) and silk fibroin (SF), mixed with various content of cellulose nanocrystals (CNC, up to 5% in weight), were electrospun to prepare nanofibrous scaffolds with heterogenous microstructure in terms of fiber size. Colloidal probe atomic force microscopy (AFM) and conventional uniaxial tensile tests measured the scaffold stiffness at the micro- and macro-scale, respectively. The cell elastic modulus and cell-scaffold adhesive interaction (i.e., a chemical function) were examined through single-cell force spectroscopy using AFM. The quantitative reverse transcription-polymerase chain reaction (qRT-PCR) was used to determine if the mechanotransduction signal (i.e., Yap1, Wwr2, Rac1, MAPK8, Ptk2 and Wnt5a) is upregulated by the scaffold stiffness at the micro-scale (cellular scale).

Results: The presence of CNC produces fibrous scaffolds with bimodal distribution of fiber diameter. This structural heterogeneity, which is CNC-composition dependent, remarkably modulates the mechanical functionality of scaffolds at microscale and macroscale simultaneously, but not the chemical functionality (i.e., only a single material property is varied). In *in vitro* tests, the osteogenic differentiation and gene expression associated with mechano-sensitive cell markers correlates to the degree of micromechanical compatibility between DFSCs and the scaffold.

Conclusion: Cells require compliant scaffolds to encourage energetically favorable interactions for mechanotransduction, which are converted into changes in cellular biochemistry to direct the phenotypic evolution. The micromechanical compatibility is indeed important to the efficacy of regenerative medicine.

Keywords: Micromechanical compatibility; Osteogenesis; Cellulose nanocrystal; Nanofibrous scaffold; Dental follicle stem cells

Acknowledgement: This work was supported, in part, by Innovation and Attracting Talents Program for College and University (“111” Project, B06023) and National Natural Science Foundation of China (11532004).

Design of Experiment for Optimizing Immunoassay Microarray Printing

Alex J. Summers, Jasmine P. Devadhasan, Douglas Montgomery, Brittany Fischer, Jian Gu and Frederic Zenhausern

Abstract—Immunoassays have been utilized for several applications, including detection of pathogens. Our laboratory is in development of a tier 1 biothreat panel utilizing Vertical Flow Assay (VFA) technology for simultaneous detection of pathogens and toxins. One method of manufacturing VFA membranes is with non-contact piezoelectric dispensing which provides advantages, such as low-volume and rapid dispensing without compromising structural integrity of antibody or substrate. Challenges of this process include premature discontinuation of dispensing and misaligned spotting. Preliminary data revealed the Yp 11C7 mAb (11C7) reagent to exhibit a large angle of failure during printing which may have contributed to variable printing outputs. A Design of Experiment (DOE) was executed using this reagent to investigate the effects of hydrostatic pressure and reagent concentration on microarray printing outputs.

A Nano-plotter 2.1 (GeSIM, Germany) was used for printing antibody reagents onto nitrocellulose membrane sheets in a clean room environment. A spotting plan was executed using Spot-Front-End software to dispense volumes of 11C7 reagent (20-50 droplets; 1.5-5 mg/mL) in a 6-test spot array at 50 target membrane locations. Hydrostatic pressure was controlled by raising the Pressure Compensation Vessel (PCV) above or lowering it below our current working level. It was hypothesized that raising or lowering the PCV 6 inches would be sufficient to cause either liquid accumulation at the tip or discontinuous droplet formation. After aspirating 11C7 reagent, we tested this hypothesis under stroboscope. 75% of the effective raised PCV height and of our hypothesized lowered PCV height were used. Humidity (55%) was maintained using an Airwin BO-CT1 humidifier. The number and quality of membranes was assessed after staining printed membranes with dye. The droplet angle of failure was recorded before and after printing to determine a “stroboscope score” for each run. The DOE set was analyzed using JMP software.

Hydrostatic pressure and reagent concentration had a significant effect on the number of membranes output. As hydrostatic pressure was increased by raising the PCV 3.75 inches, or decreased by lowering the PCV -4.5 inches, membrane output decreased. However, with the hydrostatic pressure closest to equilibrium, our current working level, membrane output reached the 50-membrane target. As the reagent concentration increased from 1.5 to 5 mg/mL, the membrane output also increased. Reagent concentration likely effected the number of membrane output due to the associated dispensing volume needed to saturate the membranes. However, only hydrostatic pressure had a significant effect on stroboscope score which could be due to discontinuation of dispensing and thus the stroboscope check could not find a droplet to record. Our JMP predictive model had a high degree of agreement with our observed results.

The JMP model predicted that dispensing the highest concentration of 11C7 at our current PCV working level would yield the highest number of quality membranes, which correlated with our results.

Acknowledgements: This work was supported by the Chemical Biological Technologies Directorate (Contract # HDTRA1-16-C-0026) and the Advanced Technology International (Contract # MCDC-18-04-09-002) from the Department of Defense Chemical and Biological Defense program through the Defense Threat Reduction Agency (DTRA).

Keywords—Immunoassay, Microarray, Design of Experiment, Piezoelectric Dispensing

Alex J. Summers is with the Center for Applied NanoBioscience and Medicine, College of Medicine, University of Arizona, Phoenix, AZ 85004 (e-mail: ajsummers@arizona.edu).

Jasmine P. Devadhasan is with the Center for Applied NanoBioscience and Medicine, College of Medicine, University of Arizona, Phoenix, AZ 85004 (e-mail: jpd136@arizona.edu).

Douglas Montgomery is with the School of Computing, Informatics and Decision Systems Engineering, Arizona State University, Tempe, AZ, USA. (e-mail: doug.montgomery@asu.edu).

Brittany Fischer is with the School of Computing, Informatics and Decision Systems Engineering, Arizona State University, Tempe, AZ, USA. (e-mail: bmf5007@gmail.com).

Jian Gu is with the Center for Applied NanoBioscience and Medicine and the Department of Basic Medical Sciences, College of Medicine, University of Arizona, Phoenix, AZ 85004 (e-mail: jgu10@arizona.edu).

Frederic Zenhasuern is with the Center for Applied NanoBioscience and Medicine and the Department of Basic Medical Sciences, College of Medicine, University of Arizona, Phoenix, AZ 85004 (e-mail: fzenhaus@arizona.edu).

Application of Nanoparticles in Biomedical and MRI

Raziyeh Mohammadi

Abstract— At present, nanoparticles are used for various biomedical applications where they facilitate laboratory diagnostics and therapeutics. The performance of nanoparticles for biomedical applications is often assessed by their narrow size distribution, suitable magnetic saturation, and low toxicity effects. Superparamagnetic iron oxide nanoparticles have received great attention due to their applications as contrast agents for magnetic resonance imaging (MRI). (Processes in the tissue where the blood brain barrier is intact in this way shielded from the contact to this conventional contrast agent and will only reveal changes in the tissue if it involves an alteration in the vasculature. This technique is very useful for detecting tumors and can even be used for detecting metabolic functional alterations in the brain, such as epileptic activity. SPIONs have found application in Magnetic Resonance Imaging (MRI) and magnetic hyperthermia. Unlike bulk iron, SPIONs do not have remnant magnetization in the absence of the external magnetic field; therefore, a precise remote control over their action is possible.

Keywords— nanoparticles, MRI, biomedical, iron oxide, spions.

The Effect of Pregabalin on the Prevention of Succinylcholine-Induced Fasciculation and Myalgia

Shahryar Sane, Hedyeh Rezaei, Shima Khanahmadi

Abstract—

Purpose: This study evaluates the effect of pregabalin on fasciculation and myalgia after using succinylcholine.

Design: This randomized double-blind prospective study was conducted among 100 patients aged 20 to 60 years old.

Methods: Pregabalin (300 mg) and placebo (in capsule form) were placed in similar containers. The results were analyzed by SPSS 23 software, and statistical analysis consisted of χ^2 test and t test, and a P value less than .05 was considered significant.

Findings: The mean pain score in the group receiving pregabalin was lower than the placebo group. According to the χ^2 test, there was a significant difference between the two groups in the frequency of fasciculation ($P = .003$). Mean fasciculation severity in the pregabalin group was lower than placebo group. According to t test, there was a significant difference in the mean fasciculation severity between the two groups ($P = .002$).

Conclusions: This study showed that 300 mg of pregabalin was effective in reducing postoperative fasciculation and myalgia in patients treated with succinylcholine.

Keywords— fasciculation, myalgia, pregabalin, succinylcholine.

Evaluating the Effectiveness of Combined Psychiatric and Psychotherapeutic Care versus Psychotherapy Alone in the Treatment of Depression and Anxiety in Cancer Patients

Nathen A. Spitz, Dennis Martin Kivlighan III, Arwa Aburizik

Abstract— Background and Purpose: Presently, there is a paucity of naturalistic studies that directly compare the effectiveness of psychotherapy versus concurrent psychotherapy and psychiatric care for the treatment of depression and anxiety in cancer patients. Informed by previous clinical trials examining the efficacy of concurrent approaches, this study sought to test the hypothesis that a combined approach would result in the greatest reduction of depression and anxiety symptoms. Methods: Data for this study consisted of 433 adult cancer patients, with 252 receiving only psychotherapy and 181 receiving concurrent psychotherapy and psychiatric care at the University of Iowa Hospitals and Clinics. Longitudinal PHQ9 and GAD7 data were analyzed between both groups using latent growth curve analyses. Results: After controlling for treatment length and provider effects, results indicated that concurrent care was more effective than psychotherapy alone for depressive symptoms ($\gamma_{12} = -0.12$, $p = .037$). Specifically, the simple slope for concurrent care was -0.25 ($p = .022$), and the simple slope for psychotherapy alone was -0.13 ($p = .006$), suggesting that patients receiving concurrent care experienced a greater reduction in depressive symptoms compared to patients receiving psychotherapy alone. In contrast, there were no significant differences between psychotherapy alone and concurrent psychotherapy and psychiatric care in the reduction of anxious symptoms. Conclusions: Overall, as both psychotherapy and psychiatric care may address unique aspects of mental health conditions, in addition to potentially providing synergetic support to each other, a combinatorial approach to mental healthcare for cancer patients may improve outcomes.

Keywords— psychiatry, psychology, psycho-oncology, combined care, psychotherapy, behavioral psychology.

Nathen A. Spitz, Dennis Martin Kivlighan III, Arwa Aburizik are with the Department of Internal Medicine - Hematology, Oncology and Blood and Marrow Transplantation, University of Iowa Hospitals and Clinics, Iowa City, Iowa (e-mail: spitz@uiowa.edu).

Bias Prevention in Automated Diagnosis of Melanoma: Augmentation of a Convolutional Neural Network Classifier

Kemka Ihemelandu, Chukwuemeka Ihemelandu

Abstract— Melanoma remains a public health crisis, with incidence rates increasing rapidly in the past decades. Improving diagnostic accuracy to decrease misdiagnosis using Artificial intelligence (AI) continues to be documented. Unfortunately, unintended racially biased outcomes, a product of lack of diversity in the dataset used, with a noted class imbalance favoring lighter vs. darker skin tone, have increasingly been recognized as a problem. Resulting in noted limitations of the accuracy of the Convolutional neural network (CNN) models. CNN models are prone to biased output due to biases in the dataset used to train them. Our aim in this study was the optimization of convolutional neural network algorithms to mitigate bias in the automated diagnosis of melanoma. We hypothesized that our proposed training algorithms based on a data augmentation method to optimize the diagnostic accuracy of a CNN classifier by generating new training samples from the original ones will reduce bias in the automated diagnosis of melanoma. We applied geometric transformation, including; rotations, translations, scale change, flipping, and shearing. Resulting in a CNN model that provided a modified input data making for a model that could learn subtle racial features. Optimal selection of the momentum and batch hyperparameter increased our model accuracy. We show that our augmented model reduces bias while maintaining accuracy in the automated diagnosis of melanoma.

Keywords— bias, augmentation, melanoma, convolutional neural network.

Childhood Epilepsy: Solutions and Study Methods That Have Been Attracting Researchers

Jhonas Geraldo Peixoto Flauzino, Pedro Pompeo Boechat Araujo, Alexia Allis Rocha Lima, Giovanna Biângulo Lacerda Chaves, Victor Ryan Ferrão Chaves

Abstract— As it is commonly presented in childhood, epilepsy has clinical relevance and worrying morbidity (CARVALHO et al., 2021), considering that the problem involves organic, psychological, social and educational factors (MAIA et al., 2004). Therefore, epilepsy has many nuances that need to be considered so that the control of the disease is effective and efficient, understanding that the holistic view of the infant is essential for therapeutic success, which justifies the interest in the development of the work presented.

Keywords— childhood, epilepsy, methods, researchers.

Expert Opinions about Barriers to Physical Activity among Ghanaian Adults with Type 2 Diabetes Mellitus: A Qualitative Descriptive Study

Mohammed Amin, Debra Kerr, Yacoba Atiase, Andrea Driscoll

Abstract—

Background: Physical activity (PA) is a major component of diabetes self-care management; although this is often stressed to patients, many adults with Type 2 Diabetes Mellitus (T2DM) lead sedentary lifestyles, and barriers exist for uptake of PA.

Aim: To explore opinions of healthcare professionals about barriers to PA for adults with T2DM in Ghana.

Methods: This qualitative descriptive study included 13 healthcare professionals (3 Physiotherapists, 3 exercise physiologists, 3 nurses, and 4 physicians) who provide care to individuals with T2DM in Ghana. Data was collected by semi-structured interviews. The social cognitive theory guided the design of the interview schedule. Data was analysed using thematic analysis.

Results: Four main themes were identified: 1) Individual-related factors, 2) Interpersonal factors, 3) Environment-related factors, and 4) Health system-related factors. Fear of injury, existing comorbidities, and lack of time make it difficult for people with T2DM to engage in PA. Lack of family support, fear of social ridicule, and cultural beliefs prevent uptake of PA. Poorly designed community spaces, including safe walkways and lack of exercise facilities, inhibit PA participation. Few physical therapists and physical therapy centres exist to support PA participation among people with T2DM. Some nurses and doctors lack adequate knowledge to deliver proper PA education to clients, thereby making clients lack the needed support.

Conclusion: Adults with T2DM may be restricted from partaking in PA arising from personal and external factors. This study has identified that barriers to exercise need consideration when designing PA programs that aim to improve health outcomes for people with T2DM in Ghana.

Keywords— type 2 diabetes, physical activity, exercise, ghana, qualitative, barriers.

A Comparison of qCON/qNOX to the Bispectral Index as Indices of Antinociception in Surgical Patients Undergoing General Anesthesia with Laryngeal Mask Airway

Roya Yumul, Ofelia Loani Elvir-Lazo, Sevan Komshian, Ruby Wang, Jun Tang

Abstract—

Background: An objective means for monitoring the antinociceptive effects of perioperative medications has long been desired as a way to provide anesthesiologists information regarding a patient's level of antinociception and preclude any untoward autonomic responses and reflexive muscular movements from painful stimuli intraoperatively. To this end, electroencephalogram (EEG) based tools including BIS and qCON were designed to provide information about the depth of sedation while qNOX was produced to inform on the degree of antinociception. The goal of this study was to compare the reliability of qCON/qNOX to BIS as specific indicators of response to nociceptive stimulation.

Methods: Sixty-two patients undergoing general anesthesia with LMA were included in this study. Institutional Review Board (IRB) approval was obtained, and informed consent was acquired prior to patient enrollment. Inclusion criteria included American Society of Anesthesiologists (ASA) class I-III, 18 to 80 years of age, and either gender. Exclusion criteria included the inability to consent. Withdrawal criteria included conversion to the endotracheal tube and EEG malfunction. BIS and qCON/qNOX electrodes were simultaneously placed on all patients prior to induction of anesthesia and were monitored throughout the case, along with other perioperative data, including patient response to noxious stimuli. All intraoperative decisions were made by the primary anesthesiologist without influence from qCON/qNOX. Student's t-distribution, prediction probability (PK), and ANOVA were used to statistically compare the relative ability to detect nociceptive stimuli for each index. Twenty patients were included for the preliminary analysis.

Results: A comparison of overall intraoperative BIS, qCON and qNOX indices demonstrated no significant difference between the three measures ($N=62$, $p>0.05$). Meanwhile, index values for qNOX (62 ± 18) were significantly higher than those for BIS (46 ± 14) and qCON (54 ± 19) immediately preceding patient responses to nociceptive stimulation in a preliminary analysis ($N=20$, * $p=0.0408$). Notably, certain hemodynamic measurements demonstrated a significant increase in response to painful stimuli (MAP increased from 74 ± 13 mm Hg at baseline to 84 ± 18 mm Hg during noxious stimuli [$p=0.032$] and HR from 76 ± 12 BPM at baseline to 80 ± 13 BPM during noxious stimuli [$p=0.078$] respectively).

Conclusion: In this observational study, BIS and qCON/qNOX provided comparable information on patients' level of sedation throughout the course of an anesthetic. Meanwhile, increases in qNOX values demonstrated a superior correlation to an imminent response to stimulation relative to all other indices.

Keywords— antinociception, BIS, general anesthesia, LMA, qCON/qNOX.

Roya Yumul, Ofelia Loani Elvir-Lazo, Sevan Komshian, Ruby Wang, Jun Tang are with the Department of Anesthesiology, Cedars-Sinai Medical Center, Los Angeles, CA, USA (e-mail: Roya.Yumul@cshs.org).

Roya Yumul is with the David Geffen School of Medicine-UCLA, Los Angeles, CA, USA.

The Impact of COVID-19 on Childhood Academic Functioning and Anxiety: A Literature Review

Lindsey Giunta

Abstract— This review examines the current literature regarding the impact of COVID-19 on academic functioning and anxiety in children and adolescents. The objective was to determine the ways in which the pandemic affected youth mental health and academics, in addition to the extent that these factors were transformed as a result of the worldwide state of affairs. Twenty papers were selected and reviewed, and data showed long term consequences in youth mental health resulting from the current pandemic. The COVID-19 pandemic and its associated lockdowns led to disrupted childhood education, and data showed that the growth of cognitive executive functions was impacted to varying degrees dependent upon geographic location. The literature recommends supplemental education on the national level, as well as mental health promotion within communities and schools.

Keywords— pandemic, children, adolescents, anxiety, academic functioning.

Preoperative Anxiety Evaluation: Comparing the Visual Facial Anxiety Scale/Yumul Faces Anxiety Scale, Numerical Verbal Rating Scale, Categorization Scale, and the State-Trait Anxiety Inventory

Roya Yumul, Chse, Ofelia Loani Elvir Lazo, David Chernobylsky, Omar Durra

Abstract—

Background: Preoperative anxiety has been shown to be caused by the fear associated with surgical and anesthetic complications. However, the current gold standard for assessing patient anxiety, the STAI, is problematic to use in the preoperative setting, given the duration and concentration required to complete the 40-item extensive questionnaire. Our primary aim in the study is to investigate the correlation of the Visual Facial Anxiety Scale (VFAS) and Numerical Verbal Rating Scale (NVRS) to State-Trait Anxiety Inventory (STAI) to determine the optimal anxiety scale to use in the perioperative setting.

Methods: A clinical study of patients undergoing various surgeries was conducted utilizing each of the preoperative anxiety scales. Inclusion criteria included patients undergoing elective surgeries, while exclusion criteria included patients with anesthesia contraindications, inability to comprehend instructions, impaired judgment, substance abuse history, and those pregnant or lactating. 293 patients were analyzed in terms of demographics, anxiety scale survey results, and anesthesia data via Spearman Coefficients, Chi-Squared Analysis, and Fischer's exact test utilized for comparative analysis.

Results: Statistical analysis showed that VFAS had a higher correlation to STAI than NVRS ($r_s=0.66$, $p<0.0001$ vs $r_s=0.64$, $p<0.0001$). The combined VFAS-Categorization Scores showed the highest correlation with the gold standard ($r_s=0.72$, $p<0.0001$). Subgroup analysis showed similar results. STAI evaluation time (247.7 ± 54.81 sec) far exceeds VFAS (7.29 ± 1.61 sec), NVRS (7.23 ± 1.60 sec), and Categorization scales (7.29 ± 1.99 sec). Patients preferred VFAS (54.4%), Categorization (11.6%), and NVRS (8.8%). Anesthesiologists preferred VFAS (63.9%), NVRS (22.1%) and Categorization Scales (14.0%). Of note, the top five causes of preoperative anxiety were determined to be waiting (56.5%), pain (42.5%), family concerns (40.5%), no information about surgery (40.1%), or anesthesia (31.6%).

Conclusions: Combined VFAS-Categorization Score (VCS) demonstrates the highest correlation to the gold standard, STAI. Both VFAS and Categorization tests also take significantly less time than STAI, which is critical in the preoperative setting. Among both patients and anesthesiologists, VFAS was the most preferred scale. This forms the basis of the Yumul FACES Anxiety Scale, designed for quick quantization and assessment in the preoperative setting while maintaining a high correlation to the golden standard. Additional studies using the formulated Yumul FACES Anxiety Scale are merited.

Keywords— preoperative anxiety, visual facial anxiety scale, numerical verbal anxiety scale, state-trait anxiety inventory.

Efficacy of Gamma Knife Radiosurgery in Refractory Obsessive-Compulsive Disorder: An Indian Experience

Milind Sankhe, Sanjeev Pattankar

Abstract— Objectives: To analyze our experience with Gamma Knife Radiosurgery (GKRS) in refractory Obsessive-Compulsive Disorder (OCD) and report the safety and efficacy/long-term outcome in patients using the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS). Methods: A retrospective review of patients receiving GKRS for refractory OCD between 2000 and 2020 was carried out. Case files of the eligible (N=9) patients were reviewed for clinical, radio-therapeutic, and outcome data. Additionally, patients were contacted telephonically to enquire about their experiences and to obtain retroactive consent for GKRS in June 2021. Results: The mean age at the time of GKRS and mean duration of OCD prior to GKRS were 30.1 ± 9.4 years and 10.2 ± 5.8 years, respectively. The mean baseline Y-BOCS score was 29.6 ± 4.7 . Our first patient received cingulotomy, while the rest underwent anterior capsulotomy. The median margin dose (50% isodose) was 70 Gy. 23.8 ± 7.7 was the mean Y-BOCS score at the last follow-up (median=30 months). 44.4% of patients showed full/partial response ($\geq 25\%$ reduction in Y-BOCS score) at the last follow-up. In anterior capsulotomy (8 patients), 50% of patients showed full/partial response, with extreme OCD patients faring poorer. A single case of cingulotomy resulted in no response. No adverse radiation effects were noted. 55.6% of patients gave retroactive consent telephonically. Conclusions: GKRS is a safe and effective non-invasive treatment modality for refractory OCD. The ventral anterior capsule is the preferred target. Maximum radiation doses of 120 to 160 Gy are well tolerated. Extremely severe OCD cases fared poorer. Proper awareness about the availability and efficacy of GKRS in refractory OCD is required in India.

Keywords— refractory obsessive-compulsive disorder, gamma knife radiosurgery, efficacy, Indian experience.

Molecular Profiles of Microbial Etiologic Agents Forming Biofilm in Urinary Tract Infections of Pregnant Women by RTPCR Assay

B. Nageshwar Rao

Abstract— Urinary tract infection (UTI) represents the most commonly acquired bacterial infection worldwide, with substantial morbidity, mortality, and economic burden. The objective of the study is to characterize the microbial profiles of uropathogenic in the obstetric population by RTPCR. Study design: An observational cross-sectional study was performed at a single tertiary health care hospital among 50 pregnant women with UTIs, including asymptomatic and symptomatic patients attending the outpatient department and inpatient department of Obstetrics and Gynaecology. Methods: Serotyping and genes detection of various uropathogens were studied using RTPCR. Pulse field gel electrophoresis methods were used to determine the various genetic profiles. Results: The present study shows that CsgD protein, involved in biofilm formation in Escherichia coli, VIM1, IMP1 genes for Klebsiella were identified by using the RTPCR method. Our results showed that the prevalence of VIM1 and IMP1 genes and CsgD protein in E.coli showed a significant relationship between strong biofilm formation, and this may be due to the prevalence of specific genes. Finally, the genetic identification of RTPCR results for both bacteria was correlated with each other and concluded that the above uropathogens were common isolates in producing Biofilm in the pregnant woman suffering from urinary tract infection in our hospital observational study.

Keywords— biofilms, Klebsiella, E.coli, urinary tract infection.

Data Augmentation for Early-Stage Lung Nodules Using Deep Image Prior and Pix2pix

Qasim Munye, Juned Islam, Haseeb Qureshi, Syed Jung

Abstract— Lung nodules are commonly identified in a computed tomography (CT) scan by experienced radiologists at a relatively late stage. Early diagnosis can greatly increase survival. We propose using a pix2pix conditional generative adversarial network to generate realistic images simulating early-stage lung nodule growth. We have applied deep image prior to 2341 slices from 895 computed tomography (CT) scans from the LIDC dataset to generate pseudo-healthy medical images. From these images, 819 were chosen to train a pix2pix network. We observed that for most of the images the pix2pix network was able to generate images where the nodule increased in size and intensity across epochs. To evaluate the images, 400 generated images were chosen at random and shown to a medical student beside its corresponding original image. Of these 400 generated images, 384 were defined as satisfactory - meaning it resembled a nodule and was visually similar to the corresponding image. We believe that this generated dataset could be used as training data for neural networks to detect lung nodules at an early stage or to improve the accuracy of such networks. This is particularly significant as datasets containing the growth of early-stage nodules are scarce. This project shows that the combination of deep image prior and generative models could potentially open the door to creating larger datasets than currently possible and has the potential to increase the accuracy of medical classification tasks.

Keywords— medical technology, artificial intelligence, radiology, lung cancer.

Ultrasound-Guided Intraarticular Electrical Stimulation, a New Intervention for Improving Physical Activity and Quality of Life in Patients with Severe Knee Osteoarthritis: Cases Report

Azar Behnamfar^{1*}

¹ Doctor of physical therapy student, Arizona school of health science, A.T. Still University, Arizona, USA.

Corresponding Author: Azar Behnamfar E-mail: azarbehnamfar@gmail.com.

ARTICLE INFO	ABSTRACT
Received: Accepted: Volume: Issue: KEYWORDS knee osteoarthritis, knee pain, physical therapy, rehabilitation, physical activity, COVID-19, noninvasive treatment, osteoarthritis	Knee osteoarthritis is one of the diseases that cause degeneration of articular cartilage, subchondral bone, and chronic joint pain and stiffness. While there is no effective treatment for inhibiting the progression of its advancement, there are various interventions for managing symptoms of this condition in its early stages and a few treatments for advanced stages. In patients with severe knee osteoarthritis, the most recommended treatment is surgery. However, this procedure is contraindicated for many patients because of their comorbidities. In this study, a new physical intervention for managing the symptoms of advanced knee osteoarthritis in patients whose surgery is not indicated had been investigated and its results had been reported. These data demonstrate that this method seems to be effective in pain and stiffness reduction and improves physical activity and quality of life. More studies is needed to insure the result of this study in the future.

1. Introduction

Osteoarthritis (OA) is the first cause of disability and is the number two reason that causes a reduction in physical activity (Badley, 1995). The most affected joint by osteoarthritis is the knee joint (Thomas et al, 2003) which can lead to gait impairments and disability (Davis et al, 1991), musculoskeletal complications like losing strength (Liikavainio et al, 2008), limitation in range of motion and physical activity, and increased risk and fear of falling (Cumming et al, 2000). In the end stages of knee OA, patients experience a great deal of pain during physical activity (Atamaz et al, 2006) and most of them respond poorly to recommended noninvasive treatments such as manual therapy techniques (Pollard et al, 2008), non-steroidal anti-inflammatory drugs (NSAID) (Cooper et al, 2019), intraarticular injections (Ayhan et al, 2014), Transcutaneous Electrical Nerve Stimulation (TENS), orthoses, therapeutic exercises, the use of assisted devices, education, weight loss, acupuncture, laser therapy, and therapeutic ultrasound (Jamtvedt et al, 2008). Normally, For these patients surgical procedures are recommended. However, some patients are not candidates for surgery because of other underlying diseases like severe cardiovascular complications, insufficient bone stock for reconstruction, active knee sepsis, active infection, extensor mechanism dysfunction, and medically unstable cases (Hsu & Siwec, 2021) or when patients refuse surgery for any reason. So what is the solution for patients with severe knee OA whose surgery is not indicated for them? Advancing Pain and dysfunction in this group of patients leads to disability and poor quality of life. In this report, a new method for managing symptoms of advanced knee OA and its impact on the quality of life of these patients will be discussed.

2. Literature Review

Knee osteoarthritis is one of the main reasons of reduction in functional performance and physical activity secondary to pain (Thomas et al, 2003). Based on Davies et al (1991) study, patients with radiographic signs of knee osteoarthritis experience difficulties and pain while performing physical activities. Reduction in strength of muscles that are in charge of function in knee joint like quadriceps femoris muscle (Liikavainio et al, 2008) is another one of the factors that cause impaired physical function. There are several interventions for treating symptoms of the knee osteoarthritis Like exercise (reduction in pain and improvement in physical activity, moderate-quality evidence), braces and orthoses (unclear, low-quality evidence), electromagnetic field (no difference between electromagnetic fields and placebo in improving pain and function, moderate-quality evidence), acupuncture (reduces pain, moderate-quality evidence), transcutaneous electrical nerve stimulation (pain reduction, moderate-quality evidence), low-level laser therapy (reduction in pain and improvement in function, moderate-quality evidence), therapeutic ultrasound (unclear, low-quality evidence), and electrical muscle stimulation (unclear, low-quality evidence) (Jamtvedt et al, 2008). Intraarticular corticosteroid injections are other treatments for short term management of moderate to severe pain in people with osteoarthritis. Also, hyaluronic acid injections might be beneficial for pain reduction in mild OA of the knee for up to 2

years (Ayhan et al, 2014). NSAIDs can be beneficial in some cases however they have several adverse effects that influences the gastrointestinal, cardiovascular, and renal systems (Pollard et al, 2008).

In cases of failure of all conservative management the suggested treatment is surgical procedures. But it must be noted that in patients with the existence of comorbidities or active infection, ineffective quadriceps muscle, neuropathic arthropathy, poor soft-tissue coverage, psychiatric disease or alcohol, and drug abuse, inadequate bone stock, inappropriate for major surgery and anesthesia, poor motivation, and idealistic expectations surgical treatment is contraindicated (Hsu & Siwec, 2021).

In patients with severe knee OA, advancements in the levels of pain and stiffness leads to immobility which causes muscle wasting, overweighting and decreased quality of life (Alhassan et al, 2022). This situation causes an impaired cycle since exercising is limited due to pain, stiffness and fear of falling (Cumming et al, 2000) and at the same time exercising is one of the best interventions for reducing pain and stiffness and improving quality of life (Jamtvedt et al, 2008).

3. Methodology

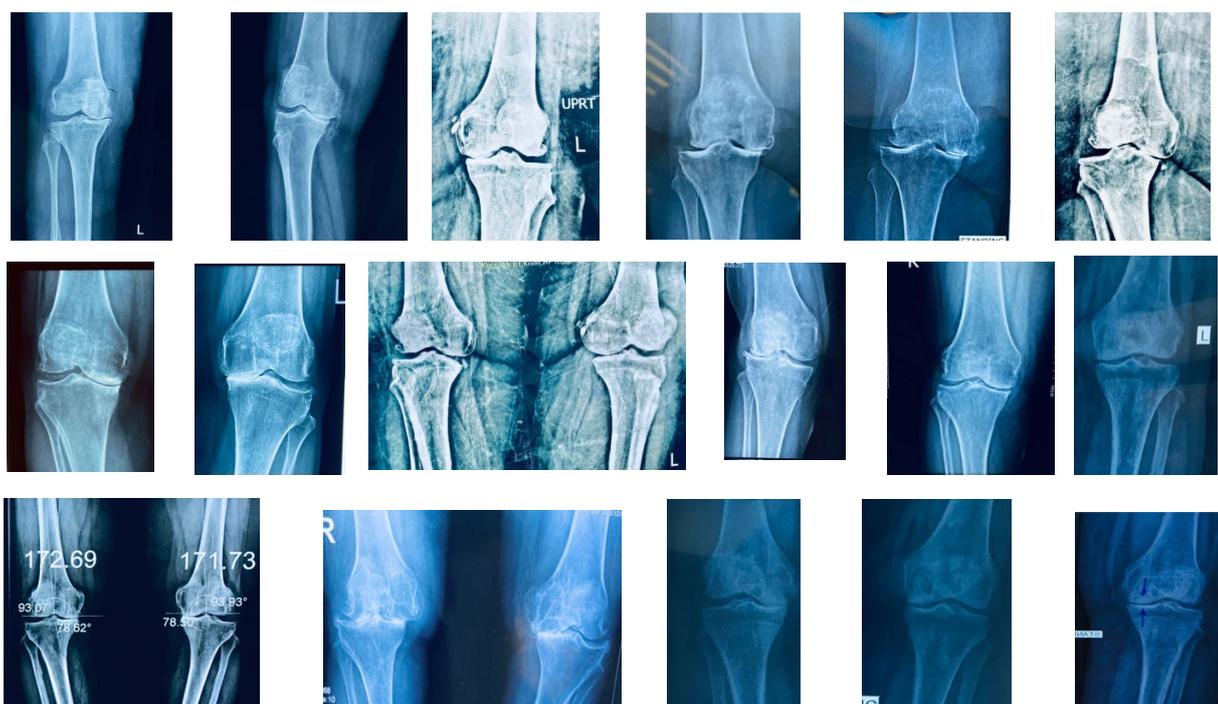
3.1 patient's characteristics :

The chosen sample consists of 17 patients (20 knee joints in total) with symptoms of knee OA confirmed by x-ray. All of them were aged 60 years or older (table 1) and based on Kellgren and Lawrence's classification (Kohn et al, 2016) they all had stage 4 knee OA. All of them were referred to orthopedic surgeons but based on their underlying diseases, surgery was not indicated and they were referred back to rehabilitation centers for non-invasive treatments. All of these cases have tried physical therapy interventions including laser therapy, extracorporeal shockwave therapy, manual techniques, TENS, and other treatments like intraarticular hyaluronic acid injection, intraarticular steroid injection, and PRP injection in the less than two years before starting this treatment. None of the above-mentioned procedures were effective for managing their pain and improving their quality of life. None of the cases had any therapeutic procedure at least 5 months before starting this course of treatment. None of them were able to perform any therapeutic exercise as they reported that any exercise would increase their pain hence they were afraid of doing it. Four of these patients were wheelchair-bound, eight of them walked with the assistance of a walker, and the rest of them walk with a noticeable limp. All of them were experiencing fear of falling which caused higher levels of immobility. Nine of them reported that the onset of a sudden increase in their symptoms was after they were infected with COVID-19. Patients were completely aware of the characteristics of the study and provided written informed consent. All of the treatment sessions were accrued in the Naji cooperative rehabilitation center and all of the patients were examined and supervised by a physical therapist and a physical medicine specialist. Patients with diabetic polyneuropathy, severe lumbar canal stenosis, and obesity were excluded from the study.

Table 1: patient's characteristics.

Patient	Sex/ Age	BMI (kg/m ²)	Duration of symptoms (M)	Walking assistance	Post COVID-19
1	F/ 69	24.8	72	walker	Yes
2	F/ 73	26.1	120	walker	Yes
3	M/ 65	23.4	66	walker	No
4	M/ 70	25.1	78	walker	No
5	M/ 74	24.7	96	wheelchair	No
6	M/ 69	27.1	90	wheelchair	No
7	F/ 66	24.2	78	–	Yes
8	M/ 75	26.1	144	wheelchair	No
9	M/ 63	23.9	72	walker	No
10	F/ 69	24.4	96	–	No
11	F/ 62	25.4	84	–	Yes

12	M/ 65	23.6	76	–	Yes
13	F/ 76	25.3	120	walker	Yes
14	F/ 68	23.4	168	wheelchair	No
15	F/ 71	26.8	84	walker	Yes
16	F/ 71	26.2	88	–	Yes
17	F/ 70	25.3	96	walker	Yes
Total	69.17(4.05)	25.04 (1.15)	95.76 (27.53)		



Figures 1- 17: Patient’s radiography: AP knee joint X-ray, patients 1 to 17 from left to right.

3.2 Intervention:

All of the patients undergo 10 sessions of intraarticular electrical stimulation during six weeks: the first week 3 sessions every other day, the second and third weeks two sessions per week, and the rest of the sessions were once a week. After 10 sessions, all of them had a maintenance treatment plan which was one session once a month for follow-up and getting the treatment if it was necessary. The electrical stimulation was provided by a nerve and muscle stimulator device and it was applied through 1 inch needles which were entered with ultrasound guidance since in this way it is more accurate than blind insertion (Fang et al, 2021). The area of entry was cleansed. Patients were supine with 90 degrees of knee flexion and an ultrasound knee examination was performed with a 6–15 MHz linear array transducer. Needles were inserted into joint space with an anterolateral, anteromedial, and mid-patellar approaches (figure 18) under ultrasound guidance. Because the needles were extremely thin (0.25mm) none of the patients reported any pain during the insertion and electrical stimulation. All of the patients were treated by one physical therapist. Each session included 15 minutes of treatment (frequency: 100 Hz) and for the rest of the time, each of them was evaluated and advised to be more physically active as much as they can without worsening their symptoms. After the fifth session, therapeutic exercises for strengthening the muscles that are in charge of movement in

lower extremities joints were added to the treatment plan. Patients were told that if any of them experience an increase in their level of pain they stop the exercise immediately and report to the physical therapist.



Figure 18: photograph of needles placement in one of the subject's knee joint. The needle insertion was under ultrasound guidance.

4. Results and Discussion

4.1 Results and reports

17 patients (10 females and 7 males, mean age 69.17 ± 4.05) with pain and stiffness in the knee joint line secondary to grade 4 knee osteoarthritis were enrolled and underwent an ultrasound guided intraarticular electrical stimulation. According to VAS and WOMAC scores, after 10 sessions all of the patients reported a significant ($p < 0.0001$) decrease in the level of pain in both rest and activity (Table 2). In addition to that, all of the subjects show a significant ($p < 0.0001$) improvement in physical activity (Table 2) and fear of falling as well as fear of performing therapeutic exercises and a reduction in stiffness level. After the 5th session when the therapeutic exercise was added to the program, patient satisfaction from being able to perform the therapeutic exercises was expressive. Two of the wheelchair-bound patients became able to walk with the assistance of a walker and one of the patients who walked with the assistance of a walker became able to walk without any assistance.

In the one-year follow-up, 12 of the patients were required to have one session per month of intraarticular electrical stimulation and they reported that this maintenance treatment keeps their symptoms and activities at an acceptable level. Another of the wheelchair-bound patients could walk with the assistance of a walker.

One of the patients got infected with COVID-19 during the one-year follow-up which cause a sudden increase in their symptoms, however, based on patients' reports the maximum level of pain that they were experiencing was much less than the time before getting treated. For this particular patient after healing from the COVID-19, six sessions of intraarticular electrical stimulation (3 sessions per week) were performed. The patient continued to do their exercises for a shorter duration during COVID, but after 6 sessions they returned to the regular program that they had before getting infected. After 6 sessions, their once a month maintenance session was resumed as the patient reported their symptoms were at the bearable level that they can perform their daily activities. Rest of the patients continue to doing therapeutic exercises and they only needed intraarticular electrical stimulation treatment when they had an unexpected increase in their level of physical activity. All of the patients stop taking NSAID medications on regular bases and they only take them when they had an episode of pain due to a sudden increase in the level of physical activity or emotional stress which was not reported more than two times in a month. None of them experienced falling in the one year. All of the scores (after 10 sessions and after the follow-up) were significant ($p < 0.0001$).

Table 2: Results.

	VAS mean [0-10] (SD)	WOMAC mean [0-96] (SD)	Physical activity (Exercise and walking)[minutes per day] (SD)
Before treatment	7.11 (0.78)	79.17 (1.74)	0
After 10 sessions	4.76 (0.90)	72.29 (2.93)	35.58(7.26)
1 year followup	4.11 (0.85)	70.82(2.83)	60.88(10.49)

WOMAC: Western Ontario and McMaster Universities Arthritis Index (pain:20 points, stiffness:8 points, physical function: 68 points). VAS: Visual Analog Scale.

After treatment patients reported that their pain had become localized. 14 cases reported that the concentration of pain was more in the medial aspect of the joint line, especially during walking which could be because of the high medial contact loads in the early stance phase (Kumar et al, 2013). However, all of the cases became able to at least perform their personal activities by themselves. It could be said that making these patients independent enough to take care of their basic personal activities is one of the most important achievements and it has a great impact on their quality of life.

4.2 Discussion

Knee osteoarthritis could be defined as a progressive degenerative disease with a low possibility of regression and repair in damaged parts of the joint (Page et al, 2011). Hence all the modalities and interventions for the management of this condition are only for controlling the symptoms of patients and improving their quality of life (Mora et al, 2018). This report concentrates on the possible ways of managing symptoms of this condition and promoting their ability to perform physical activities in patients

who are suffering from advanced stages of osteoarthritis but since they have multiple comorbidities, some systemic medications and surgical procedures are not recommended for them. Physical interventions can be the chosen treatment for this group of patients since there are no adverse effects and there is no limitation for long-term use, unlike most medications that have adverse effects on gastrointestinal, renal, cardiac, and hematological systems (Mora et al, 2018). The most important part of physical treatment is a therapeutic exercise and there are strongly recommended (both land and water-based exercises) (Mora et al, 2018. Rooij et al, 2017). But when patients' osteoarthritis is so advanced that they are not capable of performing these exercises because of their high level of pain and stiffness, what is the solution?

In this report, we investigated the influence of a new intervention that contains applying electrical stimulation into the knee joint directly using needles for patients with grade 4 knee OA (according to the Kellgren and Lawrence criteria) who did not respond well to other nonsurgical interventions and are not capable of performing therapeutic exercise because of advancement of pain level and stiffness. All of the patients responded well based on their VAS and WOMAC scores before and after the treatment and this improvement remained after a one-year follow-up. It must be added all of them had several maintenance treatment sessions. One of the most important achievements of this method was that all of the patients became capable of performing therapeutic exercises which could be one of the reasons that they experienced an improvement in performing physical activities. This achievement raised from 0 minutes to 60.88 (\pm 10.49) minutes in one year. They also reported and showed a noticeable reduction in their fear of falling and became able to walk more confidently.

Another important note is that in patients with advanced knee OA other conditions might affect patients' inability to perform physical activity and pain in the lower extremities. Although patients with neuromuscular conditions that might affect lower limbs were excluded, cardiovascular diseases also might affect the patients' ability to perform physical activity. Since one of the most important contraindications of surgery is cardiovascular diseases and they are the target group for this treatment, exercise prescription must be very careful. It must be acknowledged that none of the patients were healed completely and all of them still had symptoms of knee OA since the best treatment for patients with this advancement in knee OA only could be a surgical procedure. It is only could be recommended for patients who are not a candidate for surgery because of other comorbidities, or for patients who refuse surgery for any reason. Future studies and research is needed to confirm the effectiveness and mechanism of this method.

5. Conclusion

Based on this report, it could be concluded that this method is effective in the reduction of pain and stiffness in patients with severe knee OA and this leads to improvement in their ability to perform the therapeutic exercise and their activities of daily living. This method cannot be a substitute for surgical treatment but it can improve the quality of life in patients that cannot have surgery.

Funding: This research received no external funding.

Acknowledgments: The author thanks Dr. Naji (M.D. physical medicine specialist) for her consultation and patient referrals.

Conflicts of Interest: The author declares no conflict of interest.

References

- [1] Alhassan, E., Siaton, B. C., & Hochberg, M. C. (2022). Did COVID-19 impact osteoarthritis - clinical perspective?. *Current opinion in rheumatology*, 34(1), 68–72. <https://doi.org/10.1097/BOR.000000000000085>
- [2] Atamaz, F., Kirazli, Y., & Akkoc, Y. (2006). A comparison of two different intra-articular hyaluronan drugs and physical therapy in the management of knee osteoarthritis. *Rheumatology international*, 26(10), 873–878. <https://doi.org/10.1007/s00296-005-0096-x>
- [3] Ayhan, E., Kesmezacar, H., & Akgun, I. (2014). Intraarticular injections (corticosteroid, hyaluronic acid, platelet rich plasma) for the knee osteoarthritis. *World journal of orthopedics*, 5(3), 351–361. <https://doi.org/10.5312/wjo.v5.i3.351>
- [4] Badley E. M. (1995). The effect of osteoarthritis on disability and health care use in Canada. *The Journal of rheumatology*. Supplement, 43, 19–22.
- [5] Cooper, C., Chapurlat, R., Al-Daghri, N., Herrero-Beaumont, G., Bruyère, O., Rannou, F., Roth, R., Uebelhart, D., & Reginster, J. Y. (2019). Safety of Oral Non-Selective Non-Steroidal Anti-Inflammatory Drugs in Osteoarthritis: What Does the Literature Say?. *Drugs & aging*, 36(Suppl 1), 15–24. <https://doi.org/10.1007/s40266-019-00660-1>
- [6] Cumming, R. G., Salkeld, G., Thomas, M., & Szonyi, G. (2000). Prospective study of the impact of fear of falling on activities of daily living, SF-36 scores, and nursing home admission. *The journals of gerontology. Series A, Biological sciences and medical sciences*, 55(5), M299–M305. <https://doi.org/10.1093/gerona/55.5.m299>

- [7] Davis, M. A., Ettinger, W. H., Neuhaus, J. M., & Mallon, K. P. (1991). Knee osteoarthritis and physical functioning: evidence from the NHANES I Epidemiologic Followup Study. *The Journal of rheumatology*, 18(4), 591–598.
- [8] Fang, W. H., Chen, X. T., & Vangsness, C. T., Jr (2021). Ultrasound-Guided Knee Injections Are More Accurate Than Blind Injections: A Systematic Review of Randomized Controlled Trials. *Arthroscopy, sports medicine, and rehabilitation*, 3(4), e1177–e1187. <https://doi.org/10.1016/j.asmr.2021.01.028>
- [9] Hsu, H., & Siwiec, R. M. (2021). Knee Osteoarthritis. In *StatPearls*. StatPearls Publishing.
- [10] Jamtvedt, G., Dahm, K. T., Christie, A., Moe, R. H., Haavardsholm, E., Holm, I., & Hagen, K. B. (2008). Physical therapy interventions for patients with osteoarthritis of the knee: an overview of systematic reviews. *Physical therapy*, 88(1), 123–136. <https://doi.org/10.2522/ptj.20070043>
- [11] Kohn, M. D., Sassoon, A. A., & Fernando, N. D. (2016). Classifications in Brief: Kellgren-Lawrence Classification of Osteoarthritis. *Clinical orthopaedics and related research*, 474(8), 1886–1893. <https://doi.org/10.1007/s11999-016-4732-4>
- [12] Kumar, D., Manal, K. T., & Rudolph, K. S. (2013). Knee joint loading during gait in healthy controls and individuals with knee osteoarthritis. *Osteoarthritis and cartilage*, 21(2), 298–305. <https://doi.org/10.1016/j.joca.2012.11.008>
- [13] Liikavainio, T., Lyytinen, T., Tyrväinen, E., Sipilä, S., & Arokoski, J. P. (2008). Physical function and properties of quadriceps femoris muscle in men with knee osteoarthritis. *Archives of physical medicine and rehabilitation*, 89(11), 2185–2194. <https://doi.org/10.1016/j.apmr.2008.04.012>
- [14] Mora, J. C., Przkora, R., & Cruz-Almeida, Y. (2018). Knee osteoarthritis: pathophysiology and current treatment modalities. *Journal of pain research*, 11, 2189–2196. <https://doi.org/10.2147/JPR.S154002>
- [15] Page, C. J., Hinman, R. S., & Bennell, K. L. (2011). Physiotherapy management of knee osteoarthritis. *International journal of rheumatic diseases*, 14(2), 145–151. <https://doi.org/10.1111/j.1756-185X.2011.01612.x>List the reference here
- [16] Pollard, H., Ward, G., Hoskins, W., & Hardy, K. (2008). The effect of a manual therapy knee protocol on osteoarthritic knee pain: a randomised controlled trial. *The Journal of the Canadian Chiropractic Association*, 52(4), 229–242.
- [17] Rooij, M., van der Leeden, M., Cheung, J., van der Esch, M., Häkkinen, A., Haverkamp, D., Roorda, L. D., Twisk, J., Vollebregt, J., Lems, W. F., & Dekker, J. (2017). Efficacy of Tailored Exercise Therapy on Physical Functioning in Patients With Knee Osteoarthritis and Comorbidity: A Randomized Controlled Trial. *Arthritis care & research*, 69(6), 807–816. <https://doi.org/10.1002/acr.23013>
- [18] Thomas, S. G., Pagura, S. M., & Kennedy, D. (2003). Physical activity and its relationship to physical performance in patients with end stage knee osteoarthritis. *The Journal of orthopaedic and sports physical therapy*, 33(12), 745–754. <https://doi.org/10.2519/jospt.2003.33.12.745>

Implementing a Multimodality, Multidisciplinary Approach to Novice Nurse Success in the Emergency Department

Rocio Garcia, Guillermo Herrera, Rhiannon Rasmussen, Nicole Fraire

Abstract— University Hospital (UH) in San Antonio, Texas is a Magnet designated facility and the primary academic partner for University of Texas Health San Antonio. As the regions only combined pediatric and adult level 1 trauma and academic medical center, the emergency departments (ED) average over 87,000 visits annually.

Historically the ED at UH utilized a standardized approach to novice nurse (NN) orientation. Orientation included 12 weeks with a preceptor, bi-weekly meetings with the educator and completion of competencies. In late 2019, an analysis revealed deficiencies in the orientation process resulting in significant NN turnover and overall dissatisfaction.

High NN turnover results in negative outcomes such as decreased staff morale, decreased quality patient care, and preceptor burn out. These outcomes have effects on operations and quality throughout the health system such as financial losses that can exceed \$120K per NN and an increase in patient safety events.

A review of evidence through CINAHL, PubMed and the Cochrane library recovered 25 articles using keywords “new nurse” “orientation” “mentor,” and “education.” Filtering for dates from 2010-2020 returned 10 articles used for development of a new orientation. They consisted of a systematic review, a quasi-experimental study and a randomized experimental study. The remaining evidence were program reviews and expert opinions.

The quality improvement (QI) project stemmed from the PICOT question: Will updating the current novice nurse orientation to a multimodality and multidisciplinary approach improve retention and confidence in their clinical practice within one year? The QI project goals were decreased patient safety events, increase nurse retention, and improved confidence in their nursing practice.

In 2020, the Multimodality and Multidisciplinary Approach for Novice Nurse Success (MMANNS) was implemented. The MMANNS orientation is a program that includes 20 weeks with an experienced preceptor on the unit and bi-weekly check-ins between nursing educator, orientee and preceptor utilizing the independence rating scale. Additional requirements include attendance at emergency medicine resident grand rounds and monthly multidisciplinary lectures, completion of medication reviews, and 40 + hours of the Emergency Nurses Association: Emergency Nursing Orientation online modules. Rapid Plan-Do-Study-Act (PDSA) cycles were completed in order to implement each modality. The planning phase of the PDSA cycles included process mapping and completion of a fishbone diagram. Participants included recent NN, resident physicians, pharmacists, respiratory therapists and nursing leadership (educators and directors). The multidisciplinary approach resulted in developing a collaborative orientation for the NN.

Novice Nurse turn-over rate decreased from 53% to 16% within six months of independent practice post implementation of MMANNS.

Use of the independence rating scale throughout the NN orientation demonstrated measurable growth in knowledge, skill and self-confidence. Concurrently patient safety events have also shown a steady decrease.

Evaluation of results demonstrated effectiveness of new process to meet all goals and continued improvement. Implications: MMANNS provides variety of learning environments for addressing learners’ needs while improving peer communication and relationships.

Keywords— nurse, nurse education, nursing orientation, new nurse, novice nurse.

In Vitro Dissolution Profile of Different Brands of Low Solubility Drugs Available in the Nigerian and Brazilian Pharmaceutical Market

Ezeagu Henry Ifeanyi^{1*}, Issa Michele Georges², Ofoefule Sabinus Ifeanyi³, Ferraz Humberto Gomes⁴

Abstract— The present work describes the dissolution profiles of different brands of furosemide, glibenclamide, albendazole, ibuprofen, carvedilol and hydrochlorothiazide tablets available in the Nigerian and Brazilian pharmaceutical markets. The dissolution test was performed by following the recommendation of the United States Pharmacopeia (USP), FDA and developed method (carvedilol). A filter selection test was done for the drugs; the results obtained indicated that cannula filter was ideal for all the drugs with the exception of carvedilol which was centrifuged. The ANOVA of the filter selection showed no significant retention of drug with cannula filter ($p > 0.05$), with the exception of carvedilol ($p < 0.05$). The data obtained from the dissolution test was subjected to statistical analysis in a Microsoft excel and Minitab 17 (USA). Dissolution efficiency (%DE) was calculated for the formulations to evaluate their *in vitro* biopharmaceutical features. Tukey grouping, ANOVA and confidence interval (CI) were obtained for the comparison of the results. The ANOVA of the results indicated that the brands of (albendazole, ibuprofen, furosemide, glibenclamide, & carvedilol) were statistically different ($p < 0.05$). Hydrochlorothiazide brands were pharmaceutical equivalents ($p > 0.05$). A comparison of the results showed that 94.1% and 58.8% of the Brazilian and Nigerian brands passed respectively.

Keywords— Dissolution profile, Dissolution test, Low solubility drugs, Tablets

INTRODUCTION

Drug dissolution testing is an analytical technique used to assess release profiles of drugs from medicines, generally, from oral products such as tablets and capsules. It plays an important role as a routine quality control test, for characterizing the quality of a product^[1]. Dissolution profile comparison has been extensively used to assess drug product similarity after scale-up and post approval changes^[2].

Dissolution testing is considered to be one of the important parameters for ascertaining drug quality^[3]. This process involves the breakdown of tablets into granular or fine particulate form followed by de-aggregation, leading to its availability for systemic circulation^[4]. It also involves the interaction of solid drug with the medium resulting in the movement of drug molecules into the bulk solution^[5].

Systemic absorption of drugs is a prerequisite for eliciting their therapeutic activity, whenever given non-instantaneously. Drugs of low solubility have to be evaluated for *in vivo* bioavailability, thus, generic manufacturers must provide detailed bioequivalence evidence showing head-to-head comparative performance of their product against reference^[6].

Considering that dissolution is an *in vitro* method that characterizes how an API is extracted out of a solid dosage form, it is related with the initial stages after oral administration, that comprises the disintegration and dissolution process, thus it is a very useful tool to evaluate the quality of immediate release tablet formulations, mainly derived from class II and IV (low solubility drugs) of BCS^[7].

Eraga and his collaborators did a dissolution test of 15 different brands of metformin hydrochloride tablets marketed in South-East Nigeria, the result showed that out of the 15 different brands of metformin hydrochloride tablets analyzed, 12 of the brands passed the dissolution test while 3 failed the dissolution test^[14]. It shows that the situation of the quality of drugs marketed in the Nigerian Pharmaceutical market is not optimal when compared to the reference standard.

Based on the results of the dissolution tests shown in Table 1, it can be observed that most of the low solubility drugs marketed in the Nigerian market do not pass the dissolution test. A literature review of the dissolution tests of the Brazilian products done by Madureira et al. (2016)^[15], Mahle et al. (2008)^[16], Ferreira et al. (2016)^[17] and Gianotto et al. (2007)^[18], showed that 100% of the Ibuprofen, hydrochlorothiazide, furosemide and glibenclamide tablets passed.

It was also observed that most of the dissolution tests had lasted for 5, 7, 8, 10 and 13 years. Therefore, a current quality evaluation of the available products was done in this present work.

The objective of this present work was to elaborate a diagnosis of the dissolution test in Nigeria in comparison with Brazil, evaluating the dissolution profile of the low solubility drugs to ascertain their quality.

Dissolution test method

The dissolution test for albendazole, ibuprofen, furosemide and hydrochlorothiazide was done by following the criteria recommended by the United States Pharmacopeia^[20]. Glibenclamide was assayed following the FDA recommendation^[21]. A dissolution test method was developed for carvedilol. Apparatus 2 (paddle) was utilized for all the dissolution tests. A quantification of each drug was realized through a calibration curve which was previously constructed for each of the drugs, a straight-line graph generated from the calibration curve was used to calculate the percentage of drug

Henry Ifeanyi Ezeagu is with the University of Sao Paulo, Sao Paulo, Brazil (e-mail: henryifeanyi77@yahoo.com).

released; the percentage of drug released was plotted against time to generate a dissolution profile.

Results and Discussions

The ibuprofen, furosemide, glibenclamide, and hydrochlorothiazide drugs respectively, yielded a 100 % recovery with the cannula filter. Carvedilol did not yield 100% recovery at all volumes of discard. An analysis of variance (ANOVA) of the % recovery was done for all the drugs and a p-value of about 0.851, 0.680, 0.284, 0.129, 0.147 and 0.003 were obtained for the albendazole, ibuprofen, furosemide, glibenclamide, hydrochlorothiazide and carvedilol respectively.

The filter selection test showed that for the albendazole, the cannula filter has to be saturated by discarding about 4 mL of the dissolution medium to obtain a 100% recovery. Furosemide, glibenclamide, hydrochlorothiazide and ibuprofen yielded 100% recovery with saturation of the filter. The carvedilol was centrifuged and a 100% recovery was obtained. With a p-value greater than 0.05 ($p > 0.05$), cannula is ideal for the filtration of the drugs with the exception of carvedilol with a p-value of 0.03 ($p < 0.05$).

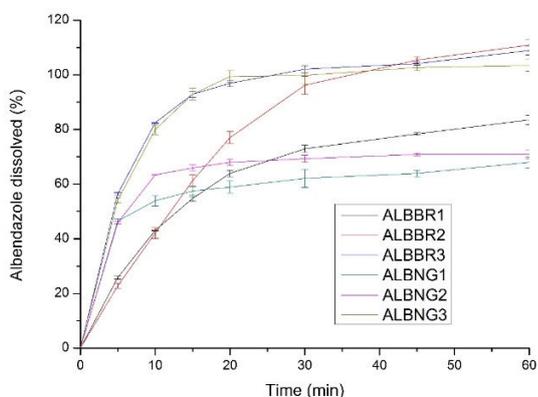


Figure 1: Dissolution profile of the brands of albendazole

Figure 1 showed the dissolution profile of the brands Nigerian (ALBNG1, ALBNG2 & ALBNG3) and Brazilian (ALBBR1, ALBBR2, ALBBR3) a more rapid release pattern followed by the ALBNG3 and ALBBR2 respectively.

Table 1: Nigerian and Brazilian products utilized in the study

The profile also showed that both the ALBBR3 and ALBNG3 brands had similar characteristics release pattern. At 45min of the assay, ALBBR3, ALBNG3 and ALBBR2 have an equivalent amount of drug released. %DE was 58.7, 64.9, 90.3, 60.4, 73.7 and 91.9% for ALBNG1, ALBNG2, ALBNG3, ALBBR1, ALBBR2 and ALBBR3 products respectively. The

results of the dissolution test showed that ALBBR1, ALBNG1, ALBNG2, ALBBR2, ALBBR3 and ALBNG3 brands released 72.9 ± 1.1 , 62.1 ± 2.5 , 69.3 ± 2.2 , 96.2 ± 0.5 , 102.1 ± 0.2 and $99.8 \pm 0.9\%$ respectively, of their active ingredients within 30 min of the assay. The analysis of variance (ANOVA) of the % drug released was evaluated for the brands and a p-value of about 0.006 ($p < 0.05$) was obtained.

The result of the Tukey grouping showed that ALBBR3, ALBNG3, ALBBR2, ALBNG2, ALBBR1 and ALBNG1 were placed in groups A, AB, ABC, ABC, BC and C respectively. From the confidence interval shown in figure 2; ALBBR2-ALBBR1, ALBNG1-ALBBR1, ALBNG2-ALBBR1, ALBNG3-ALBBR1, ALBBR3-ALBBR2, ALBNG1-ALBBR2, ALBNG2-ALBBR2, ALBNG3-ALBBR2, ALBNG2-ALBBR3, ALBNG3-ALBBR3, ALBNG2-ALBNG1 and ALBNG3-ALBNG2 crossed the zero line (limit of significance). ALBBR3-ALBBR1, ALBNG1-ALBBR3 and ALBNG3-ALBNG1 did not cross the zero line.

The rapid release pattern shown by ALBBR3 signified a rapid dissolution of the drug. The similar characteristics release pattern exhibited by ALBBR3 and ALBBR2 indicated that they are similar. The equivalent amount of drug released by ALBBR3, ALBNG3 and ALBBR2 at 45 min indicated that they can be considered as pharmaceutical equivalents. Dissolution efficiency (%DE) was calculated for the formulations and this value represents the area under the dissolution curve up to 60 min.

The % DE result showed that there was a similarity between ALBNG2 and ALBBR1; ALBNG3 and ALBBR3. Following the United States Pharmacopeia (USP) recommendation of not less than 80% release of the drug in 30 min of the assay [20]; it therefore, indicated that ALBBR2, ALBBR3 and ALBNG3 passed the test while ALBBR1, ALBNG1 and ALBNG2 failed. The p-value ($p < 0.05$) signified that the brands had statistical significant difference between them. The Tukey grouping indicated that the products ALBBR2 and ALBNG2 are pharmaceutical equivalents.

The different grouping (A, AB, BC and C) signified that the brands had statistical significant difference between them, which validated the p-value of the ANOVA result shown above. A comparison between two brands was evaluated by the confidence interval shown in figure 2. The intervals which crossed the zero line (limit of significance) as shown in figure 2 means that they had some similarity between them, while the intervals which did not cross the line had statistical significant difference.

Drug	Brazil			Nigeria		
ALBENDAZOLE	ALBBR1	ALBBR2	ALBBR3	ALBNG1	ALBNG2	ALBNG3
IBUPROFEN	IBUBR1	IBUBR2	IBUBR3	IBUNG1	IBUNG2	IBUNG3
FUROSEMIDE	FURBR1	FURBR2	FURBR3	FURNG1	FURNG2	FURNG3
GLIBENCLAMIDE	GLIBR1	GLIBR2	GLIBR3	GLING1	GLING2	GLING3
HYDROCHLOROTHIAZIDE	HYDBR1	HYDBR2	HYDBR3	HYDNG1	HYDNG2	HYDNG3
CARVEDILOL	CARBR1	CARBR2	CARNG1	CARNG2

The profile showed that IBUBR2 brand had a more rapid drug release pattern followed by IBUBR3. The graph also showed that within 20 min of the assay, all the medicines had a steady

release pattern until the 60 min time. It was also observed that from 45 until 60 min of the assay, both IBUBR2 and IBUBR3 had almost the same percentage of drug released; the same was obtainable at 15 and 20 min of the assay.

IBUNG3 and IBUBR1 showed an equivalent amount of drug released within 20 min of the assay. IBUNG2 and IBUNG1 showed the same % of drug released within 20 min of the dissolution test. Dissolution efficiency (% DE) of the IBUNG1, IBUNG2, IBUNG3, IBUBR1, IBUBR2 and IBUBR3 products gave 68.2, 67.9, 81.9, 85.9, 99.4 and 99.2% respectively. Ibuprofen dissolution showed that within 60 min of the assay, IBUBR1, IBUBR2, IBUBR3, IBUNG1, IBUNG2 and IBUNG3 had released 98.5±1.1, 108.8±2.1, 108.9±2.2, 87.2±1.3, 72.2±3.3 and 103.9±0.6% of the API content respectively.

All the medicines released more than 80% of the active ingredients with the exception of IBUNG2. The analysis of variance (ANOVA) of the ibuprofen brands gave a p-value of about 0.000 (p<0.05). A Tukey grouping of the brands placed IBUBR1, IBUBR2, IBUBR3, IBUNG1, IBUNG2 and IBUNG3 into groups AB, A, A, B, B and AB, respectively.

From the 95% confidence interval of the different brands of ibuprofen shown in figure 4, IBUBR3-IBUNG1, IBUBR2-IBUNG1, IBUBR3-IBUNG2 and IBUBR2-IBUNG2 did not cross the zero line (limit of significance), while, IBUBR2-IBUBR1, IBUBR3-IBUBR1, IBUNG1-IBUBR1, IBUNG2-IBUBR1, IBUNG3-IBUBR1, IBUBR3-IBUBR2, IBUNG3-IBUBR2, IBUNG2-IBUNG1, IBUNG3-IBUNG1 and IBUNG3-IBUNG2 crossed the zero line.

The more rapid drug release pattern shown by IBUBR2 signified rapid drug dissolution. This means that IBUBR2 had a dissolution rate faster than the other products. The almost same amount of drug released by IBUBR2 and IBUBR3 from 45 until 60 min of the assay indicated that the both medicines are pharmaceutical equivalents. Similarly, IBUNG3 and IBUBR1; IBUNG2 and IBUNG1 were also equivalents. The % DE result showed that there was a similarity between IBUNG1 and IBUNG2; IBUNG3 and IBUBR1; and IBUBR2 and IBUBR3.

The result indicated that IBUBR2 and IBUBR3 had more efficient dissolution, followed by IBUBR1, IBUNG3 and subsequently, IBUNG1 and IBUNG2. This dissolution result

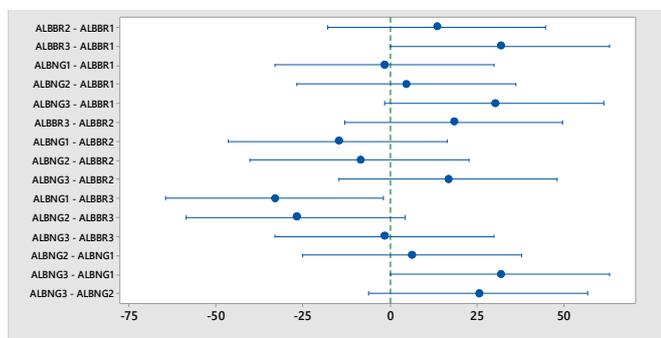


Figure 2: Confidence Interval of the Tukey grouping of the brands of albendazole.

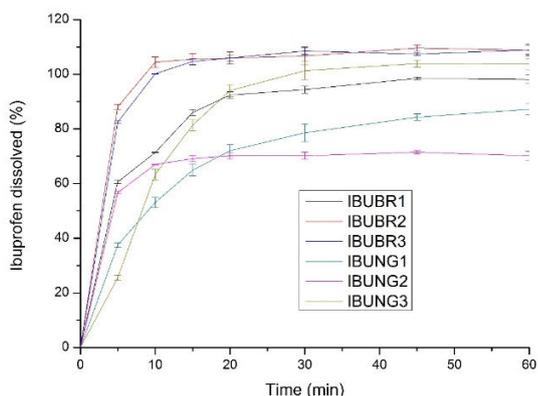


Figure 3: Dissolution profiles of the brands of ibuprofen medicines

Figure 3 showed the dissolution profile of the brands of Nigerian (IBUNG1, IBUNG2 & IBUNG3) and Brazilian (IBUBR1, IBUBR2 & IBUBR3) ibuprofen.

showed that all the brands but IBUNG2, passed the dissolution test as they complied with the recommendation of the USP method, (NLT 80 %) of the API within 60 min of the assay, should be released [20].

The almost same amount of drug (108 ± 2.1 and 108 ± 2.2) released by IBUBR2 and IBUBR3 respectively, within 60 min of the test indicated that they are pharmaceutical equivalents. The p-value of about 0.000 ($p < 0.05$) signified that the brands had statistical significant difference. The result of the Tukey grouping shown above indicated that the products IBUBR1 and IBUNG3, IBUBR2 and IBUBR3, and IBUNG1 and IBUNG2 were pharmaceutical equivalents. The different grouping validated the significant difference observed from the p value of about 0.000 ($p < 0.05$). The brands which crossed the zero line as shown in the 95% confidence interval in figure 4 did not have significant difference between them while the brands that did not cross the line had significant difference.

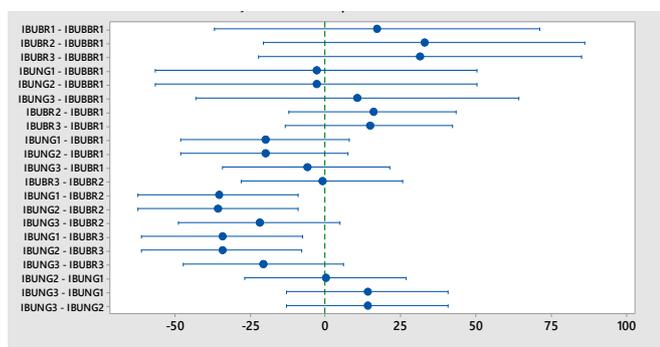


Figure 4: Confidence Interval of the Tukey of the brands of ibuprofen

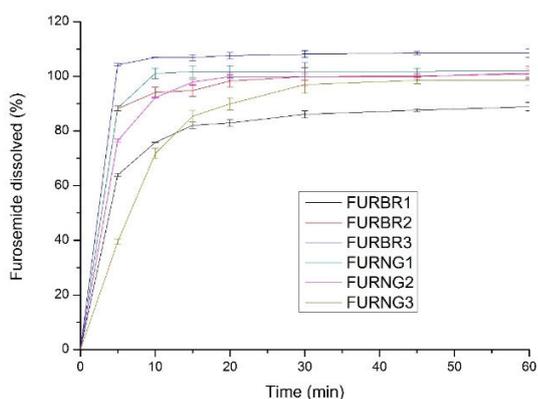


Figure 5: Dissolution profile of the brands of furosemide

Figure 5 showed the dissolution profile of the brands of Nigerian (FURNG1, FURNG2 & FURNG3) and Brazilian (FURBR1, FURBR2 & FURBR3) furosemide. The figure showed that FURBR3 had a more rapid release pattern followed by FURNG1, FURBR2, FURNG2 and FURNG3, respectively. A steady release of the drugs were observed between 30 min of the assay until 60 min. Figure 5 also showed that within 30 min

of the assay, FURBR3 and FURNG1 had released more than 100 % of their active ingredient. It was observed from the profile that within 30 min of the assay until 60 min, FURBR2, FURNG1 and FURNG2 released the same amount of active ingredients.

Dissolution efficiency (% DE) of the FURBR1, FURBR2, FURBR3, FURNG1, FURNG2 and FURNG3 products gave 81.1, 96.6, 99.1, 98.0, 95.3 and 82.9% respectively. Furosemide dissolution showed that within 60 min of the assay, FURBR1, FURBR2, FURBR3, FURNG1, FURNG2 and FURNG3 released about 88.9 ± 3.2 , 101.2 ± 0.1 , 108.5 ± 0.3 , 101.2 ± 0.2 , 100.9 ± 0.4 and $98.6 \pm 2.1\%$ respectively. The ANOVA of the brands gave a p-value of about 0.000 ($p < 0.05$). A Tukey grouping of the brands classified FURBR1, FURBR2, FURBR3, FURNG1, FURNG2 and FURNG3 into groups C, ABC, A, AB, ABC and BC, respectively.

The 95% confidence interval in figure 6 showed that FURBR3-FURNG2, FURNG1-FURNG2, FURBR2-FURNG2, FURBR2-FURBR1 and FURNG3-FURBR1 crossed the zero line (limit of significance), while, FURBR3-FURBR1, FURNG1-FURBR1, FURNG3-FURBR3 and FURNG3-FURNG1 did not cross the line.

The more rapid release pattern shown by FURBR3 indicated that it had the fastest dissolution rate compared to the other brands. The steady release of the drugs observed between 30 min until 60 min meant that the drugs had similar release characteristics. The 100% release of the API observed within 30 min by FURBR3 and FURNG1 showed that they are similar. The same amount of drug released by FURBR2, FURNG1 and FURNG2 within 60 min of the assay showed that they are pharmaceutical equivalents. The % DE result showed that there was a similarity between FURBR3 and FURNG1; FURBR2 and FURNG2; and FURBR1 and FURNG3. The dissolution values indicated that all the brands passed the test, based on USP recommendation of not less than 80% in 60 min [20]. The ANOVA value signified that the medicines had significant statistical difference. The Tukey grouping showed that FURBR2 and FURNG2 are similar. FURBR3 and FURBR1 show some variations. FURBR3 and FURNG3 had significant difference. From figure 6, the intervals that crossed the zero line did not have significant difference while, those that did not cross the line had significant difference.

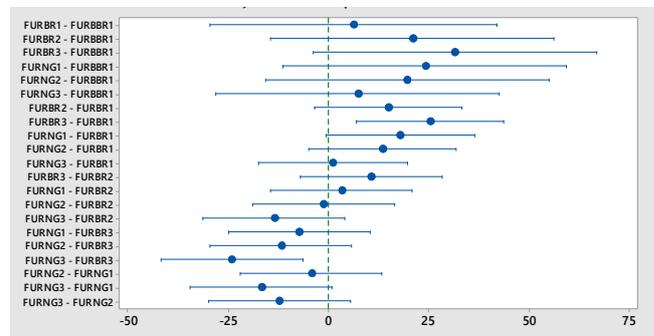


Figure 6: Confidence Interval of the Tukey of the brands of furosemide medicines.

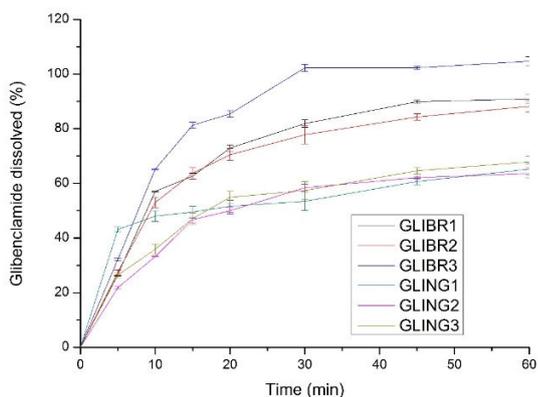


Figure 7: Dissolution profile of the brands of glibenclamide medicine

Figure 7 showed the dissolution profile of the brands of Nigerian (GLING1, GLING2 & GLING3) and Brazilian (GLIBR1, GLIBR2 & GLIBR3) glibenclamide. The profile showed that, GLING1 had the most rapid release pattern, GLIBR1 and GLIBR2 had similar release characteristics and GLING1, GLING2 and GLING3 showed equivalent release pattern. The figure also showed that the brands of GLING1, GLING2 and GLING3 released less than 80% of the API within 60 min of the assay. A calculation of the dissolution efficiency (% DE) of the GLIBR1, GLIBR2, GLIBR3, GLING1, GLING2 and GLING3 brands gave 68.9, 66.4, 81.9, 53.1, 47.9 and 50.6% respectively.

Dissolution test showed that within 60 min, GLIBR1, GLIBR2, GLIBR3, GLING1, GLING2 and GLING3 released 90.9±1.3, 88.2±2.5, 104.7±0.1, 65.2±5.2, 63.6±6.1 and 67.9±3.2% of their active ingredients respectively. The analysis of variance of the glibenclamide brands gave a p-value of about 0.014 ($p < 0.05$). A Tukey grouping of the brands placed the GLIBR1, GLIBR2, GLIBR3, GLING1, GLING2 and GLING3 brands into groups AB, AB, A, AB, B and B respectively. From the 95% confidence interval shown in figure 8, GLING2-GLING1, GLING3-GLING1, GLING2-GLIBR1, GLING3-GLIBR1 and GLIBR2-GLIBR1 crossed the zero line, GLIBR3-GLING2 and GLIBR3-GLING3 did not cross the line.

The rapid release pattern shown by GLING1 indicated that it had the most rapid drug dissolution rate. The similar release characteristics shown by GLIBR1 and GLIBR2 indicated that they are similar. Less than 80% of the API released by GLING1, GLING2 and GLING3 within 60 min of the assay showed that they did not pass the test and had inadequate API content also. The % DE result showed that there was a similarity between GLIBR1 and GLIBR2; GLING1 and GLING3; GLIBR3 and GLING2 show some variations to the other brands.

According to the FDA, not less than 80% of the API is dissolved in 60 min [21], therefore, based on the result of the dissolution test, GLIBR1, GLIBR2 and GLIBR3 passed the test, while GLING1, GLING2 and GLING3 failed. The p-value

(0.014) indicated a statistical significant difference between the brands.

The difference in Tukey grouping justified the statistical difference observed by the p-value of the ANOVA shown above. GLIBR1, GLIBR2 and GLING1 had some similarity as they were placed in group AB, similarly, GLING2 and GLING3 had some similarity as both were placed in the same group B. The intervals that crossed the zero line in as shown in figure 8 did not have significant difference, while the intervals which did not cross the line, had significant difference.

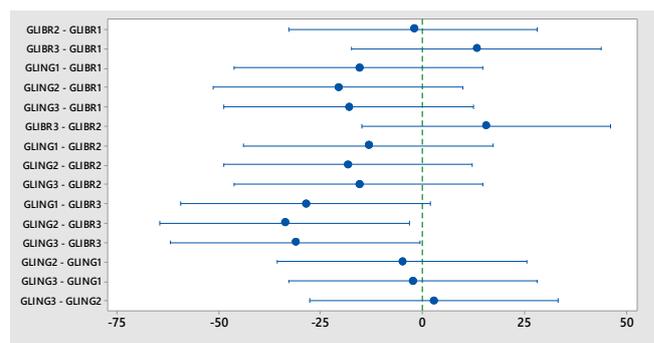


Figure 8: Confidence Interval of the Tukey of the brands of glibenclamide

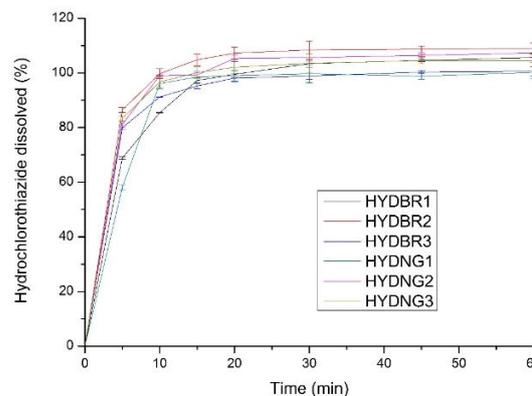


Figure 9: Dissolution profiles of the brands of HCTZ medicines

Figure 9 showed the dissolution profile of the brands of Nigerian (HYDNG1, HYDNG2 & HYDNG3) and Brazilian (HYDBR1, HYDBR2 & HYDBR3) hydrochlorothiazide. The profile showed that within 20 min of the assay until 60 min, there was a steady drug release for all the brands. Dissolution efficiency (% DE) was calculated for the brands and 94.9, 99.9, 94.9, 92.9, 99.3 and 99.2% was obtained for the HYDBR1, HYDBR2, HYDBR3, HYDNG1, HYDNG2 and HYDNG3 respectively.

Dissolution test showed that within 60 min of the assay, HYDBR1, HYDBR2, HYDBR3, HYDNG1, HYDNG2 and HYDNG3 had released 105.6±0.1, 108.9±0.3, 100.6±1.1,

100.2±2.1, 107.3±2.2 and 104.4±1.0% of the active ingredients respectively. The ANOVA of the HCT brands gave a p-value of about 0.424(p > 0.05). A Tukey grouping of the brands placed all (HYDBR1, HYDBR2, HYDBR3, HYDNG1, HYDNG2 and HYDNG3) into group A. Figure 10 showed that all the brands crossed the zero line.

The steady drug release observed for the drugs indicated that they have similar dissolution behavior. The products showed equivalent release characteristics.

The % DE showed a high degree of similarity between HYDBR2, HYDNG2 and HYDNG3. Similarly, a similarity was observed between HYDBR1 and HYDBR3. HYDNG1 was different from the other brands but not statistically significant. The dissolution results complied with the USP recommendation of not less than 80% release of the API in 60 min [20]. Thus, all the products passed the test.

This ANOVA value signified that the brands had no statistical significant difference. The Tukey grouping of the medicines into group A had validated the ANOVA value of above 0.05, which signified that the different brands had no statistical significant difference. The 95% confidence interval of the brands of medicines shown in figure 10 indicated that the products did not have significant difference.

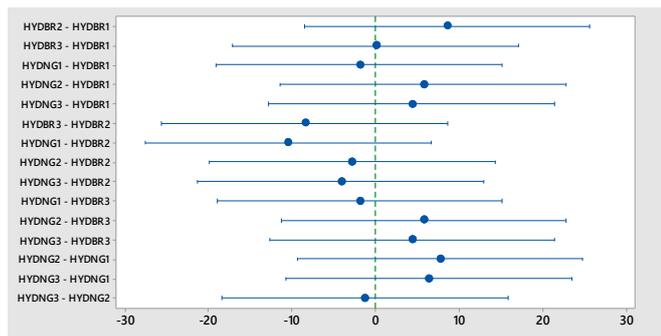


Figure 10: Confidence interval of the brands of hydrochlorothiazide medicine

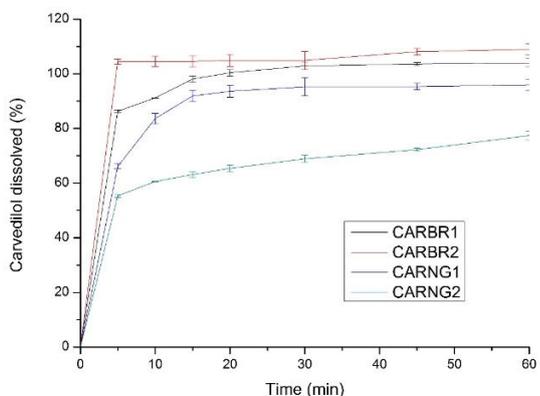


Figure 11: Dissolution profiles of the brands of carvedilol medicine

Figure 9 showed the dissolution profile of the brands of Nigerian (CARNG1, CARNG2) and Brazilian (CARBR1, CARBR2) carvedilol.

The profile showed that CARBR2 brand had the most rapid release pattern followed by CARBR1. It also showed that CARNG2 had the lowest release pattern. The medicines had steady drug release within 15 min of the assay until 60 min. All the medicines, except CARNG2 released more than 80% of their API within 15 min of the assay. CARBR2, CARBR1 and CARNG1 showed similar release characteristics.

Dissolution efficiency (% DE) of the CARBR1, CARBR2, CARNG1 and CARNG2 gave 98.1, 99.5, 88.8 and 66.1% respectively. Carvedilol dissolution showed that CARBR2 released more than 100% of its active ingredients within 5 min of the assay, CARBR1 released 100% of the API within 20 min, CARNG1 did not release 100% of the active content even at infinity time of the assay but released more than 80% within 10 min. CARNG2 did not release 100% of the API at infinity, as 83.817% release was observed. Within 60 min of the assay, CARBR1, CARBR2, CARNG1 and CARNG2 released 104.0, 108.9, 95.9 and 77.4% of the API respectively. The ANOVA of the brands gave a p-value of about 0.000 (p<0.05). A Tukey grouping placed the CARBR1, CARBR2, CARNG1 and CARNG2 into group AB, A, B and C respectively. Figure 12 showed that CARBR2-CARBR1 and CARNG1-CARBR1 crossed the zero line while, CARNG2-CARBR1, CARNG1-CARBR2, CARNG2-CARBR2 and CARNG2-CARNG1 did not cross the line.

The rapid release pattern shown by CARBR2 signified a rapid dissolution of the drug. The lowest release pattern shown by CARNG2 indicated a slow response to dissolution. The release profile shown by CARBR2, CARBR1 and CARNG1 indicated that they are similar. The % DE result showed that there was a similarity between CARBR1 and CARBR2. CARNG1 was different from CARBR1 and CARBR2 but not statistically significant. CARNG2 had statistical significant different from the other brands. Based on the result of the dissolution, CARBR1, CARBR2, and CARNG1 passed the test but CARNG2 failed.

The ANOVA p-value signified that the brands had statistical significant difference. The difference in Tukey grouping had affirmed the p-value of the ANOVA shown above, meaning that the brands had statistical significant difference. From the 95% confidence interval shown in figure 12, the brands that crossed the line did not have significant difference while the brands which did not cross the line had significant difference.

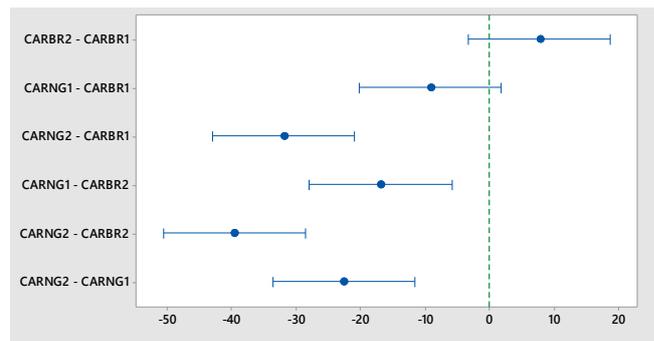


Figure 12: Confidence interval of the brands of carvedilol medicine

General evaluation of the Brazilian and Nigerian brands

Table 4: Percentage of products which passed the dissolution test.

Drug	Brazilian Brands	Nigerian Brands
Albendazole	66.7%	33.3%
Ibuprofen	100%	66.7%
Furosemide	100%	100%
Glibenclamide	100%	0%
Hydrochlorothiazide	100%	100%
Carvedilol	100%	50%

Table 4 showed the summary of the dissolution test of the brands of medicines in this present work. It was observed that out of the 17 brands each of the Brazilian and Nigerian medicines tested, 94.1% and 58.8% passed respectively. However, none of the glibenclamide products from Nigeria passed. The furosemide and hydrochlorothiazide samples from both countries passed. It was obvious that a higher percentage of Brazilian products passed, which may be attributed to the higher availability of generic medicines in the Brazilian pharmaceutical market which makes it difficult for adulteration or faking to take place. As most Nigerian medicines were imported from foreign countries, it makes it easier for adulteration or falsification of the products.

CONCLUSION

From the results of the dissolution studies performed on the Brazilian and Nigerian brands of medicines, it can be concluded that a higher percentage of Brazilian medicines (about 94.1%) passed the dissolution test as compared to a lower percentage of the Nigerian medicines (about 58.8%) that passed. The brands of medicines tested showed statistical significant difference with the exception of hydrochlorothiazide. The presence of generic medicines in the Brazilian pharmaceutical market may be one of the contributory factors to the differences in dissolution observed between the Nigerian and Brazilian brands. This observation has enabled a diagnosis of the dissolution test in Nigeria in comparison with Brazil.

REFERENCES

1. FDA Guidance for industry: Dissolution Testing of Immediate Release Solid Oral Dosage Form. *Rockville, MD: FDA.* 2015.

2. FDA Guidance for Industry; Dissolution Testing of Immediate Release Solid Oral Dosage Forms; U.S. Department of Health and Human Services, Food and Drug Administration, U.S. Government Printing Office: *Washington, DC , FDA,* 2007.
3. Zhu H, Zhou H, Seitz K. Bioavailability & Bioequivalence in Developing Solid Oral Dosage Forms. USA. *Academic Press,* 2009: 341-364.
4. Rahul M, Piyoosh S, Tekade M, Atneriya U, Dua K, Hansbroe PM. Microsponge Embedded Tablet for Sustained Delivery of Nifedipine. *Pharm. Nanotechnol,* 2017.
5. Qiu Y, Chen Y, Zhang GG, Yu L, Mantri RV. Developing Solid Oral Dosage Forms: Pharmaceutical Theory and Practice. *Academic Press,* 2016.
6. FDA Guidance for industry: Waiver of *in vivo* Bioavailability and Bioequivalence Studies for Immediate Release Solid Oral Dosage Forms Based on a Biopharmaceutics Classification System. *Rockville, MD: FDA,* 2017.
7. Ferraz HG, Carpentieri LN, Watanabe SP. Dissolution Profile Evaluation of Solid Pharmaceutical Forms Containing Chloramphenicol Marketed in Brazil. *Brazilian Archives of Biology and Technology,* 50(1), 2007: 57-65.
8. Eraga SO, Idemili CAD, Iwuagwu MA. Pharmaceutical Quality Assessment of Brands of Metformin Hydrochloride Tablets Available in South-East Nigeria. *African Journal of Pharmaceutical Research & Development,* 12(3), 2020:333-341.
9. Madureira GLN, Brigida CRB, Fernando TAN. Avaliação do Perfil de Dissolução de Comprimidos de Ibuprofeno. *Jornada de Farmácia,* 2016.
10. Mahle F, Goelzer F, Adriano J, Felipe M, Vier N, Carli RBG, Rosa T, Couto AG, Lucinda SRM. Evaluation of the Dissolution Profile of Hydrochlorothiazide Tablets Marketed in Brazil. *Brazilian Journal of Basic and Applied Pharmaceutical Sciences,* 28(3), 2008:265-271
11. Ferreira DSM, Viana CMGL, Matos AMR, Ramon RS, Fernanda ADSS, Milleno DM, Edith CLC, Hemerson IFM, Anibal DF, Santos J. Comparative *In vitro* Analysis of Dissolution Profiles of Furosemide Tablets Marketed in Bahia, Brazil. *Latin American Journal of Pharmacy,* 35(9), 2016: 2064-2070.
12. Gianotto ADSE, Renata PA, Maurilio JLF, Alexandre CSCF, Marlene MFN. Dissolution Test for Glibenclamide Tablets. *Quim. Nova,* 30(5), 2007:1218-1221.
13. UNITED STATES PHARMACOPEIA (USP). Dissolution of Solid Pharmaceutical Dosage

Forms. *Pharmacoepial Forum*, 27(3), 2016:711-2505.

14. UNITED STATES FOOD AND DRUG ADMINISTRATION (USFDA). FDA's *Dissolution Guidance*, 2018.

Computer-Aided Drug Repurposing for Mycobacterium Tuberculosis by Targeting Tryptophanyl-tRNA Synthetase

Neslihan Demirci, Serdar Durdađı

Abstract— Mycobacterium tuberculosis is still a worldwide disease-causing agent that, according to WHO, led to the death of 1.5 million people from tuberculosis (TB) in 2020. The bacteria reside in macrophages located specifically in the lung. There is a known quadruple drug therapy regimen for TB consisting of isoniazid (INH), rifampin (RIF), pyrazinamide (PZA), and ethambutol (EMB). Over the past 60 years, there have been great contributions to treatment options, such as recently approved delamanid (OPC67683) and bedaquiline (TMC207/R207910), targeting mycolic acid and ATP synthesis, respectively. Also, there are natural compounds that can block the tryptophanyl-tRNA synthetase (TrpRS) enzyme, chuangxinmycin, and indolmycin. Yet, already the drug resistance is reported for those agents. In this study, the newly released TrpRS enzyme structure is investigated for potential inhibitor drugs from already synthesized molecules to help the treatment of resistant cases and to propose an alternative drug for the quadruple drug therapy of tuberculosis. Maestro, Schrodinger is used for docking and molecular dynamic simulations. In-house library containing ~8000 compounds among FDA-approved indole-containing compounds, a total of 57 obtained from the ChemBL were used for both ATP and tryptophan binding pocket docking. Best of indole-containing 57 compounds were subjected to hit expansion and compared later with virtual screening workflow (VSW) results. After docking, VSW was done. Glide-XP docking algorithm was chosen. When compared, VSW alone performed better than the hit expansion module. Best scored compounds were kept for ten ns molecular dynamic simulations by Desmond. Further, 100 ns molecular dynamic simulation was performed for elected molecules according to Z-score. The top three MMGBSA-scored compounds were subjected to steered molecular dynamic (SMD) simulations by Gromacs. While SMD simulations are still being conducted, ponesimod (for multiple sclerosis), vilanterol (β_2 adrenoreceptor agonist), and silodosin (for benign prostatic hyperplasia) were found to have a significant affinity for tuberculosis TrpRS, which is the propulsive force for the urge to expand the research with in vitro studies. Interestingly, top-scored ponesimod has been reported to have a side effect that makes the patient prone to upper respiratory tract infections.

Keywords— drug repurposing, molecular dynamics, tryptophanyl-tRNA synthetase, tuberculosis.

N.Demirci is with Computational Biology and Molecular Simulations Laboratory, Department of Biophysics, School of Medicine, Bahçeşehir University, 34353, Istanbul, Turkey (e-mail: neslihan.demirci@bahcesehir.edu.tr).

S.Durdađı is with Computational Biology and Molecular Simulations Laboratory, Department of Biophysics, School of Medicine, Bahçeşehir University, 34353, Istanbul, Turkey (e-mail: serdar.durdagi@med.bau.edu.tr).

Accurate Bluetooth Proximity Estimation with Machine Learning Algorithms for Private Automated Contact Tracing to limit the spread of COVID-19

Collin Li

Collin Li is with the Great Neck, United States (e-mail: collinli2022@gmail.com).

Abstract

The human virus SARS-CoV-2 has resulted in high fatality rates and incapacitated health systems. Thus, the ability to limit transmission is a priority. Although COVID-19 is spreading too fast to be contained by manual contact tracing, it seems plausible that if this method was more efficient and happened at scale, transmissions could be significantly reduced. A contact-tracing app that stores a memory of proximity contacts and immediately notifies positive case contacts can contribute to epidemic control if used by enough people. By targeting only those at risk that could have been exposed to the virus, epidemics could be contained without harmful mass quarantines. This approach is known as Private Automated Contact Tracing, which attempts to develop exposure detection functions while preserving privacy by using Bluetooth Low Energy for proximity detection. This study analyzes the inner workings of Bluetooth to construct the most accurate proximity detection by testing various environments and factors that may impact Bluetooth signal strength. Along with experiments conducted with raspberry pi computers, data from MITRE Corporation was used to train supervised machine learning algorithms for Bluetooth proximity detection, which accounted for possible environmental factors that affect Bluetooth signals such as physical environments, the orientation of the devices, etc. By accounting for environmental factors, model accuracy improved by around 30%, and the Kalman filter used during feature engineering improved accuracy by an additional 5%. Decision Trees and Random Forest machine learning models performed the best, with accuracies of around 83%. Finally, PACT is also applicable in reducing the spread of any future potential highly infectious and life-threatening diseases.

Keywords: machine learning, artificial intelligence, automated contact tracing, covid-19

1. Introduction

In an attempt to limit the spread of the COVID-19 crisis, Private Automated Contact Tracing (PACT) arose to enhance Contact Tracing by designing exposure detection functions in personal digital communication devices that have maximal public health utility while preserving privacy.

Contact Tracing is the process by which medical professionals trace the route of transmissions starting from a positively diagnosed patient. This way, they aim to break the chain of transmission by notifying and taking preventative measures against further transmissions. Current contact tracing techniques are slow and inefficient and require a person to recall everybody they came into contact with (Ferretti, 2020).

Countries like Singapore, Germany, Iceland, China, and South Korea have already developed contact tracing apps to limit the spread of COVID-19 (Bay, 2020). However, some of these technologies come at the expense of privacy as they are centralized and use GPS to track one's position. Therefore, a decentralized approach for using personal digital communication devices for automating exposure detection using Bluetooth® Low Energy signaling was proposed to maintain the user's privacy (Rivest, et al. 2020). By using Bluetooth instead of GPS, the user's location cannot be determined, only nearby devices can be identified for proximity detection.

For this to work, accurate and efficient proximity algorithms must be developed to determine the distance between two devices. According to studies of the transmission of COVID-19, maintaining a 6ft distance for social distancing is crucial as a large portion of transmissions "were community members who spent at least 10 minutes within 6 feet of a patient with confirmed disease." (Burke, 2020). Thus, the general concept is that when an algorithm determines that the two devices are too close (≤ 6 feet) for a duration of time, the other's universally unique identifier (UUID), an integer number, would be stored for later use. When the other user is diagnosed with COVID-19, he or she may anonymously publish their phone's UUID numbers to a database where others may check if they have gone in close contact by checking if their phone saved one of the UUID numbers. The proximity detection aspect is based on the received signal strength indication (RSSI) between two Bluetooth devices/chirps (STMicroelectronics, 2020). A strong received signal strength (RSS) would suggest a close distance, while a weak RSS would suggest a far distance between the two devices. This project will investigate the inner workings of RSS while building a successful algorithm to determine the proximity between two Bluetooth devices.

2. Bluetooth Low Energy

Bluetooth, first introduced in 1989, has had a profound impact on our digital world. This technology was "designed for a short-range wireless transmission while maintaining low energy consumption, small size, and low cost" (Sthapit, 2018). Now, almost every device is Bluetooth compatible.

The detection concept relies on the fact that the Bluetooth system provides a device discovery protocol for allowing a device to find neighbor devices before establishing a piconet link (Naya, 2005). One device acts as a beacon, which advertises, while another device scans for beacons. This project uses the iBeacon advertising format due to its simplicity and its ability to work on both IOS and Android operating systems. During this discovery protocol, information is sent in 30-byte packets.

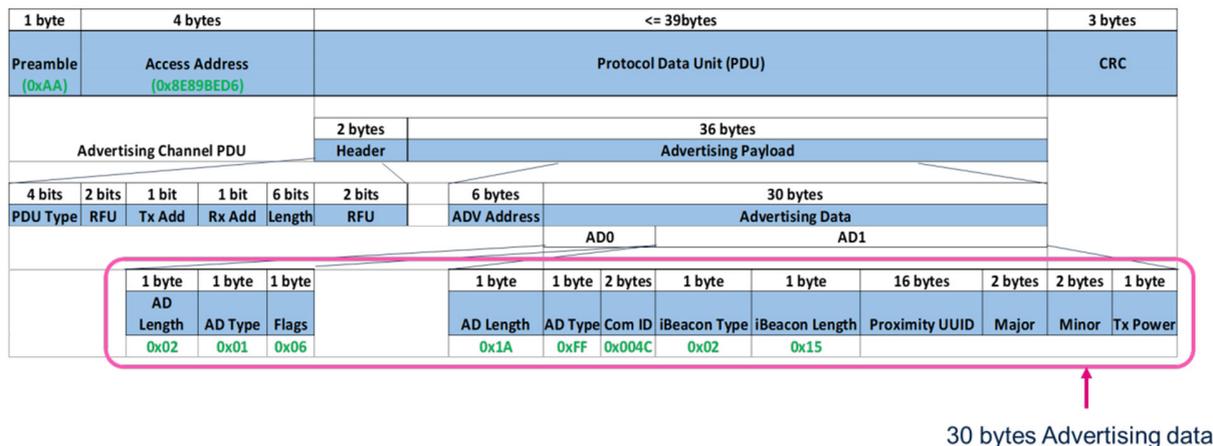


Figure 1 shows the iBeacon advertising packet format.
Source: (Singhal, 2020)

The center of the proximity detection is based on the received signal strength indication (RSSI). A strong RSSI would suggest a close distance, while a weak RSSI would suggest a far distance between the two devices. Duration could also be measured by subtracting the first time from the last time the device/Bluetooth chirp was seen.

3. NY State Contact Tracing

However, this Bluetooth proximity detection is not so straightforward. For a number of reasons, RSSI does not include a simple and consistent measure of distance. In a recent study done by TraceTogether for Singapore’s proximity detection app, Singapore researchers found that due to differences in bluetooth hardware, antenna layout, and even operating system configurations led to RSSI variations by as much as 20 dBm in highly controlled environments (Robinson & Waldo, 2021). In terms of measured distance, a difference of 20 dBm approximately corresponds to a predicted output that could range from 1 meter to 10 meters. The NY State Contact Tracing system does not account for these environmental factors. Instead, it uses a static threshold setting shown in figure 2 which is an inaccurate way of calculating proximity and exposure to COVID-19.

Minutes-at-attenuation weight



Bluetooth attenuation (dB)

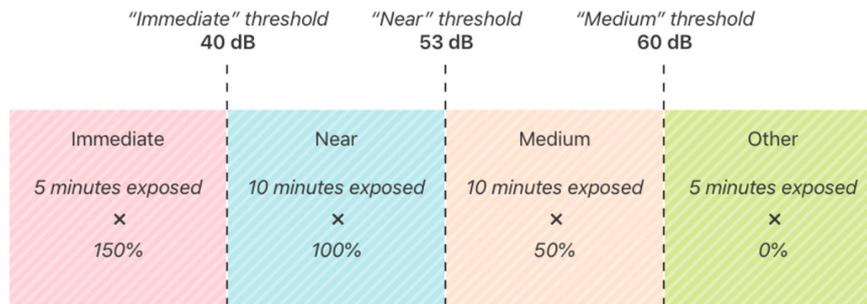


Figure 2 indicates how Exposure Risk Value is calculated

Source:

<https://developer.apple.com/documentation/exposurenotification/enexposureconfiguration>

4. Kalman filter

The Kalman filter is a set of mathematical equations for combining information in the presence of uncertainty. It provides an efficient computational (recursive) solution of the least-squares method, supporting estimations of past, present, and even future states (Welch, 1995). For this project, the Kalman filter was used to process and smooth the Bluetooth RSS signal as RSS can greatly fluctuate depending on the environment, caused by signal refraction.

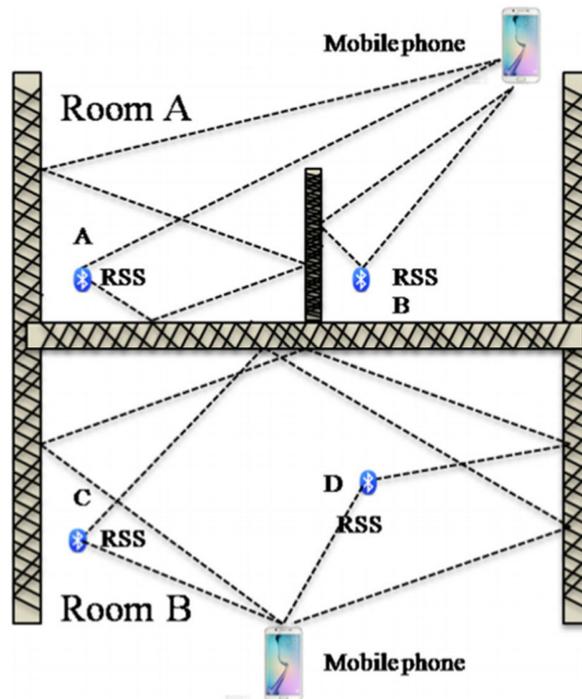


Figure 3 shows an example of the refraction of RSS resulting in a weaker RSS.
Source: (Singhal, 2020)

5. (Supervised) Machine Learning

Machine learning is a method of data analysis that automates analytical model building and automatically improves with experience through a learning process (Pedregosa, 2011). It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention. Various machine learning algorithms and BLE, along with fingerprinting and triangulation can be used together to accurately estimate the location of a mobile device (Sthapit, 2018). However, this requires an extensive amount of setup and a controlled environment. Other possible methods for positioning include k-Nearest Neighbor and Naive Bayes' Classifier. (Bekkelien, 2012). Yet again, this requires an extensive amount of setup and a controlled environment for fingerprinting and triangulation.

This project utilized multiple popular supervised learning algorithms—Linear Regression, K-Nearest Neighbor, Random Forests, Decision Trees, and Neural Networks—to compare the performance of these algorithms and ultimately construct a valid system for Bluetooth proximity detection for various environments.

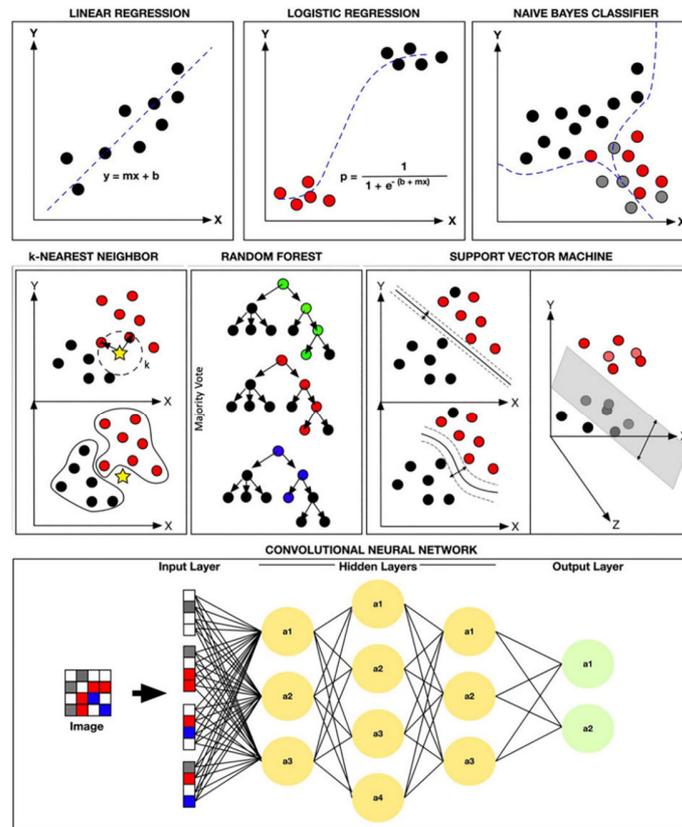


Figure 4 shows a comparison of popular supervised learning methodologies. Source: (Rashidi, 2019)

6. Deep Learning & Neural Networks

Since the advent of the 21st century, computer scientists have been researching ways to create a true artificial intelligence system. In other words, computer scientists are trying to create a machine capable of thinking and processing information just like humans, only in a way that is faster and more advanced than the human mind. One major step towards a true A.I system has been the creation of deep learning, a computer technique located under the branch of machine learning. Machine learning is the idea that computers are capable of performing specific tasks without being given any instruction and can only rely on patterns and inferences. Deep-learning helps computers recognize and utilize patterns and inferences among data. The way that this is accomplished is through the use of artificial neural networks.

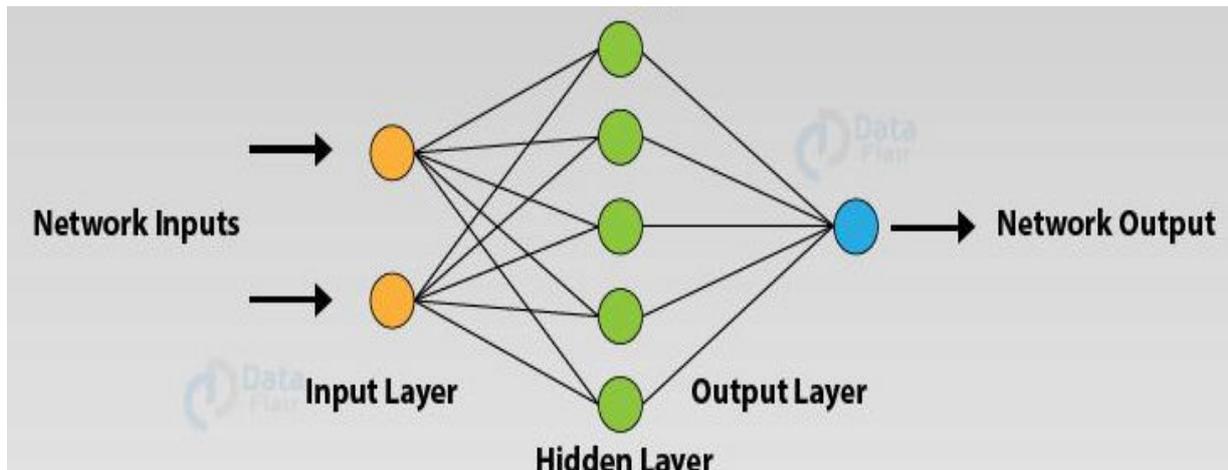


Figure 5 shows the general layout of every artificial neural network.
 Source: <https://data-flair.training/blogs/artificial-neural-networks-for-machine-learning/>

Every neural network has three types of layers: an input layer, a hidden layer, and an output layer. The input layer consists of neurons that each hold a value between 0 and 1. Typically, each neuron is used to represent a pixel of a picture. The value each neuron receives from the pixel is determined by how active the pixel is. In other words, the higher the RGB value of a pixel, the higher the value of a neuron will be.

Each neuron in the input layer is used to determine the value for one neuron in the next sequential layer. Hidden layers are used to determine what features are contained in an image. However, these values are altered through the use of weights and biases. Weights represent the importance of each input neuron to a neuron in the first hidden layer. Each neuron has a different weight from the other neurons in the same layer and different weights depending on which neuron in the second layer it is corresponding to. This procedure repeats until the network reaches the output layer. At this point, the network simply returns the neuron with the highest value.

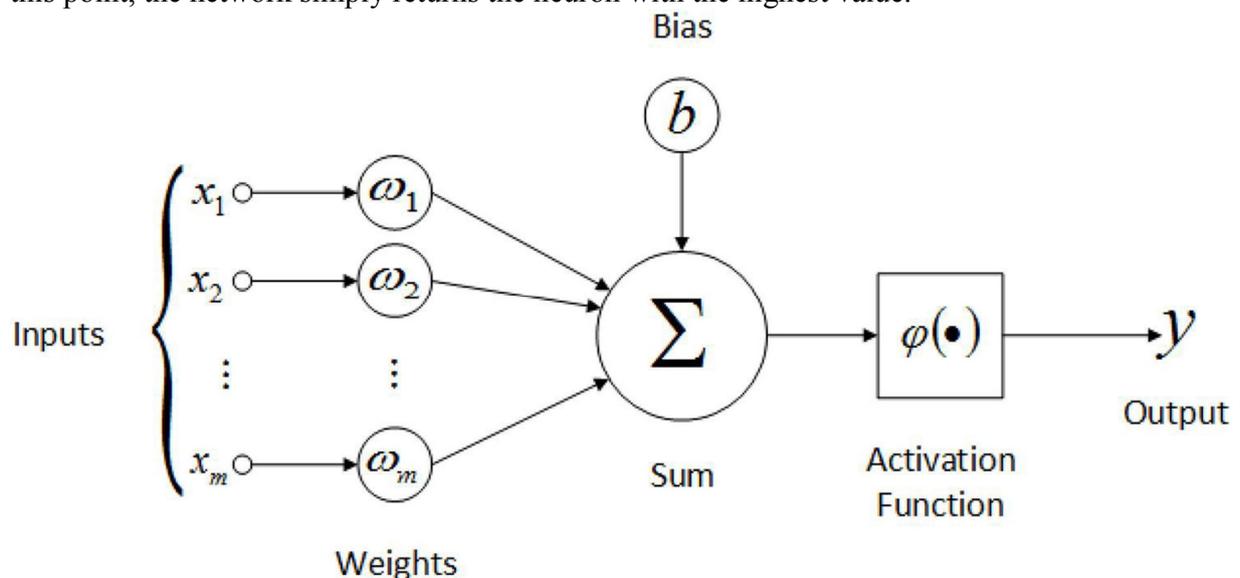


Figure 6 shows the general layout of a neuron in a neural network with an input layer, hidden layers, and output layers. In each neuron, there are usually inputs, weights, biases, a net input function, and an activation function.
 Source: <https://towardsdatascience.com/deep-learning-versus-biological-neurons-floating-point-numbers-spikes->

and-neurotransmitters-6eebfa3390e9

7. Hypotheses

RSSI can range from anywhere between -20 decibel milliwatts (dBm) (a few inches) to over -100dBm (tens to hundreds of feet). In a perfect scenario with no outside factors, each RSSI value should map to a constant distance. However, due to many environmental factors, unstable beacon signals, different Bluetooth chips, the direction the chip is facing, et cetera, a close to perfect distance calculation is extremely difficult to achieve.

Thus, is it possible to develop an efficient and relatively accurate (machine learning) detection algorithm that can account for the many different factors that impact RSSI?

8. Methodology

8.1 Experimental Section

To test and experiment with the possible independent factors that may affect RSSI, experiments were performed using two Bluetooth-capable Raspberry Pi 4s. Using the reference code provided by Beaver Works Summer Institute piPact, which utilized the python library pybluez[ble], one device acted as a Bluetooth iBeacon while the other device scanned for beacons. The result would be a comma-separated values file that stored the RSSI values of each scan, along with other values such as TX power, timestamp, a universally unique identifier (UUID), etc. With this setup, I could experiment and test the various effects of various factors. I have collected data on varying angles, materials, and heights interfering with the Bluetooth beacon and scanner. The materials I have tested include Cardboard, Cloth, Blankets, Walls, and Jackets.

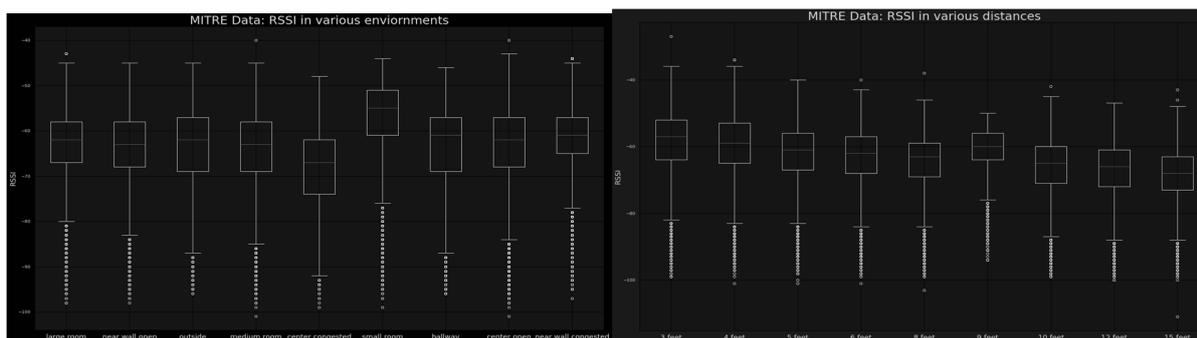
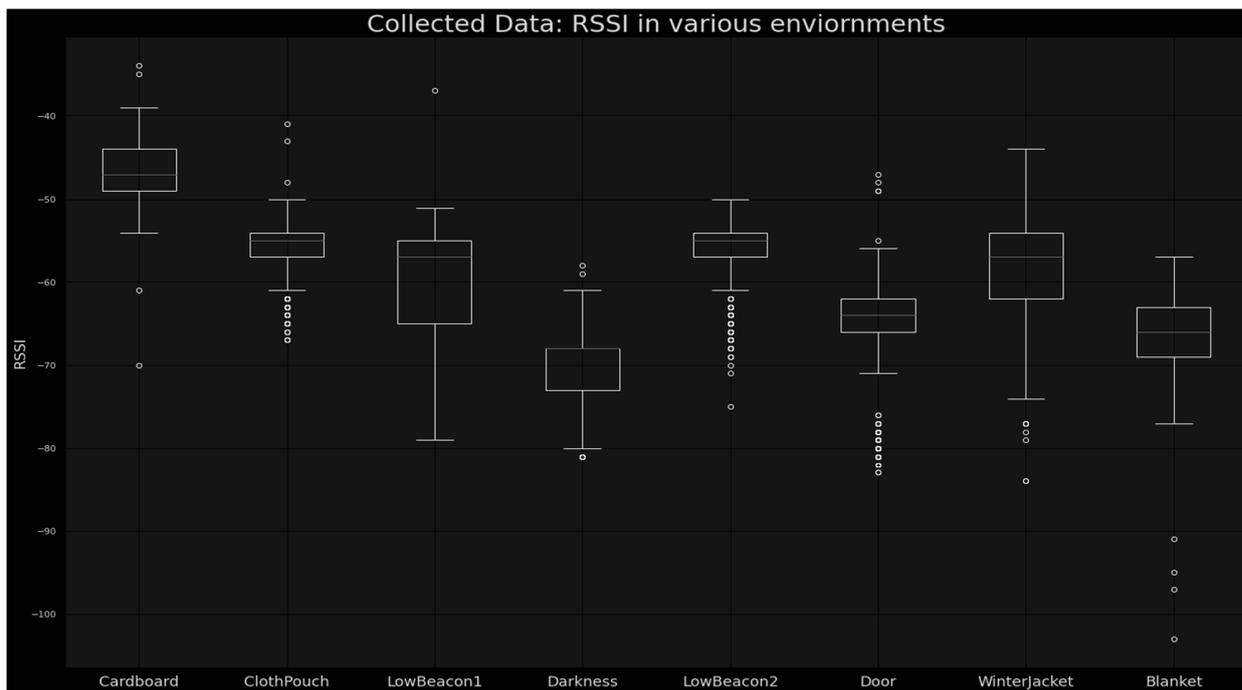


Figure 7a (top) shows 13,207 RSSI values in various environments of devices 6 feet apart. Data collected by student.

Figure 7b (bottom left) shows 123,063 RSSI values in various environments of devices 6 feet apart. Data from MITRE Corporation.

Figure 7c (bottom right) shows 923338 RSSI in various distances. Data from MITRE Corporation.

Sources: Student Generated

Graphs 7a, 7b, and 7c show RSSI collected from devices 6 feet apart in various environments and demonstrate that the RSSI, in a real-world environment, does not easily correlate with distance. Focusing on the experiment of a Cardboard vs Blanket blocking the bluetooth signal, shown in figure 7a, the RSSI medians have a significant 20 dBm RSSI difference. This indicates that a proximity detector that does not account for environmental factors could be off by 1 to 10 meters.

8.2 Model Construction Section

All machine learning models trained in this project were supervised learning algorithms. Machine learning was used because there were many inputs to be accounted for and the relationships were complex. A crucial part of machine learning is data preparation and manipulation. The original dataset, collected from The MITRE Corporation, was collected based on their *Range-Angle Collection Protocol*. This dataset included an abundance of values:

- Range
- RSSI
- The types of devices
- The type of environment
- The pose of the person holding the devices
- The location of the phone on the person (pants pocket, purses)

Data Preparation

Some argue that the most crucial step in training a machine learning algorithm is analyzing the data, also known as feature engineering. There were multiple steps in this complex process.

- I. Combine Datasets
 - A. The dataset from The MITRE Corporation was split into hundreds of comma-separated delimited files that were sorted depending on their testing conditions (independent variables such as range, type of devices, etc). Thus, the first step was to combine all of these values into one file. This step was completed using python libraries such as the python os module and CSV module.
- II. Create target variable
 - A. The dataset from The MITRE Corporation provided RSSI collected from a variety of ranges: 3, 4, 5, 6, 8, 9, 10, 12, 15 (in feet). Through thoughtful reasoning, it was decided that a binary classification algorithm would be better suited for this project. This would result in a simpler training process and does not significantly impact contact tracing as research showed that COVID is most likely to be transmitted within a 6ft radius.
 - B. To accomplish this, all ranges less than or equal to 6 feet were set to 0 and all ranges greater than 6 were set to 1.
- III. Remove outliers
 - A. The dataset was then analyzed through various graphs to spot outliers or any patterns. It was quickly noticed that there were 1600 RSS values equal to 127, the indication that the Bluetooth beacon and scanner lost connection. They were deleted from the dataset.

- B. The RSS values were also inverted (multiplied by -1) so all the RSS values would be positive.

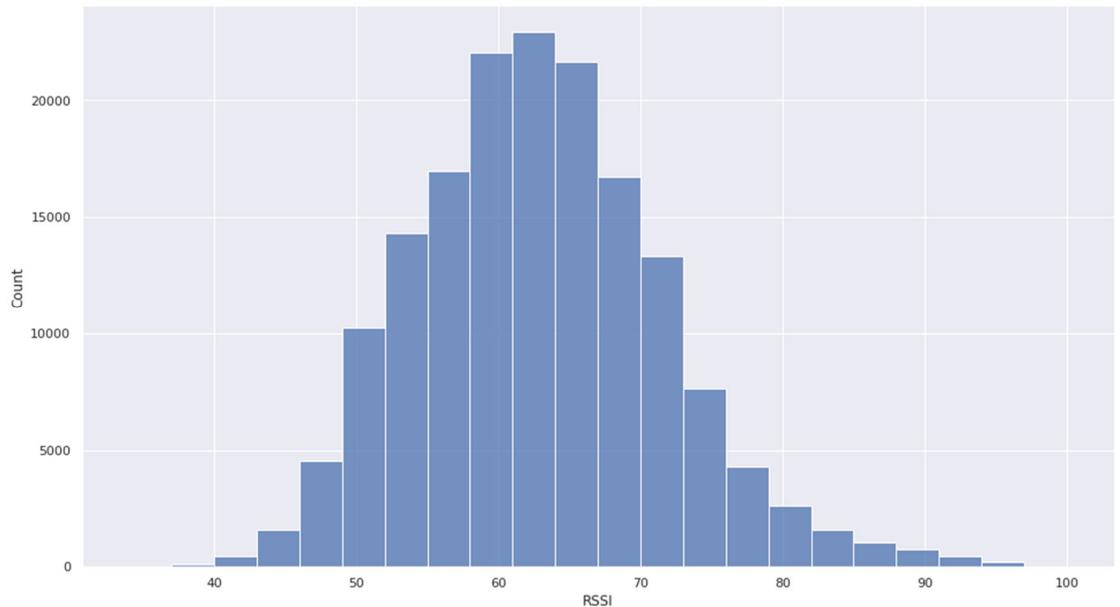


Figure 8 shows the occurrence of the RSS values of devices 6 feet apart following a bell shape with a skew to the right.

Source: Student Generated

IV. Kalman Filter

- A. A simple implementation of the Kalman Filter for 1D data was used for filtering the RSSI. This implementation is a recursive algorithm that takes the history of the previous measurements into account.

$$x_t = A_t x_{t-1} + B_t u_t + \epsilon_t$$

Figure 9 shows the main mathematical equation used in the Kalman filter.

Source: <https://www.wouterbulten.nl/blog/tech/kalman-filters-explained-removing-noise-from-rssi-signals/>

- B. The equation above shows a general form of the transition model. A is a transformation matrix, x_t is the current state (current RSSI), x_{t-1} is the previous state (previous RSSI), u is the control input, and ϵ is the noise.

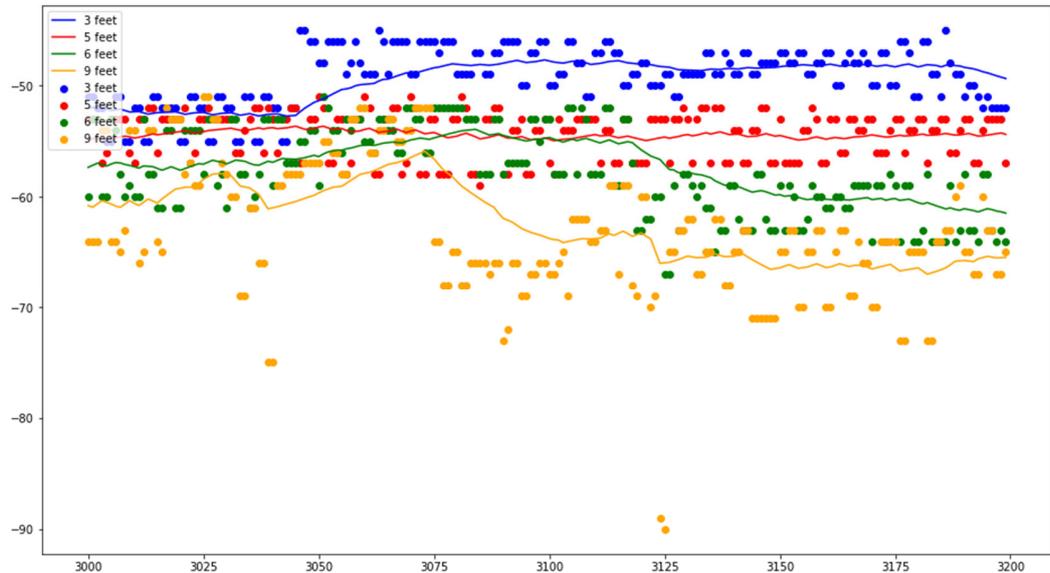


Figure 10 shows the Kalman Filter Visualized with scatter points representing the original RSS values while the line represents the filtered RSS.

Source: Student Generated

V. Normalization

- A. Normalization is necessary for a good machine learning model. For my model, I normalized the filtered RSSI using Z Normalization, also known as Standardization. This improves the numeral stability of my model and speeds up the training process. The formula for this normalization is shown below.

$$z = \frac{x - \mu}{\sigma}$$

$$\mu = \text{Mean}$$

$$\sigma = \text{Standard Deviation}$$

Figure 11 shows the formula for Z-Normalization

Source: <https://www.thoughtco.com/z-scores-worksheet-3126534>

VI. Label Encoding for strings

- A. The dataset from the MITRE Corporation involves many inputs in a string format, such as the types of devices: iPhone 11, iPhone 6, iPhone 7, etc. Machine learning cannot take in string/categorical data so it is necessary to convert them into numerical. The python library Sklearn makes this a very easy process with their label encoding which encodes categorical features as a one-hot numeric array. For example, the strings in the array ["paris", "tokyo", "amsterdam"] would be replaced with a number with "paris" = 0, "tokyo" = 1, and "amsterdam" = 2.

8.3 Model Construction and Training

In summary, thirty-three machine learning algorithms were constructed. The best and most unique models were Logistic Regression, Decision Trees, K-Nearest Neighbor, Random Forests, and Neural Networks (in Scikit learn and Keras/Tensorflow). All the machine learning algorithms were constructed with the robust python library, Scikit-learn.

Model	Accuracy	Balanced Accuracy	ROC AUC	F1 Score	Time Taken
RandomForestClassifier	0.81	0.81	0.81	0.81	2.26
BaggingClassifier	0.80	0.80	0.80	0.80	0.48
ExtraTreesClassifier	0.80	0.80	0.80	0.80	2.03
KNeighborsClassifier	0.80	0.80	0.80	0.80	1.95
DecisionTreeClassifier	0.80	0.80	0.80	0.80	0.13
LabelPropagation	0.80	0.80	0.80	0.80	33.63
LabelSpreading	0.80	0.80	0.80	0.80	43.33
ExtraTreeClassifier	0.80	0.80	0.80	0.80	0.09
XGBClassifier	0.78	0.78	0.78	0.78	1.09
LGBMClassifier	0.76	0.76	0.76	0.76	0.34
NuSVC	0.76	0.76	0.76	0.76	115.40
SVC	0.72	0.72	0.72	0.72	31.91
AdaBoostClassifier	0.69	0.69	0.69	0.69	0.84
QuadraticDiscriminantAnalysis	0.67	0.67	0.67	0.67	0.16
BernoulliNB	0.66	0.66	0.66	0.66	0.13
GaussianNB	0.66	0.66	0.66	0.66	0.08
SGDClassifier	0.65	0.65	0.65	0.65	0.20
NearestCentroid	0.65	0.65	0.65	0.65	0.13
LinearDiscriminantAnalysis	0.65	0.65	0.65	0.65	0.17
RidgeClassifier	0.65	0.65	0.65	0.65	0.13
RidgeClassifierCV	0.65	0.65	0.65	0.65	0.22
CalibratedClassifierCV	0.65	0.65	0.65	0.65	7.61
LogisticRegression	0.65	0.65	0.65	0.65	0.18
LinearSVC	0.65	0.65	0.65	0.65	2.03
Perceptron	0.63	0.62	0.62	0.62	0.15
PassiveAggressiveClassifier	0.61	0.61	0.61	0.61	0.14
DummyClassifier	0.50	0.50	0.50	0.50	0.07

Figure 12 shows a list of some models used and accuracies

Source: Student Generated

Additionally, this project utilized Keras and Tensorflow for the neural network approach. Keras is a deep learning library from TensorFlow, which is a neural network library that is written in Python.

The model construction starts with importing the complete csv file with all the values including the RSSI, filtered RSSI, ranges, and other factors obtained from the dataset from the MITRE Corporation.

Decision Trees are a non-parametric supervised learning method for classification and regression. Some benefits are that the cost of using the tree is logarithmic in the number of data points used to train the tree, and able to handle multi-output problems. However, decision trees can be unstable because small variations in the data might result in a completely different tree being generated.

Logistic regression is a linear model for classification rather than regression, despite its name. Logistic regression performs well when the dataset is linearly separable, less prone to overfitting, and easy/efficient to train and utilize. Additionally, this type of model not only gives a measure of

how relevant a predictor (coefficient size) is, but also its direction of association (positive or negative). Ultimately, this model was not the best for RSSI prediction since this model is based on the assumption of linearity between the dependent variable and the independent variables while in the real world data is rarely linearly separable.

K-Nearest Neighbors (KNN) provides functionality for unsupervised and supervised neighbors-based learning methods. This project uses supervised neighbors-based learning for classification for data with discrete labels. KNN constantly evolves as it is a memory-based approach and is simple to implement, but it is a slow algorithm.

Random Forests are an ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes or mean/average prediction of the individual trees. This type of model handles both categorical and continuous data well and are not sensitive to outliers. However, random forests are found to be biased while dealing with categorical variables and are slow to train.

Neural networks are computing systems vaguely inspired by the biological neural networks that constitute animal brains. This project used a Multi-layer Perceptron (MLP) which is a supervised learning algorithm that learns a function by training on a dataset. They are good to model with nonlinear data with large numbers of inputs and can be trained with any number of inputs and layers. However, Neural networks are black boxes, meaning the weights on the independent variables are unknown. This project constructed two Neural Networks using both Sklearn and Tensorflow.

8. Results & Discussion

Experimental and Analysis of data

Through experimentation, the data demonstrated that RSSI, in a real-world environment, does not easily correlate with distance. In addition, environments demonstrated to have a significant effect on the RSSI. However, summaries were difficult to conclude as it is hard to specifically determine how RSSI is impacted by the various environments. It can be hypothesized that the more objects in an environment blocking the two devices, the weaker the signal.

Models

There were 738670 train examples (data used to train the models) and 184668 test examples (data used to test the models). Three models for each algorithm were trained to compare a model that accounted for external factors versus a model that just accounted for the RSSI.



Figure 13a shows the accuracy for the models that accounted for external variables but without filtered RSSI. Figure 13b shows the accuracy of the models that just accounted for the filtered RSSI. Figure 13c shows the accuracy of the models that accounted for both Environmental Factors and filtered RSSI. Source: Student Generated

From the graphs above we can determine that accounting for external variables increases the accuracy by more than 30%. Decision trees had the best accuracy of 83%. Additionally, the Kalman filter was crucial to remove noise in RSSI which improved accuracy by more than 5%. Furthermore, the training accuracy and validation accuracy were both very similar which is a good sign because it means the models were not overfitted. Logistic regression showed the worst result and seems reasonable as this dataset was not linearly separable. However, it was an important model as it provides a reference, being the algorithm most closely related to the simple linear regression algorithm. Furthermore, it allows for analysis of the various weights/biases for the various inputs.

Table 1 shows the biases of the Logistic Regression model that accounted for the external factors. Source: Student Generated

Environment 1	Environment 2	Phone type1	Phone type2	Position 1	Position2	Pose1	Pose2	Angle	RSSI (filtered)
0.07027277	0.11012068	- 0.0340353 1	- 0.01421 856	0.19738 068	- 0.057550 04	0.172 15611	0.0741 2156	0.0072 3099	0.879592 76

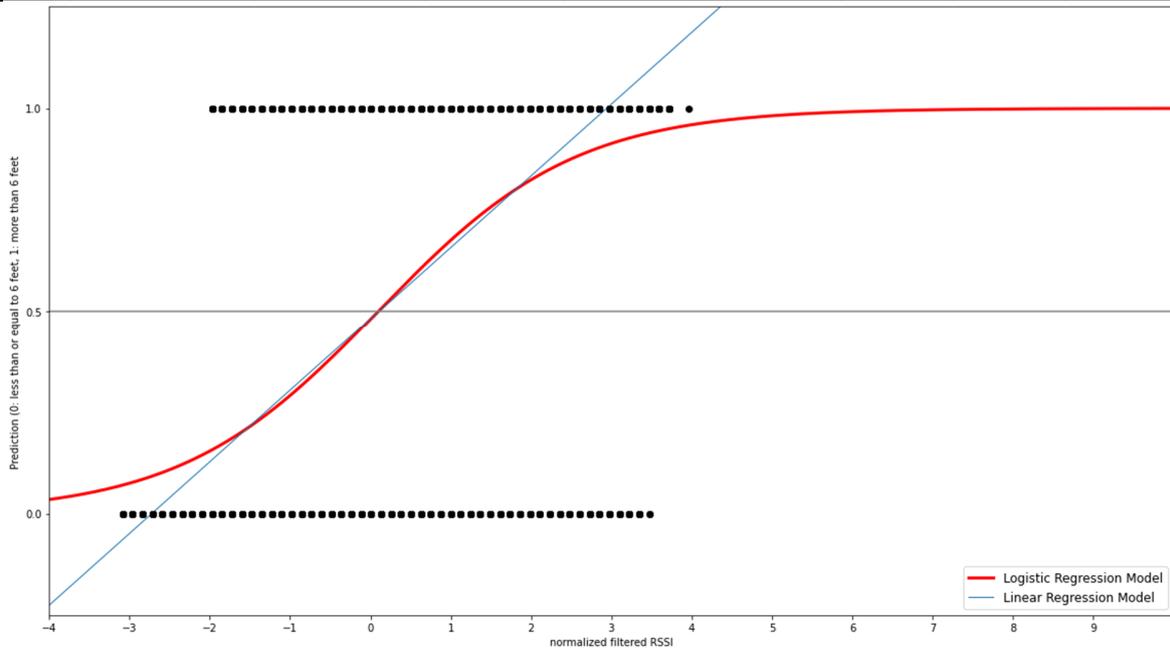


Figure 14 shows a visual of the Logistic Regression Model and a Linear Regression model.
Source: Student Generated

From table 1, we can see that the RSSI is the most crucial variable followed by the physical environments, suggesting that the most impactful factor was the physical environment. Second most important factor was the position of the phone (in a bag, pocket, etc.).

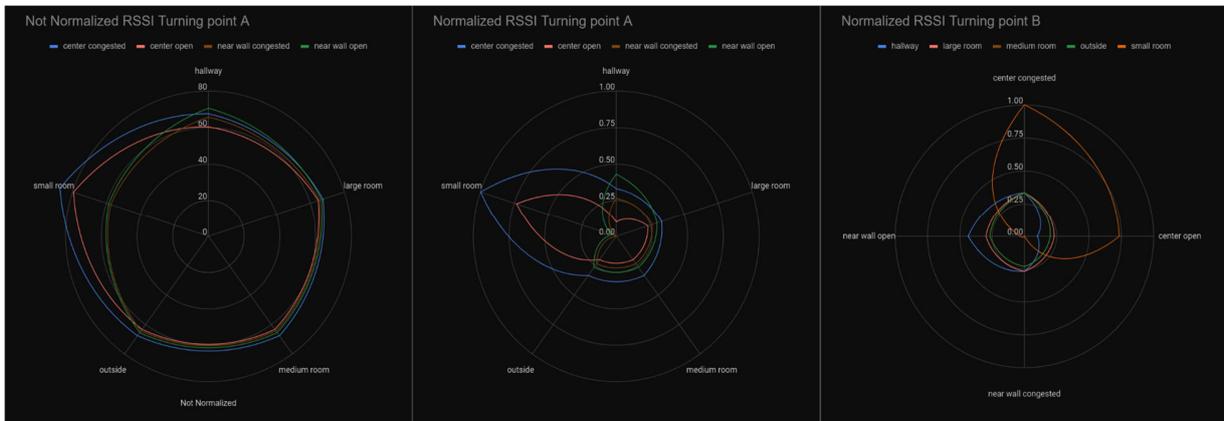


Figure 15a, 15b, & 15c demonstrates the effect of environments on the inflection point of RSSI values of where the range is determined to be “close contact” or not.

Source: Student Generated

To analyze the effect of the various environments on RSSI, the turning point was calculated for every combination of environments. A turning point is a tipping point for the RSSI when a value can be considered to be “close contact” or not (within predicted distance of 6feet or not). For example, the turning point for specifically an environment of a small near-wall congested room is -57.5 dbm. This means a stronger RSSI than -57.5 dbm would be predicted as close proximity and a weaker RSSI such as -100 dbm would be predicted as a far proximity. These turning points were then graphed on these radar charts. Figure 15a demonstrates the turning points with the strongest RSSI turning point of -57.5dBm and the weakest RSSI turning point of -87.5dbm.

To better view this data, the turning points were normalized to numbers between 0 and 1 to better demonstrate the relationships of the environments and RSSI. Figure 15b demonstrates the Normalized Inflection points of figure 15a and figure 15c reversed the X and Y values. While it is hard to determine a conclusive relationship between the physical environments and RSSI, we can determine that congested small rooms resulted in the weakest RSSI signals. Ultimately, this demonstrates that environmental factors play various roles in Bluetooth RSSI and need to be accounted for. The State Contact Tracing systems method of using a static and unchanging threshold is not accurate for proximity detection which is used by many State Contact Tracing systems against COVID-19.

Finally, a short real-life test was performed in various environments inside my house and outside with combinations of all other factors. However, the phone type factor was not used since the experiment was performed with two raspberry pi 4s. This short experiment’s purpose was to see the prediction in action in the real world environment. The results were 100% accurate as the changing environments were accounted for when calculating proximity. However, more testing will be needed for a more conclusive analysis of effectiveness.

10. Conclusion

Machine Learning allows computers to recognize patterns and inferences in large amounts of data and then utilize them to solve problems. This project proves that it is essential to include outside variables in predicting the proximity between two Bluetooth devices. Out of the six models trained, all of them except logistic regression showed potential. This project demonstrates that it is essential to include outside variables in predicting the proximity between two Bluetooth devices with the models trained without environmental factors having accuracies significantly lower at 66% as opposed to 83%. Thus, machine learning proves to be a successful approach in developing a private automated contact tracing system.

This study also applies to future widespread disease and pandemics. Bluetooth Contact Tracing, being privacy protecting, can attract more people willing to participate, making tracing more widespread and successful. With a successful and accurate automated contact tracing system on a global scale, quarantine may become only a thing of the past.

11. Future Research

There are many other approaches to improving the proximity detector's accuracy and reliability. Increasing the number of independent variables to account for all sorts of possible factors that influence RSSI would definitely improve the accuracy and reliability. Furthermore, other machine learning models can be tested such as Recurrent Neural Networks (using Long short-term memory cells). This approach would take multiple RSS values (in sequential order based on their time) as inputs for the neural network. Ultrasound is another technique that could be utilized to confirm close proximity. One device can emit a high-pitched sound for the other device to receive through its microphone. This would confirm if the two devices are in the same room or within range, allowing for a simple proximity checker.

Acknowledgements

I would like to thank Dr. Carol Hersh for providing me with insightful advice and comments on this paper.

References

- [1] Bay, J., Kek, J., Tan, A., Hau, C. S., Yongquan, L., Tan, J., & Quay, T. A. (2020). BlueTrace: A privacy-preserving protocol for community-driven contact tracing across borders. *Semantic Scholar*. <https://pdfs.semanticscholar.org/b460/fbaf2041cec41ee1266fcf1b60d3683d137b.pdf>
- [2] Bekkelien, A., Deriaz, M., Marchand-Maillet, S. (2012, March). Bluetooth Indoor Positioning. <http://cui.unige.ch/~deriazm/masters/bekkelien/Bekkelien.pdf>
- [3] Robinson, A., & Waldo, J. (2021). Technical Difficulties of Contact Tracing. Belfer Center for Science and International Affairs. <https://www.belfercenter.org/index.php/publication/technical-difficulties-contact-tracing>.
- [4] Bisong, E. (2019). TensorFlow 2.0 and Keras. *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, 347-399. doi:10.1007/978-1-4842-4470-8_30
- [5] Bonaccorso, G. (2017). *Machine learning algorithms*. Packt Publishing Ltd.
- [6] Burke RM, Midgley CM, Dratch A, et al. Active Monitoring of Persons Exposed to Patients with Confirmed COVID-19 — United States, January–February 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:245–246. doi: <http://dx.doi.org/10.15585/mmwr.mm6909e1>
- [7] Faragher, R., & Harle, R. (2015). Location Fingerprinting With Bluetooth Low Energy Beacons. *IEEE Journal on Selected Areas in Communications*, 33(11), 2418-2428. doi:10.1109/jsac.2015.2430281
- [8] Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., . . . Fraser, C. (2020). Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*, 368(6491). doi:10.1126/science.abb6936
- [9] Gomez, C., Oller, J., & Paradells, J. (2012). Overview and Evaluation of Bluetooth Low Energy: An Emerging Low-Power Wireless Technology. *Sensors*, 12(9), 11734–11753. doi:10.3390/s120911734
- [10] Hatke, G. F., Montanari, M., Appadwedula, S., Wentz, M., Meklenburg, J., Ivers, L., ... & Fiore, P. (2020). Using Bluetooth Low Energy (BLE) signal strength estimation to facilitate contact tracing for COVID-19. *arXiv preprint arXiv:2006.15711*.

- [11] Kramer, O. (2013). K-nearest neighbors. In Dimensionality reduction with unsupervised nearest neighbors (pp. 13-23). Springer, Berlin, Heidelberg.
- [12] Liu, Jia & Chen, Canfeng & Ma, Yan & Xu, Ying. (2013). Energy Analysis of Device Discovery for Bluetooth Low Energy. Vehicular Technology Conference, 1988, IEEE 38th. 1-5. doi:10.1109/VTCFall.2013.6692181
- [13] Naya, F., Noma, H., Ren Ohmura, & Kogure, K. (2005). Bluetooth-based Indoor Proximity Sensing for Nursing Context Awareness. Ninth IEEE International Symposium on Wearable Computers (ISWC'05). doi:10.1109/iswc.2005.13
- [14] Paul, A. S., & Wan, E. A. (2009). RSSI-Based Indoor Localization and Tracking Using Sigma-Point Kalman Smoothers. IEEE Journal of Selected Topics in Signal Processing, 3(5), 860–873. doi:10.1109/jstsp.2009.2032309
- [15] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. the Journal of Machine Learning research, 12, 2825-2830.
- [16] Rashidi, Hooman & Tran, Nam & Betts, Elham & Howell, Lydia & Green, Ralph. (2019). Artificial Intelligence and Machine Learning in Pathology: The Present Landscape of Supervised Methods. Academic Pathology. 6. 237428951987308. 10.1177/2374289519873088.
- [17] Rivest, S., Ivers, L., Soibelman, I., Weitzner, D., Zissman, M. (2020, May 19). PACT: Private Automated Contact Tracing. <https://pact.mit.edu/wp-content/uploads/2020/05/PACT-Mission-and-Approach-2020-05-19-.pdf>
- [18] Sthapit, P., Gang, H., & Pyun, J. (2018). Bluetooth Based Indoor Positioning Using Machine Learning Algorithms. 2018 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia). doi:10.1109/icce-asia.2018.8552138
- [19] Singhal, I., Tiwari, R. (2020). Social Distancing Detection using Bluetooth® Low Energy. STMicroelectronics. https://www.st.com/resource/en/application_note/dm00712091-social-distancing-detection-using-bluetooth-low-energy-stmicroelectronics.pdf
- [20] Welch, G., & Bishop, G. (1995). An introduction to the Kalman filter.

Individualized Emotion Recognition Through Dual- Representations and Group-Established Ground Truth

Valentina Zhang

Phillips Exeter Academy, Exeter NH, USA
vzhang1@exeter.edu

Abstract

While facial expression is a complex and individualized behavior, all facial emotion recognition (FER) systems known to us rely on a single facial representation and are trained on universal data. We conjecture that: (i) different facial representations can provide different, sometimes complementing views of emotions; (ii) when employed collectively in a discussion group setting, they enable more accurate emotion reading which is highly desirable in autism care and other applications context sensitive to errors. In this paper, we first study FER using pixel-based DL vs semantics-based DL in the context of deepfake videos. Our experiment indicates that while the semantics-trained model performs better with articulated facial feature changes, the pixel-trained model outperforms on subtle or rare facial expressions. Armed with these findings, we have constructed an adaptive FER system learning from both types of models for dyadic or small interacting groups and further leveraging the synthesized group emotions as the ground truth for individualized FER training. Using a collection of group conversation videos, we demonstrate that FER accuracy and personalization can benefit from such an approach.

KEYWORDS

neurodivergence care, facial emotion recognition, deep learning, ground truth for supervised learning

1. INTRODUCTION

In medical practice, emotion recognition is crucial to accurate clinical decision-making [1]. However, there are several obstacles. Patients' emotional behaviors could often time be affected by an underlying condition such as neurodivergence. On the other hand, physicians could be biased by their own emotions and limited by their cognitive ability.

Deep learning (DL) offers a unique value in this context as a DL-based system does not have an emotional bias and could detect patterns and changes too subtle for human cognition at real-time. Yet, there are many challenges for building an accurate DL-based facial emotion recognition (FER) system. Leveraging a small conversing group setting in autism care, our research investigates the potentials of composing DL-based FER system with dual-representations and automatically deriving the ground truth for quality training data.

This paper hypothesizes that for comparable accuracy rates, pixel-based DL and semantics-based DL sometimes deliver complementing predictions in FER. In the context of small conversing groups, the two types of DL models could be orchestrated to deliver more accurate group and individual emotion recognition. Our adaptive group emotion recognition system includes three components: individual emotion recognition, adaptive group emotion synthesis, and group vs individual emotion modeling. We aim to understand the relationship between pixel-based DL and semantics-based DL and how they could work together to potentially outperform humans in FER.

The main contributions of this paper are (i) a comparatively study of DL models trained with different facial representations; (ii) an adaptive approach toward accurate individual and group FER leveraging discussion group context; and (iii) a proposal to use group emotion as ground truth labels for FER personalization.

The remainder of the paper is organized as follows: In section II, we outline the two neural networks we used for this work and how their different performances inspired us. Section III describes our system architecture, face detection mechanism and working model for adaptive group emotion recognition in detail. Recognizing the established group emotion as a robust ground truth, section IV outlines how it could be leveraged to improve emotion recognition for individuals and automatic training data labelling. The section then analyzes the test results of our experiment. Section V discusses the threats to the validity of our research. In section VI, a few closely related works are presented. Section VII gives the conclusion and future work of the paper.

2. ANALYSIS OF TWO TYPES OF FER MODELS

When training a DL model for image analysis tasks, there are two general approaches. One is to use the full images as training input data, which is called in this paper as pixel-based DL or image-based training. The other approach is semantics-based, which extracts semantics from the images and uses extracted features such as facial landmarks as the training data. We choose to comparatively study these two different approaches as the human emotion recognition system operates similarly.

In the human brain, a section called the fusiform face area looks at the whole face holistically, which is similar to how a model trained on full images would function. On the other hand, a part of the human brain called the occipital face area recognizes the eyes, nose, and mouth as individual pieces, which would be very similar to a model trained on facial landmarks. The landmark-based training is generally considered more effective because it helps the neural networks to focus on the essence of the problem, which is the outline of a face. In facial emotion recognition where the outlines to emotion mapping may not be well defined [2], we assume that image-based training may outperform landmark-based training in some cases. This section explains our experiment for validating our assumption.

Our first model is a standard fully-convolutional neural network composed of 60 convolutional and separable convolutional layers, ReLUs, Batch Normalization, Dropout, Flatten, and Global Average Pooling layers. It is trained with the ADAM optimizer, and achieved a validation accuracy of 60% on the FER-2013[3] dataset.[4]

Our second model is trained on facial landmarks which are extracted by solving the shape prediction problem. In this approach, each face consists of a few shapes which outline the face, eyes, mouth, and nose. Through extracting these shapes, this model will ignore the other features and details of the face, which may have introduced noise into the data of pixel-based training. We use a standard 68-point facial landmark system. There are 6 landmarks for each eye, 9 for the nose, 20 for the lips, and the remaining 27 outline the face.

Landmark-based training uses information about these 68 landmarks on the face to correlate the shape of these landmarks with a certain emotion. Four pieces of information from each landmark is extracted: the x and y coordinates, its distance from the mean of all the points, and its vector angle. Collectively, this representation of the 68 landmarks gives a comprehensive summary of facial features. Our second model is a simple neural network consisting of two hidden layers of 128 neurons using the Rectified Linear Unit activation function, and Adamax optimizer. This produces a validation accuracy of 58%.

For our experiment, we analyzed a well-known deepfake video of Robert Downey Jr. created from a speech by Elon Musk using DeepFaceLab 2.0 Quick96 at 1 million iterations. We apply our two trained models to every frame of the video and produce a probability value for each of our seven core emotions. Subsequently, we correlated the results from both models and constructed correlation heatmaps as shown in Fig. 1.

The correlation analysis reveals some interesting results. On the two heatmaps, the correlation between the original and deepfake videos’ corresponding emotions is shown as a diagonal orange line of square cells. Our data shows that the landmark-based training model detected a much higher correlation between the original and the deepfake, even in the “Disgust” and “Surprise” components, which are very minimal in the videos.

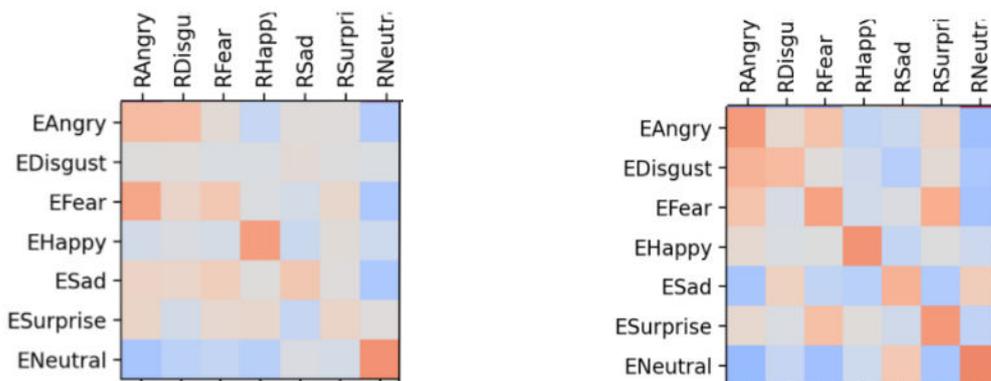


Fig. 1. Emotion Correlation by Image-Trained (left) vs. Landmark-Trained Model (right)

We believe that this is due to the fact that subtle facial expressions are difficult to accurately capture through facial landmarks. On the other hand, the image-based training model detected details that were lost during the creation of the deepfake. To confirm this belief, we employed three human evaluators of the video. Given the manually labeled video from each of the three evaluators, we computed the two-judges agreement to obtain the true labels (e.g., a label was marked as a true positive if at least two of the three evaluators classified it as such). As our assumption predicted, the emotion ground truth on the two faces, as shown in Fig. 2, do not match well. Generally, Elon Musk appeared to have many more positive emotions than the deepfake did.

Our manual analysis also confirmed some subtle or mixed emotions lost in the deepfake video.

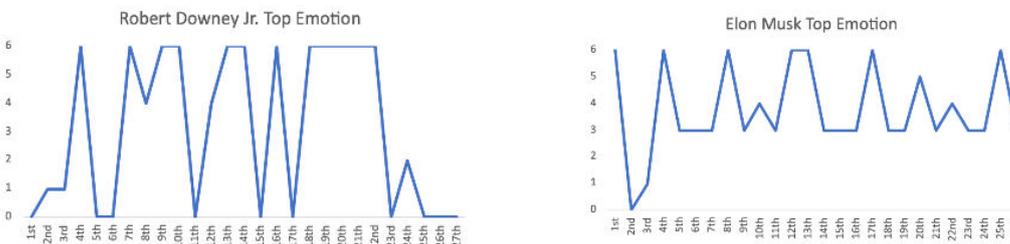


Fig. 2. Ground Truth for Deepfake (left) vs. Original (right)

To analyze our assumption that our two models have complementing strengths in detecting facial emotions, we constructed the complementarity metrics for the recognized emotions as shown in TABLE I.

TABEL I. DL Models Complementary metrics

Deepfake	Landmark Correct	Landmark Incorrect	Exclusive	%
Image Correct	12	20	Landmark	42%
Image Incorrect	16	4	Image	33%

Original	Landmark Correct	Landmark Incorrect	Exclusive	%
Image Correct	8	10	Landmark	42%
Image Incorrect	6	28	Image	25%

With the two tables on the left, we report the intersection of sets of true positive emotion detected by two models. For example, 23% of the true positives in the original video and 15% of the true positives in the deepfake are detected by both the image-trained model and landmark-trained model. The relatively small overlap (no more than 51%) suggests that these representations complement each other.

The two tables on the right shows the difference in the sets of true positives detected by the two representations. For example, 42% of the true positives detected by the landmark-trained model were not correctly identified by the image-trained model for both the original and the deepfake.

In summary, while the image-trained model picks up more textual details and recognizes more subtle emotions, the landmark-trained model detects less noise and has better accuracy in detecting well-articulated emotions. Thus, there is a potential to create a more effective emotion recognition system by combining the two models.

3. ADAPTIVE FER LEVERAGING GROUP EMOTION CHANGE CADENCE

With its wider adoption, DL-based FER today is used in technologies interacting with humans where accurate detection of individual and group emotion becomes desirable or even necessary [5]. Our motivating interest in better autism care is one example. Group emotion is a complex function of group members' emotions, group context, and environmental context. While the context information is relatively easy to acquire, people reflect their emotions facially in different ways and varying degrees of intensity. It is a challenge to design a system that accounts for individual behaviors.

Taking on the challenge, we construct a system combining the strength of both models from the previous section. Since group emotion is defined by the cadence of individual emotion changes in an engaging environment, our basic idea is to use the cadence to find an optimal trade-off between the two models.

We chose small (4 people) conversing groups as the experiment context of our system for four reasons: First, compared to static images, videos are more redundant for robust emotion recognition. Second, emotional changes within a conversation group tend to be synchronous which simplifies our system design. Third, small groups are easier for manual emotion evaluation and annotation. Lastly, we found a good amount of 2x2 grid view conversation videos for analysis.

The architecture of our system is shown in Fig. 3. There are three important components: noise reduction in emotional change recognition, group emotion synthesis and change alignment, and adaptive weights of two models' outputs.

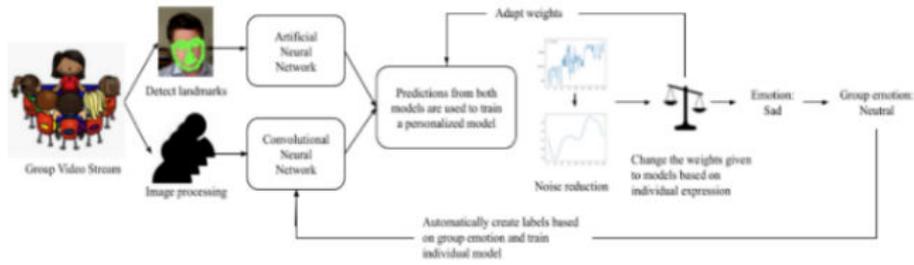


Fig. 3. Architecture of the Adaptive FER System

3.1 Noise Reduction in Emotional Change Recognition

One important limitation observed of the DL-based FER models is that they are sensitive to noisy inputs. For example, image quality jitters from frame to frame in the video stream could cause brief false emotion readings. To combat this, we computed the weighted arithmetic mean of detected emotions for every frame in a second. This averaging allows us to ensure that correlation between changes in emotion are in fact caused by participant’s emotions, and reduces the impacts from various input noises.

3.2 Group Emotion Synthesis and Change Alignment

A group’s emotion should reflect all its members’ emotions. It is more diverse than individual emotions and could consist of more than one dominating emotion reflecting a polarizing group sentiment. To account for that, for any given point in time, our system sums up the probabilities of each emotion type from all group members and uses Euclidean distance to check for deviation. This is adequate for the small group size we analyze. For larger groups, a clustering algorithm such as K-means could be used to separate polarized sections.

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2 + \dots + (p_m - q_m)^2}$$

Fig. 4. Euclidean Distance

In detecting emotion changes, we separate the group-wise events where most members change emotions from the individual events where only one member changes. This allows us to treat group-level noises differently from individual-level noises. For group-level noise, we use DFT to convert the time series into frequency-domain and the noise reduction is achieved by eliminating the low-energy frequencies.

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N} \quad k = 0, \dots, N - 1,$$

Fig. 5. Discrete Fourier Transform

For individual-level noises, we want to handle them more carefully because we do not want to miss any real emotion readings obscured by personalized behaviors. We adopted a collection of heuristics such as checking for recurring patterns before dismissing it as a noise.

3.3 Adaptive Weights of Two Models’ Output

By our observations, the most common inconsistency in recognized emotions is the intensity in which people express their emotions. One's faint twitch of the cheek may convey the same amount of happiness as another's wide grin. As we demonstrated in earlier discussion, the image-based model is better equipped for recognizing these subtleties which the landmark abstraction could overlook. Thus, if not enough changes in emotions were detected in a certain group member, we increase the weight of the image-trained model to account for more subtle changes. On the other hand, if the emotion readings are noisier than the group average, the group member's facial expressions may be exaggerated, or the image quality may be low. In this case, we amplify the landmark-trained model to focus on the key emotions.

In Fig. 6, we describe our adaptive algorithm. For example, as a baseline, both models are given the same weight. Every minute, we track the amounts of changes in emotion during that minute. Out of the four participants, the ones that were detected to have above-average amounts of changes in emotion were given an extra weight to the landmark-based model and the same weight change is reduced from the image-based model. The opposite was done to the participants who were detected to have below-average amount of changes in the same time frame. Overtime, the weight for each participant settles into equilibrium as the group emotion dynamics converges. To facilitate the convergence, the weight adjustment value is a function of emotion change amount distribution among the group members. The higher the deviation, the higher the adjustment value.

```

Initialize weights for image-based model to 0.5
Initialize weights for landmark based model to 0.5
for every 10 seconds
    for every participant in meeting
        if the amount of emotion changes of member >
            group number of emotion changes
                Increase weight for landmark based model
                Decrease weight for image-based model
            else
                Decrease weight for landmark based model
                Increase weight for image-based model
    Emotion of each participant = image model prediction * image model weight +
        landmark model prediction * landmark model weight
    Group emotion = average of individual emotions

```

Fig. 6. Pseudo-code for the Adaptive Algorithm

To illustrate the results of our system, the graph to the left in Fig. 7 shows that the adaptive algorithm starts taking corrective action after 20 seconds and has eliminated three transient noises the non-adaptive algorithm classified as "Sad". The graph to the right in Fig. 7 shows that the adaptive algorithm corrected the false "fear" while bringing to surface a subtle angry emotion that would have been missed otherwise.

An issue our adaptive system does not handle well is when an individual's facial structure or neurodivergence makes the person show unintended emotions. In some recordings we reviewed as a part of this work, there were participants consistently misclassified as having a sad or angry emotion component. A more advanced online learning algorithm could aim to detect such persistent patterns in people's emotions and systematically remove the misleading structural component. Alternatively, considering our discussion group context, one could automatically generate personalized DL training data using the group emotion as the ground truth. The next section explores the latter as a solution.

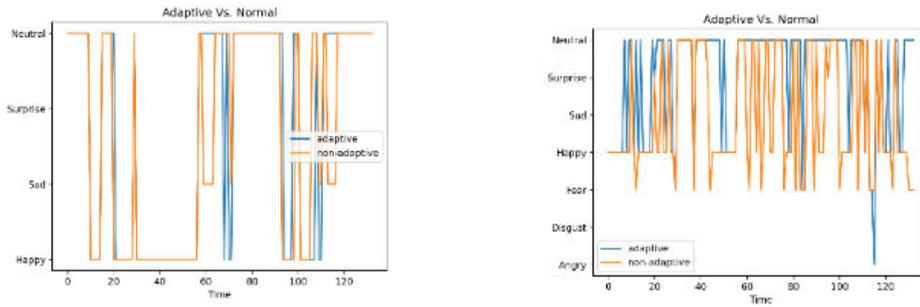


Fig. 7. Emotions Recognized by Adaptive vs Non-adaptive Algorithm

4. GROUP EMOTION AS THE GROUND TRUTH

Group emotion recognition has applications in social psychology [6], shot selection [7], image retrieval [8], surveillance [9], event detection [10], and event summarization [7]. We propose that group emotion be also used for personalizing individual emotion recognition.

For our group emotion recognition purpose, we classify discussion groups into four types: (i) unengaged groups, (ii) engaged groups, (iii) synchronous groups, and (iv) homogenous groups. An engaged group can establish more information about the emotion correlation among group members than an unengaged group. A synchronous group is a strongly engaged group whose members share the same emotion change cadence, but the specific emotions at a given time may not be the same. At the highest level of engagement, in a homogenous group, group members tend to share the same emotions at any given time.

In the previous section, we discussed how group emotion cadence assists our adaptive algorithm to recognize individual emotion more accurately. It assumes that emotions of the members of the group are all affected by group-wise events, which would indicate that it is a synchronous group. This assumption required us to identify and eliminate the unengaged members from the group emotion calculation.

If we further restrict our application context to a homogenous group, such as in classrooms, or movie theatres, or concert halls, we hypothesize that the group emotion could be treated as the ground truth and used to label new training data. For example, when a group member's images are labelled with this ground truth, DL model could be trained against the member's personal facial expression patterns. In this section, we discuss our experiment designed to validate this idea.

Our experiment consists of the following four steps:

Step 1. We identified group discussion videos of knowledge-sharing nature where group members' emotions are highly synchronized without any divergence of opinions.

Step 2. Using the first half of the video, we compute the group emotion G_x using 3 of the 4 people in the group and label all image frames of the fourth person using G_x as the ground truth.

Step 3. Train our emotion recognition model with the labelled images obtained in Step 2.

Step 4. Using the second half of the video, we apply the newly trained model to the fourth person and check its performance against the models described earlier in this paper.

In principle, Step 3 could be achieved through active learning techniques so that the training could happen online in real-time. However, for this work, we have not tried that because our main goal is to show the validity of using group emotion as the ground truth and the viability of such an approach.

For the validation in Step 4, while a quantitative and general analysis is very difficult, we consider as a qualitative indicator whether automatic labeling increases the correlation between the individual and group emotion. Under our assumption that our group discussion videos do not produce diverging emotions and the group members are evenly engaged throughout the videos, more correlation between the individual and the group would indicate that automatic labeling improves the accuracy of the model. Our correlation metrics are calculated using the Pearson algorithm. For the videos we experimented with, we observed an anecdotal correlation improved of 15% - 20%. While this is encouraging, we believe more data is needed for a thorough quantitative analysis.

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}$$

Fig. 8. Pearson Correlation Coefficient

As a specific example, Fig. 9 compares two time series obtained in our experiment. The blue line is generated by the adaptive FER system described in the previous section. The orange line is generated by the model trained with the automatically labeled data. The main difference between the two is that the blue line was trained with universal data (FER-2013) and the orange line was trained with personal data. Rather than relying on a universal facial emotion model, the orange line is able to identify the inherent “Sad” component in this particular group member’s facial expression. Additionally, the false readings of “Fear” and “Angry” due to unrelated facial changes were also detected and compensated.

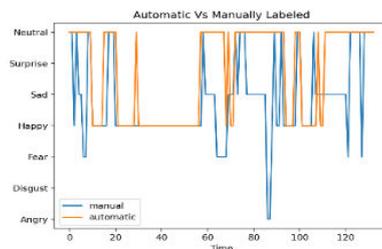


Fig. 9. Model Trained with Auto Labels vs Manual Labels

One important prerequisite for this approach is that the group videos used to collect training data need to exercise the full range of emotion states. This allows the personalized DL model to be trained with labels of all possible emotion values.

Labeling training data is an expensive, error-prone, yet critical step in developing DL-based systems. Our experiment shows an example of how this step could be automated to some extent through the knowledge of group context. Another interesting observation made in our experiment is that data labelled this way captures individual facial emotion patterns which could lead to more accurate and personalized emotion detection. We believe this finding could be applied to other group behavior analysis contexts where DL could play a bigger role in both synthesizing group-level behavior properties and creating individualized DL models.

5. THREATS TO VALIDITY

Construct validity. The main threat is related to how we assess the complementarity of the facial representations: image vs landmark. We support this claim by performing two different analyses: (i) complementarity metrics; and (ii) correlation test.

Internal validity. This is related to possible subjectiveness when evaluating the group emotion in the video fragments used. To mitigate such a threat, we employed three evaluators who independently checked the emotion. Then, we computed two-judges agreement on the evaluated videos. We also qualitatively discuss false positives and borderline cases.

External validity. The results obtained in our study used a small number of selected videos that may not generalize to other small conversing group contexts. To mitigate this threat, we applied our approach to a collection of group discussion videos of different subject areas, such as art, technology, politics, entertainment and sports. Another threat in this category is related to the fact that we apply our approach on pre-recorded videos only. While we do not yet have data to show the effectiveness of our approach in a live group meeting context, the focus of this paper is to show a general technique rather than to build a tool.

6. RELATED WORKS

In the area of individual FER, this work benefitted from studying the pioneering working such as [3][4][14] and holistically understanding their principles and limitations. In the area of group FER, this work has been informed by a diverse set of existing research settings ranging from a four-person UNO game [10] to public crowds [12]. In the area of comparative research in different facial representations and different DL architectures, this work drew its inspirations from [13]. Last but not least, this work relies on [15][16] for the complete and up-to-date survey of all published research works in FER and group FER. With deep appreciation, in this section, we review these representative papers that are most related to and influenced our work.

Alex Krizhevsky et al [14] provides a first detailed description and analysis of architecture and design variables of DL-based image classification. Ian Goodfellow et al [2] has an early yet insightful discussion on the challenges of facial emotion recognition and outlines some important considerations in designing a successful solution. Octavio Arriaga et al [4] explains one of the first and very successful CNN-based individual FER system. The system does not rely on facial landmarks. Tatsuya Hayamizu et al. [10] is one of the earliest works in group emotion recognition. It also studies a 4-person group, but it relies on classic statistics-based AI techniques instead of DL. Liwei Wang et al. [13], presents a first systematic approach quantitatively characterizing what representations do deep neural networks, and how similar are the representations learned by two networks with identical architecture but trained from different initializations.

7. CONCLUSIONS

In this paper, we show that accurate emotion recognition can be informed by different facial representations. We evaluated the performance of two dominant facial representations and showed their complementary values. Our adaptive group emotion recognition system is flexible and could be reused for different group sizes and contexts. This avoids retraining which eliminates a large time sink native to some DL approaches, and broadens the applicability of our approach. Moreover, as an on-going effort, the adaptive algorithm used by our system is being replaced with an adaptive machine learning (ML) model. Further analysis is being done to assess the relative effectiveness of this ML model against our hand-crafted adaptive algorithm. We hypothesize that the two also present complementary values to some degree, and their complementarity metrics should be studied.

Our approach also highlights the values of accurate group emotion analysis. We showed that by establishing the recognized group emotions as the ground truth, individual emotion patterns such as resulting from neurodivergence could be better analyzed and modeled through automatic training data labeling. This finding speaks to the general possibility of automating the creation of certain training data in various group meeting contexts.

ACKNOWLEDGEMENTS

The author wants to express her heartfelt appreciation to Dr. Neha Keshav and Dr. Ned Sahin at Brain Power LLC for their invaluable guidance throughout this research. The work also received generous support from MIT BeaverWorks Summer Institute. Thank you sincerely.

REFERENCES

- [1] Nancy R. Angoff, Making a Place for Emotions in Medicine, 2 Yale Journal of Health Policy L. & Ethics (2002).
- [2] Tuan Le Mau, etc. Professional actors demonstrate variability, not stereotypical expressions, when portraying emotional states in photographs, Nature Communications 19 August 2021.
- [3] Ian J. Goodfellow, et al. “Challenges in Representation Learning: A report on three machine learning contests.” <http://deeplearning.net/icml2013-workshop-competition>, July 2013
- [4] Octavio Arriaga et al., Real-time Convolutional Neural Networks for Emotion and Gender Classification, 2016.
- [5] J. Bullington. Affective computing and emotion recognition systems: the future of biometric surveillance. Proceedings of the 2nd annual conference on Information security curriculum development. ACM, 95–99. 2015
- [6] P.M. Niedenthal and M. Brauer. 2012. Social functionality of human emotion. Annual review of psychology 63 (2012), 259–285.
- [7] A. Dhall et al., 2015. Automatic group happiness intensity analysis. IEEE Transactions on Affective Computing 6, 1 (2015), 13–26.
- [8] A. Dhall, A. Asthana, and R. Goecke. 2010. Facial expression based automatic album creation. International Conference on Neural Information Processing. Springer, 485–492.
- [9] T. Vandal, D. McDuff, and R. El Kaliouby. 2015. Event detection: Ultra large-scale clustering of facial expressions. In IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Vol. 1. IEEE, 1–8.
- [10] Tatsuya Hayamizu, Group Emotion Estimation using Bayesian Network based on Facial Expression and Prosodic Information, 2012 IEEE International Conference on Control System, Computing and Engineering
- [11] James Bergstra and David D. Cox. Hyperparameter optimization and boosting for classifying facial expressions: How good can a “null” model be? Workshop on Challenges in Representation Learning, ICML, 2013.
- [12] V. Franzoni, G. Biondi, and A. Milani, “Crowd emotional sounds: spectrogram-based analysis using convolutional neural network.” in SAT@ SMC, 2019, pp. 32–36.
- [13] Liwei Wang, et al. Towards Understanding Learning Representations: To What Extent Do Different Neural Networks Learn the Same Representation. 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada.
- [14] Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pp. 1097–1105, 2012.

- [15] Emmeke A. Veltmeijer et al., Automatic emotion recognition for groups: a review. DOI 10.1109/TAFFC.2021.3065726, IEEE Transactions on Affective Computing
- [16] Wafa Mellouk et al., Facial emotion recognition using deep learning: review and insights. The 2nd International Workshop on the Future of Internet of Everything (FioE) August 9-12, 2020, Leuven, Belgium

Utilizing Artificial Intelligence to Predict Post Operative Atrial Fibrillation in Non-Cardiac Transplant

Alexander Heckman, Rohan Goswami, Zachi Attia, Paul Friedman, Peter Noseworthy, Demilade Adedinsewo, Pablo Moreno-Franco, Rickey Carter, Tathagat Narula

Abstract— Background: Postoperative atrial fibrillation (POAF) is associated with adverse health consequences, higher costs, and longer hospital stays. Utilizing existing predictive models that rely on clinical variables and circulating biomarkers, multiple societies have published recommendations on the treatment and prevention of POAF. Although reasonably practical, there is room for improvement and automation to help individualize treatment strategies and reduce associated complications. Methods and results: In this retrospective cohort study of solid organ transplant recipients, we evaluated the diagnostic utility of a previously developed AI-based ECG prediction for silent AF on the development of POAF within 30 days of transplant. A total of 2261 non-cardiac transplant patients without a preexisting diagnosis of AF were found to have a 5.8% (133/2261) incidence of POAF. While there were no apparent sex differences in POAF incidence (5.8% males vs. 6.0% females, $p=.80$), there were differences by race and ethnicity ($p<0.001$ and 0.035 , respectively). The incidence in white transplanted patients was 7.2% (117/1628), whereas the incidence in black patients was 1.4% (6/430). Lung transplant recipients had the highest incidence of postoperative AF (17.4%, 37/213), followed by liver (5.6%, 56/1002) and kidney (3.6%, 32/895) recipients. The AUROC in the sample was 0.62 (95% CI: 0.58-0.67). The relatively low discrimination may result from undiagnosed AF in the sample. In particular, 1,177 patients had at least 1 AI-ECG screen for AF pre-transplant above .10, a value slightly higher than the published threshold of 0.08. The incidence of POAF in the 1104 patients without an elevated prediction pre-transplant was lower (3.7% vs. 8.0%; $p<0.001$). While this supported the hypothesis that potentially undiagnosed AF may have contributed to the diagnosis of POAF, the utility of the existing AI-ECG screening algorithm remained modest. When the prediction for POAF was made using the first postoperative ECG in the sample without an elevated screen pre-transplant ($n=1084$ on account of $n=20$ missing postoperative ECG), the AUROC was 0.66 (95% CI: 0.57-0.75). While this discrimination is relatively low, at a threshold of 0.08, the AI-ECG algorithm had a 98% (95% CI: 97 – 99%) negative predictive value at a sensitivity of 66% (95% CI: 49-80%). Conclusions: This study's principal finding is that the incidence of POAF is rare, and a considerable fraction of the POAF cases may be latent and undiagnosed. The high negative predictive value of AI-ECG screening suggests utility for prioritizing monitoring and evaluation of the transplant patients with a positive AI-ECG screening. Further development and refinement of a post-transplant-specific algorithm may be warranted further to enhance the diagnostic yield of the ECG-based screening.

Keywords— artificial intelligence, atrial fibrillation, cardiology, transplant, medicine, ECG, machine learning.

Alexander Heckman is with the Mayo Clinic, United States (e-mail: heckman.alexander@mayo.edu).

Developing a Leukemia Diagnostic System Based on Hybrid Deep Learning Architectures in Actual Clinical Environments

Skyler Kim

Abstract— An early diagnosis of leukemia has always been a challenge to doctors and hematologists. On a worldwide basis, it was reported that there were approximately 350,000 new cases in 2012, and diagnosing leukemia was time-consuming and inefficient because of an endemic shortage of flow cytometry equipment in current clinical practice. As the number of medical diagnosis tools increased and a large volume of high-quality data was produced, there was an urgent need for more advanced data analysis methods. One of these methods was the AI approach. This approach has become a major trend in recent years, and several research groups have been working on developing these diagnostic models. However, designing and implementing a leukemia diagnostic system in real clinical environments based on a deep learning approach with larger sets remains complex. Leukemia is a major hematological malignancy that results in mortality and morbidity throughout different ages. We decided to select acute lymphocytic leukemia to develop our diagnostic system since acute lymphocytic leukemia is the most common type of leukemia, accounting for 74% of all children diagnosed with leukemia. The results from this development work can be applied to all other types of leukemia. To develop our model, the Kaggle dataset was used, which consists of 15135 total images, 8491 of these are images of abnormal cells, and 5398 images are normal. In this paper, we design and implement a leukemia diagnostic system in a real clinical environment based on deep learning approaches with larger sets. The proposed diagnostic system has the function of detecting and classifying leukemia. Different from other AI approaches, we explore hybrid architectures to improve the current performance. First, we developed two independent convolutional neural network models: VGG19 and ResNet50. Then, using both VGG19 and ResNet50, we developed a hybrid deep learning architecture employing transfer learning techniques to extract features from each input image. In our approach, fusing the features from specific abstraction layers can be deemed as auxiliary features and lead to further improvement of the classification accuracy. In this approach, features extracted from the lower levels are combined into higher dimension feature maps to help improve the discriminative capability of intermediate features and also overcome the problem of network gradient vanishing or exploding. By comparing VGG19 and ResNet50 and the proposed hybrid model, we concluded that the hybrid model had a significant advantage in accuracy. The detailed results of each model's performance and their pros and cons will be presented in the conference.

Keywords— acute lymphoblastic leukemia, hybrid model, leukemia diagnostic system, machine learning.

Integrated Machine Learning Framework for At-Home Patients Personalized Risk Prediction using Activities, Biometric and Demographic Features

Claire Xu, Welton Wang, Manasvi Pinnaka, Anqi Pan, Michael Han

Abstract—Hospitalizations At-home patients are lack of continuous monitoring after being discharged from hospital. Especially the heart failure patients are at high risk for hospital readmissions and mortality. Early risk detection and intervention can reduce this high cost and increase the satisfaction of both patients and physicians. Due to the lack of awareness of the potential arising risks in home environment, the opportunities for patients to seek early actions of clinical visits are dramatically reduced. Our aim is to offer a highly personalized remote patients monitoring and risk assessment Machine Learning framework to identify the potentially preventable hospitalization for both acute as well as chronic diseases. Data to train the Machine Learning integrated framework is gathered from both clinical setting, patients' surveys, as well as online databases. 20+ features are analyzed ranging from activities, biometric info, demographic info, socio-economic info, hospitalization history, medication info, lifestyle info, etc. Experimented models cover Logistic Regression, XGBoost, SVM, and Deep Learning Neural Network. An integrated ML-framework is trained for risk prediction and yields high performance of 87% accuracy and 88% sensitivity. This integrated framework is proven to be effective in identifying the potentially preventable hospitalization. Further, the high indicative features are identified by the models which guide us to a healthy lifestyle and early intervention suggestions.

Keywords—Hospitalization Prevention, Machine Learning, Remote Patient Monitoring, Risk Prediction

Claire Xu is with the Harker School, 4525 Union Ave, San Jose, CA 95124 USA (e-mail: 27clairex@students.harker.org).

Welton Wang is with Archbishop Mitty High School, 5000 Mitty Way, San Jose, CA 95129 USA (e-mail: weltonwang88@gmail.com).

Manasvi Pinnaka is with Basis Independent Silicon Valley, 1290 Parkmoor Ave, San Jose, CA 95126 USA (e-mail: maavistar@gmail.com).

Anqi Pan is with Salesianum School, 1801 N Broom St. Wilmington, DE 19802 USA (e-mail: anqi.pan@salesianum.org).

Michael Han is with William Diamond School, 99 Hancock St, Lexington, MA 02420 USA (e-mail: michael.l.han@gmail.com).

Establishing a Data Processing Pipeline Using Artificial Intelligence to Capture Human Kinematic Data: DeepLabCut as a Markerless Motion Capture System

Fatemeh Khounsarian¹, Sowmya Gopalakrishnan², Devyani McLaren³, Sara Kraeutner⁴, Lara Boyd⁵

Abstract—Investigating the neural underpinnings of motor learning is crucial to understanding human movement. However, brain data is non-interpretable without accurate behavioural information to contextualize it. Thus, to understand the brain, it is vital to gather kinematic data to characterize the complexity of movements and changes in movement behaviours. Kinematic variables are traditionally mapped with motion capture systems. Current research-grade motion capture systems are cumbersome and expensive, requiring markers that are affixed to the body to track movements. Further, these systems often lack the sensitivity to capture small movements of the hand or individual fingers and are labor intensive, requiring extensive post-processing. DeepLabCut is an open-source 3D markerless motion capture system that uses deep learning and transfers learning from neural networks for movement tracking. This new technology offers the possibility of gathering accurate kinematic movement data through inexpensive means with less effort as compared to traditional marker based motion capture systems. The purpose of this research is to investigate the feasibility of DeepLabCut for use in studies of human motor learning. Therefore, we assess DeepLabCut's ability to discriminate and generalize amongst a variety of movements in a healthy population. To do so, we initially created a project pipeline for future research studies to use DeepLabCut to analyze the movement data collected. This involved installing DeepLabCut on personal and lab computers, optimising the workflow to train the AI network initially, and recruiting the use of Sockeye, UBC's supercomputer, to help with the speed of analysis. We investigated the number of frames that had to be labeled by a human to train the AI network, and determined best practices for using sockeye to analyze DeepLabCut. Then, we tested the efficacy of DeepLabCut by training the AI network on one participant's movement video and testing the network on new, unseen participant's videos. We determined the feasibility of the machine learning software DeepLabCut and found that it was able to accurately analyze movement patterns and dramatically decreased the amount of human intervention to extract kinematic data from human movements. Furthermore, it decreased the amount of total time required to analyze a data set and output results in a way that was easily exportable to other formats. The software and process was also deemed to be user-friendly, and thus a feasible alternative to existing motion capture systems. Overall, DeepLabCut has removed the need for conventional marker-based motion capture systems while assessing physical movement accurately. Therefore, in the future, it could be used to remotely conduct physical assessments from the comfort of a patient's home with high quality and accuracy of the data obtained. This will benefit the large population of patients with movement disorders including neurological, psychiatric, and rehab patients.

Keywords—AI, Artificial Intelligence, DeepLabCut, Neuroscience,

Fatemeh Khounsarian is with the The University of British Columbia, Canada (e-mail: f.khonsar@yahoo.com).

Introduction—Investigating the neural underpinnings of motor learning is crucial to understanding human movement. However, brain data is non-interpretable without accurate behavioural information to contextualize it. Thus, to understand the brain, it is vital to gather kinematic data to characterize the complexity of movements and changes in movement behaviours^{1,2}. Kinematic variables are traditionally mapped with motion capture systems³. Current research grade motion capture systems are cumbersome, expensive, often requiring subjects to be tethered and employ markers that are affixed to the body to track movements. Further, these systems often lack the sensitivity to capture small movements of the hand and are labor intensive.

DeepLabCut is an open-source 3D markerless motion capture system that uses deep learning and transfer learning from neural networks for pose estimation and movement tracking^{4,5}. This new technology offers the possibility of gathering accurate kinematic movement data through inexpensive means with less effort as compared to traditional marker based motion capture systems^{4,5}. It would allow for videos taken on a personal phone camera to be used to capture movement patterns and thus allow the behavioural analysis of participants in their natural environment. However, the feasibility, specificity and accuracy of these new systems (such as DeepLabCut) have not been studied; especially for use on neurological and patient populations⁶.

This interdisciplinary project spans the fields of artificial intelligence (AI), Neuroscience, and Kinesiology, with a focus on mapping how movement patterns change as individuals learn a new motor skill. Importantly this approach does not just apply to healthy controls but could be used to map movement associated with recovery from stroke, i.e, relearning a movement. The use of AI technology is important in this line of research as it enables technologies that can be easily adapted to rapidly map complex and abnormal movements in natural environments without the use of a purpose-built motion capture lab⁷. The purpose of this research is to investigate the feasibility of a markerless motion capture system: DeepLabCut, for use in studies of human motor learning. Deliverables of the proposed project include:

1. Development of a data processing pipeline to analyze kinematic variables from a neural network created by DeepLabCut
2. Assess DeepLabCut's ability to discriminate and generalize amongst a variety of movements in a healthy population,

Methods—In this project we created a project pipeline for future research studies to use DeepLabCut to analyze the movement data collected. This involved installing DeepLabcut on personal and lab computers, optimizing the workflow to train the AI network initially, and recruiting the use of Sockeye, UBC’s supercomputer, to help with the speed of analysis. We investigated the number of frames that had to be labeled by a human to train the AI network, and determined best practices for using sockeye to analyze DeepLabCut. We tested the efficacy of DeepLabCut by training the AI network on one participant’s movement video and testing the network on new, unseen participant’s videos. We also carefully documented our process for future researchers wanting to use DeepLabCut. (Fig 1)

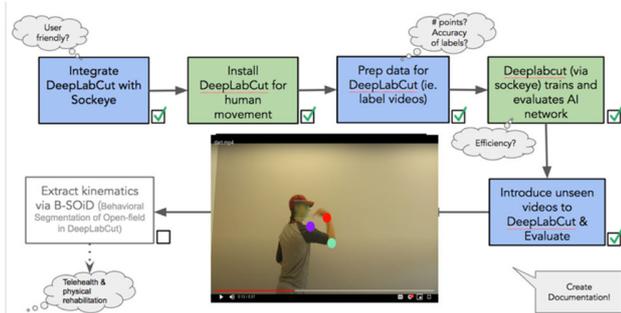


Fig 1- Steps of training and testing the DeepLabCut

Results— These figures show and compare still frames of a dart throwing video that was labelled by a human, first column, with the markers that was labelled by the DeeplabCut, second column.

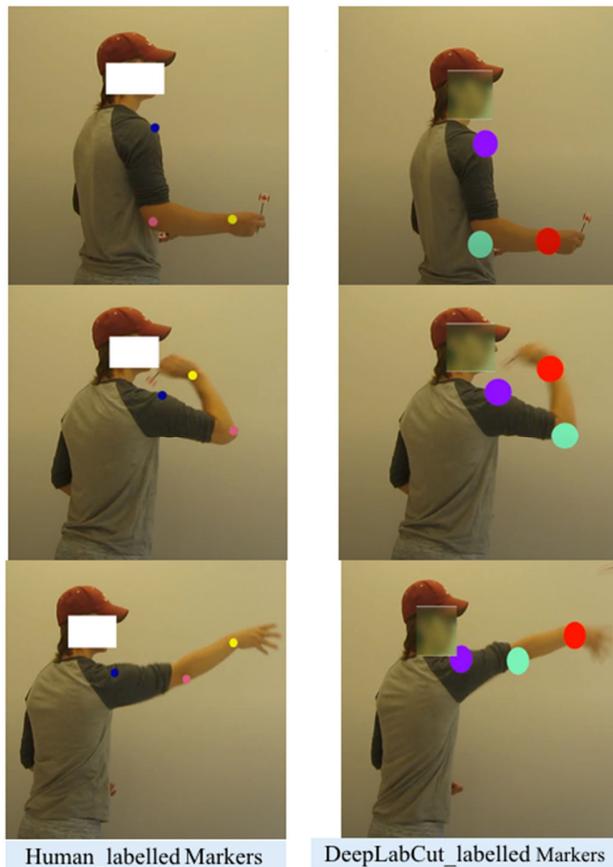


Fig 2-

Conclusion—We determined the feasibility of the machine learning software DeepLabCut and found that it was able to accurately analyze movement patterns and dramatically decreased the amount of human intervention to extract kinematic data from human movements. Furthermore, it decreased the amount of total time required to analyze a data set and output results in a way that was easily exportable to other formats. The software and process was also deemed to be user-friendly, and thus a feasible alternative to existing motion capture systems.

Next steps of this project involve analyzing multiple movement patterns and using a sister software called B-soid to aid with the behavioural segmentation of data. Furthermore, we aim to ultimately test the efficacy of this software with clinical populations and check feasibility using videos captured in a low quality camera. This is to determine the use of DeepLabCut when analyzing video data captured on a phone by a stroke patient in their own home and ultimately increase the feasibility of Telehealth based movement assessments for clinical populations.

Overall, the usefulness of DeepLabCut is in its potential to assess physical movement conveniently. DeepLabCut has removed the need for conventional marker-based motion capture systems. So, in the future, it could be used to remotely conduct physical assessments from the comfort of a patient’s home with high quality and accuracy of the data obtained. This will benefit the large population of patients with movement disorders including neurological, psychiatric, and rehab patients.

References

- 1-Van Dokkum LEH, Mottet D, Laffont I, et al. Kinematics in the brain: unmasking motor control strategies?. *Exp Brain Res*. 2017;235(9):2639-2651. doi:10.1007/s00221-017-4982-8
- 2- Blanken TF, Bathelt J, Deserno MK, Voge L, Borsboom D, Douw L. Connecting brain and behavior in clinical neuroscience: A network approach. *Neuroscience and biobehavioral reviews*. 2021;130:81-90.
- 3- Yunus MNH, Jaafar MH, Mohamed ASA, Azraai NZ, Hossain MS. Implementation of Kinetic and Kinematic Variables in Ergonomic Risk Assessment Using Motion Capture Simulation: A Review. *Int J Environ Res Public Health*. 2021;18(16):8342. Published 2021 Aug 6. doi:10.3390/ijerph18168342
- 4- Wei, Kunlin, and Konrad P. Kording. "Behavioral Tracking Gets Real." *Nature Neuroscience*, vol. 21, no. 9, 2018, pp. 1146-1147.
- 5- Mathis, Alexander et al. "DeepLabCut: markerless pose estimation of user-defined body parts with deep learning." *Nature neuroscience* vol. 21,9 (2018): 1281-1289. doi:10.1038/s41593-018-0209-y
- 6-Mathis, Alexander et al. "A Primer on Motion Capture with Deep Learning: Principles, Pitfalls, and Perspectives." *Neuron* vol. 108,1 (2020): 44-65. doi:10.1016/j.neuron.2020.09.017
- 7- Nath, Tanmay et al. "Using DeepLabCut for 3D markerless pose estimation across species and behaviors." *Nature protocols* vol. 14,7 (2019): 2152-2176. doi:10.1038/s41596-019-0176-0

Acknowledgments—This project was supported by the 2021 UBC Faculty of Medicine, Multidisciplinary Research Program in Medicine (MRPM)

Special thanks to all members of UBC the Neuroplasticity, Imagery, and Motor Behaviour Laboratory (NIMBL) and the Brain Behavior Lab (BBL)

Contact information:
Fatemeh Khounsaraian fkhounsa@student.ubc.ca

Introduction of Integrated Image Deep Learning Solution and how it brought laboratorial level heart rate and blood oxygen results to everyone

Zhuang Hou

Booth School of Business
University of Chicago
Chicago, United States
zhou10@chicagobooth.edu

Xiaolei Cao

Booth School of Business
University of Chicago
Chicago, United States
xcao2@chicagobooth.edu

Abstract—The general public and medical professionals recognized the importance of accurately measuring and storing blood oxygen levels and heart rate during the COVID-19 pandemic. The demand for accurate contact-less devices was motivated by the need for cross-infection reduction and the shortage of conventional oximeters, partially due to the global supply chain issue. This paper evaluated a contactless mini program HealthyPai’s heart rate (HR) and oxygen saturation (SpO₂) measurements compared with other wearable devices. In the HR study of 185 samples (81 in the laboratory environment, 104 in the real-life environment), the mean absolute error (MAE) \pm standard deviation was 1.4827 ± 1.7452 in the lab, 6.9231 ± 5.6426 in the real-life setting. In the SpO₂ study of 24 samples, the MAE \pm standard deviation of the measurement was 1.0375 ± 0.7745 . Our results validated that HealthyPai utilizing the Integrated Image Deep Learning Solution (IIDLS) framework can accurately measure HR and SpO₂, providing the test quality at least comparable to other FDA-approved wearable devices in the market and surpassing the consumer-grade and research-grade wearable standards.

Index Terms—remote photoplethysmography, heart rate, oxygen saturation, contactless measurement, mini program

I. INTRODUCTION

Heart Rate (HR) and oxygen saturation (SpO₂) are the most focused measurements during the COVID-19 pandemic as they are both essential metrics providing critical states of a person to determine whether a person is infected by COVID, especially at the early stage of COVID-19 due to the lack of fast test kits. Even if the COVID mortality rate in the US has in general declined following the virus mutation, HR and SpO₂ are still consistently tested among the health-focused public. Separating people from exposing them to the virus as a strategy to slow down its spread was widely leveraged across the globe. There are numerous strategies to avoid gatherings, such as social distancing, home quarantine, or centralized quarantine. Vaccination is another successful strategy slowing down the virus spread in a community. However, the vaccine effectiveness against COVID-19 declines in a nine-month period. According to [7], the two-dose regimens of messenger RNA (mRNA) vaccines BNT162b2 (Pfizer–BioNTech) (30 μ g per dose) and mRNA-1273 (100

g per dose), vaccine effectiveness against Covid-19 declined from 94.5% and 95.9% , respectively, at 2 months after the first dose and to 66.6% and 80.3%, respectively, at 7 months. And the fast mutation of COVID-19 declines the vaccine effectiveness. According to [8], no effect against the omicron variant was noted from 20 weeks after two ChAdOx1 nCoV-19 (AstraZeneca) doses, whereas vaccine effectiveness after two BNT162b2 (Pfizer–BioNTech) doses was 65.5% at 2 to 4 weeks, dropping to 8.8% at 25 or more weeks. Therefore, decentralizing the medical level devices is the key to both strategies.

The establishing methods on vital signs measurement using smartphone can be split into two categories: Photoplethysmography (PPG)-based [2], [3] and remote Photoplethysmography (rPPG)-based [4] methods. With each cardiac cycle, the heart pumps blood to the periphery. With the change of blood volume caused by the pressure, the light absorption is variant. Both methods capture the minor variation of the light reflection and generate a signal for vital signs measurement. The PPG-based method is similar to the principle of some oximeters which requires the users to put their fingers to cover the rear camera with illuminated flashlight. The rear camera can capture the variations of the flow of blood within the vessels. Kanva et al. [2] measured SpO₂ and HR with smartphone rear camera. Nemcovaa et al. [3] developed an Android application to estimate HR, SpO₂ and blood pressure simultaneously.

The rPPG method applies the same principle of PPG, while the users need to turn on the front camera to capture the face videos. Kwon et al. [4] built a smartphone application called FaceBEAT that can measure users’ heart rate with face videos. Qiao et al. [5] built an application that can measure HR and heart rate variability (HRV) by using the smartphone front camera. Nam et al. [6] utilized dual cameras to monitor HR and respiration rate. The front camera is applied to capture the motions caused by heartbeats and respirations.

The paper presents a robust workflow, Integrated Image Deep Learning Solution (IIDLS), that realized the laboratory results using real-life quality facial videos. We proved that IIDLS can provide medical level accuracy in HR and SpO₂

measurement and requiring only on a smartphone. We have tested the framework in the laboratory environment in ensuring the theoretical highest quality results that it can generate and compared the results with the medical chest strap. We have also compared the real-life results with those of the other products, such as the Apple Watch, Fitbit, etc. , claiming offering the same in the markets.

Our IIDLS framework is illustrated in Section II. Section III presents our experiments in both laboratory and real-life using environments as well as corresponding measurement results. Section IV briefly concludes our works and presents the future works.

II. APPROACH

The overall workflow of the IIDLS estimation framework is presented in Fig. 1. The IIDLS consists of two major steps, 1) data processing to have the video denoised and the region of interests (RoI) selected, and 2) the data modeling step to create features from the data and predict the outputs.

1) *Data Processing*: The first major step shown in the workflow is input data processing. The input of the system is a 20-second facial video from the user. The processing includes luminance analysis, face detection and face RoIs selection which can select multiple face patches for vital signs measurement. After selecting the RoI, a candidate blood volume pulse (BVP) signal is extracted from the RoI. We apply established algorithms to identify the best RoI from the users' faces that maximize the true signal.

2) *Data Modeling*: Once data has been extracted from the RoIs of the facial videos, it will be used as the input to the deep learning algorithms. The model has been trained by 104 videos taken in a real-life environment and 84 videos taken in a laboratory environment preparing for the corresponding comparisons.

III. RESULTS

There are two sets of comparison results depending on the testing environment including the laboratory environment and the real-life environment.

A. Laboratory Results

Nine participants, with one female, all of the East Asian descendants and of different ages (20s - 60s) enrolled in the experiment in the laboratory environment [1]. The participants were asked to sit on a chair 1.5m from the camera. The light in the video is consistent from the beginning to the end. There are three 60-second sessions, relax, exercise, and relax sessions, in each 180-second video, where the participants were asked to perform handgrip exercise in the exercise session. The video resolution is VGA (640×480), and the frame rate is 300 fps. Each video is extracted into nine 20-second sequences as the input data fitting the algorithm. Therefore, there are 81 samples for both heart rate and SpO2 comparison.

Out of all 81 samples, the mean error (ME) between the HR predicted value (through the HealthyPai algorithm) and HR actual value (through the conventional testing devices) is

-0.1951 ± 2.2876 . The mean absolute error (MAE) between the HR predicted value (through the HealthyPai algorithm) and HR actual value (through the conventional testing devices) is 1.4827 ± 1.7452 .

Compared to polar chest straps, the laboratory results demonstrate that HealthyPai, without any contact, can provide testing results comparable to the polar chest straps that measure heart rate with maximum precision.

TABLE I
HEART RATE MONITOR DIFFERENCES FROM ELECTROCARDIOGRAM
(MEAN \pm SD)

Device	N	Paired ME	Paired MAE
HealthyPai	81	-0.2 ± 2.3	1.5 ± 1.7
Polar Chest Strap	160	0.2 ± 1.4	0.7 ± 1.2

B. Real-life environment results

In total, there are 25 participants (9 female, 16 male) providing 104 samples in the heart rate comparison and 24 samples (23.08 % of all samples) in the SpO2 comparison. All 100 % of the participants are of East Asian descent and taking the videos while resting.

Out of all 104 samples, the ME between the HR predicted value (through the HealthyPai) and HR actual value (through the conventional testing devices) is 0.6538 ± 8.9332 . The MAE between the HR predicted value (through the HealthyPai) and HR actual value (through the conventional testing devices) is 6.9231 ± 5.6426 .

HealthyPai provides the quality of results that is at least comparable to other FDA-approved wearable devices and beyond the consumer-grade wearables, which gives an MAE of 7.2 ± 5.4 , and research-grade wearable equipments, which gives an MAE of 13.9 ± 7.8 . The comparison results are significant through more samples that are collected.

TABLE II
HEART RATE MONITOR DIFFERENCES FROM ELECTROCARDIOGRAM
(MEAN \pm SD)

Device	N	Paired ME	Paired MAE
HealthyPai	104	0.7 ± 8.9	6.9 ± 5.6
Apple Watch	78	-1.7 ± 10	5.0 ± 9.0
Fitbit	80	1.0 ± 8.5	5.7 ± 6.3
Garmin	80	0.8 ± 15	9.2 ± 12
TomTom	76	1.3 ± 9.8	5.4 ± 8.2
Consumer-grade wearables	-	-	7.2 ± 5.4
Research-grade wearables	-	-	13.9 ± 7.8

Out of all 24 participants' results, the mean error (ME) between the SpO2 predicted value (through the HealthyPai) and SpO2 actual value (through the conventional testing devices) is -0.3375 ± 1.2666 . The MAE (Mean Average Error) between the SpO2 predicted value (through the HealthyPai) and SpO2 actual value (through the conventional testing devices) is 1.0375 ± 0.7745 .

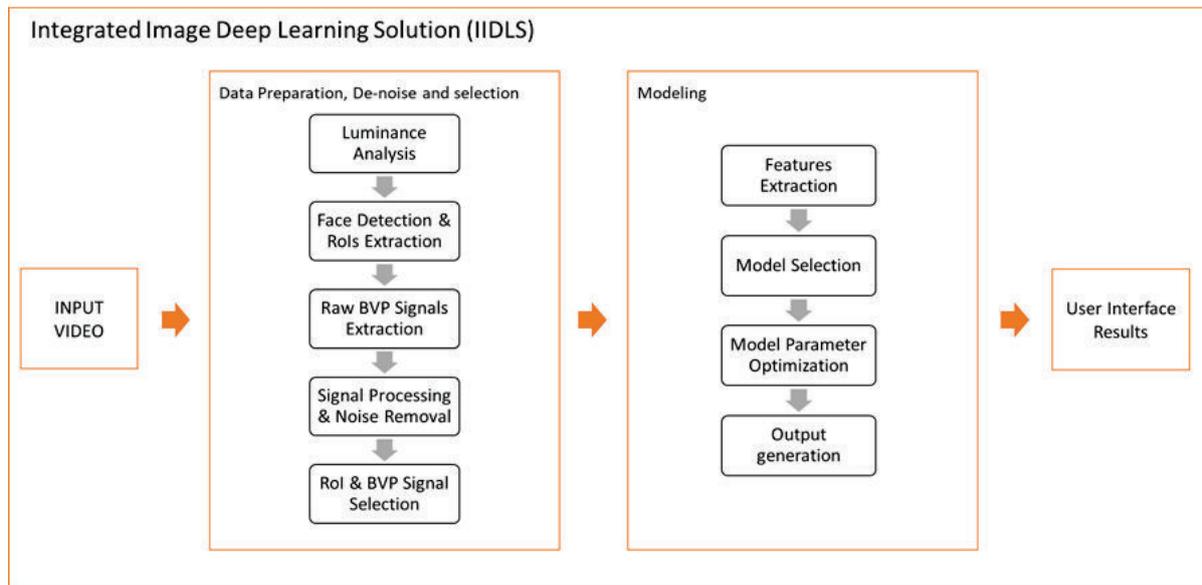


Fig. 1. Integrated Image Deep Learning Solution (IIDLS) framework.

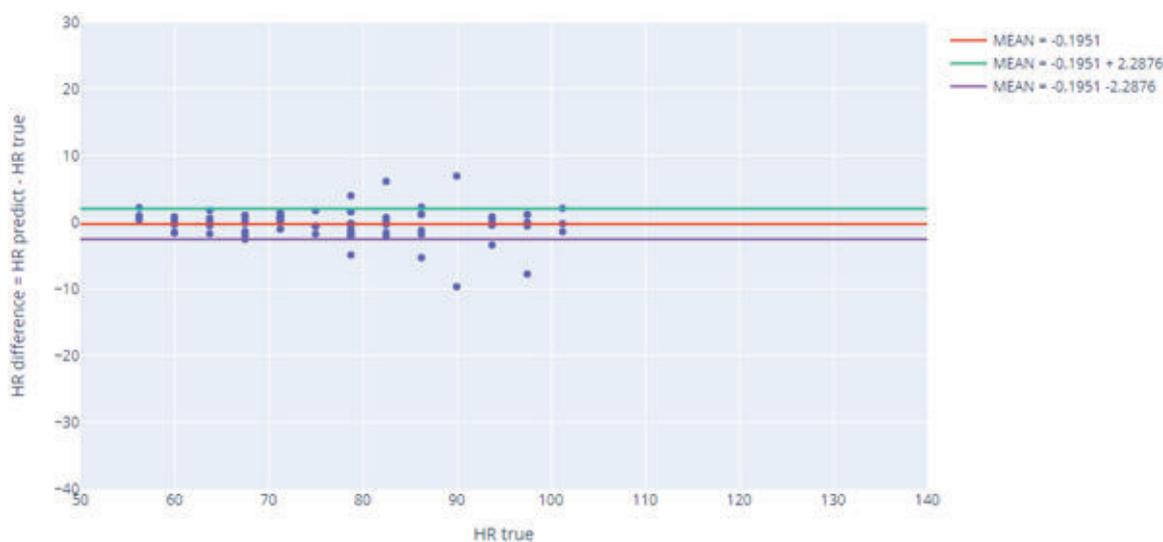


Fig. 2. HR difference as HR estimated - HR true value v.s. HR true value

IV. CONCLUSION

IIDLS has been thoroughly tested by collecting 185 samples (81 in the lab, 104 in the real-life environment). The comparison shows that IIDLS provides at least comparable quality to other FDA-approved wearable devices in the market and surpasses the consumer-grade and research-grade wearables standards. The environment in which a video was taken is a critical factor affecting the quality of the results. The closer the surrounding environment is to the laboratory, the better quality the user will receive. IIDLS is optimizing the de-noise process to cleanse the signal to the model and has proactively intervened users to either pause the video taking or re-take the

videos if the surrounding environment is less optimal.

REFERENCES

- [1] Maki Y, Monno Y, Yoshizaki K, Tanaka M, Okutomi M. Inter-Beat Interval Estimation from Facial Video Based on Reliability of BVP Signals. *Annu Int Conf IEEE Eng Med Biol Soc.* 2019 Jul;2019:6525-6528. doi: 10.1109/EMBC.2019.8857081. PMID: 31947336.
- [2] A. K. Kanva, C. J. Sharma and S. Deb, "Determination of SpO2 and heart-rate using smartphone camera," *Proceedings of The 2014 International Conference on Control, Instrumentation, Energy and Communication (CIEC)*, 2014, pp. 237-241, doi: 10.1109/CIEC.2014.6959086.
- [3] Nĕmcová, Andrea & Jordanova, Ivana & Varecka, Martin & Smisek, Radovan & Maršánová, Lucie & Smítal, Lukáš & Vitek, Martin. (2020). Monitoring of heart rate, blood oxygen saturation, and blood pressure using a smartphone. *Biomedical Signal Processing and Control.* 59. 101928. 10.1016/j.bspc.2020.101928.

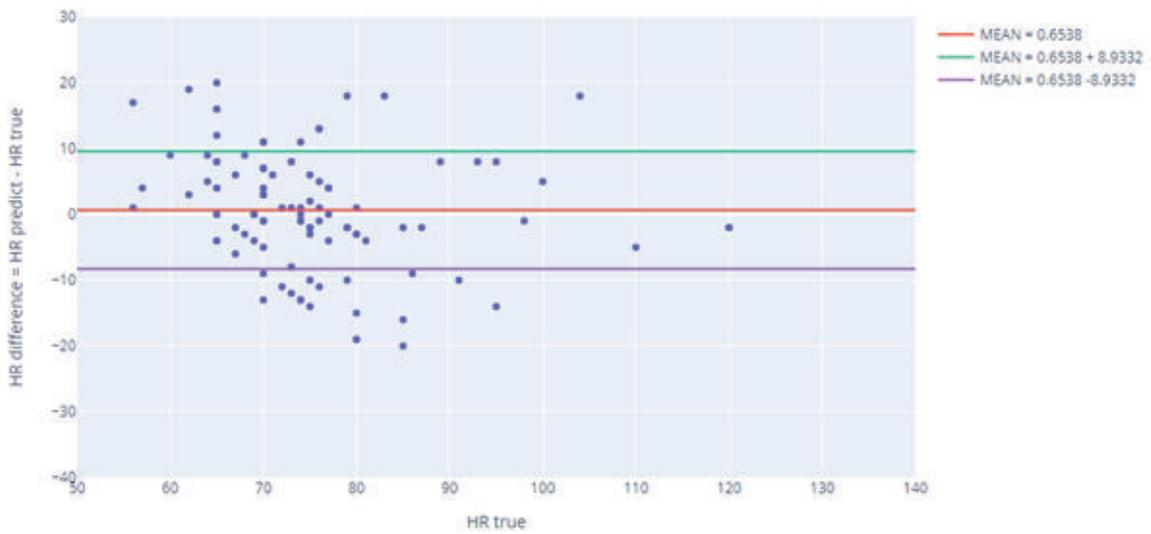


Fig. 3. HR difference as HR estimated - HR true value v.s. HR true value

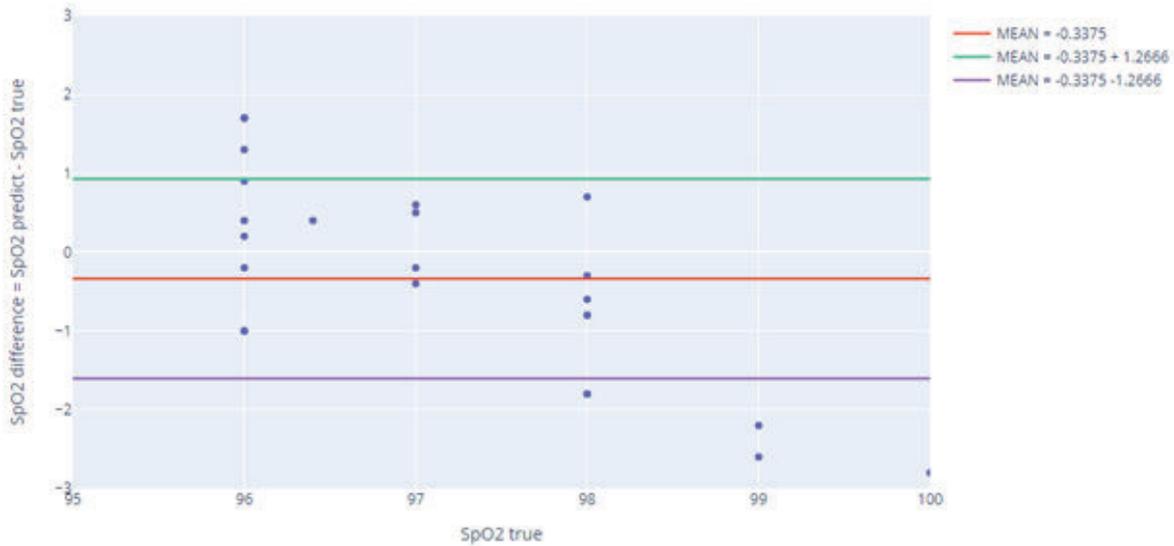


Fig. 4. SpO2 difference as SpO2 estimated - SpO2 true value v.s. SpO2 true value

- [4] S. Kwon, H. Kim and K. S. Park, "Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone," 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2012, pp. 2174-2177, doi: 10.1109/EMBC.2012.6346392.
- [5] D. Qiao, F. Zulkernine, R. Masroor, R. Rasool and N. Jaffar, "Measuring Heart Rate and Heart Rate Variability with Smartphone Camera," 2021 22nd IEEE International Conference on Mobile Data Management (MDM), 2021, pp. 248-249, doi: 10.1109/MDM52706.2021.00049.
- [6] Nam, Yunyoung & Kong, Youngsun & Reyes, Bersain & Reljin, Natasa & Chon, Kaye. (2016). Monitoring of Heart and Breathing Rates Using Dual Cameras on a Smartphone. PLOS ONE. 11. e0151013. 10.1371/journal.pone.0151013.
- [7] Lin DY, Gu Y, Wheeler B, Young H, Holloway S, Sunny SK, Moore Z, Zeng D. Effectiveness of Covid-19 Vaccines over a 9-Month Period in North Carolina. N Engl J Med. 2022 Mar 10;386(10):933-941. doi: 10.1056/NEJMoa21117128. Epub 2022 Jan 12. PMID: 35020982; PMID: PMC8781317.
- [8] Andrews, Nick and Stowe, Julia and Kirsebom, Freja and Toffa, Samuel and Rickeard, Tim and Gallagher, Eileen and Gower, Charlotte and Kall, Meaghan and Groves, Natalie and O'Connell, Anne-Marie and Simons, David and Blomquist, Paula B. and Zaidi, Asad and Nash, Sophie and Aziz, Nurin Iwani Binti Abdul and Thelwall, Simon and Dabrera, Gavin and Myers, Richard and Amirthalingam, Gayatri and Gharbia, Saheer and Barrett, Jeffrey C. and Elson, Richard and Ladhani, Shamez N and Ferguson, Neil and Zambon, Maria and Campbell, Colin NJ and Brown, Kevin and Hopkins, Susan and Chand, Meera and Ramsay, Mary and Bernal, Jamie Lopez. Effectiveness of COVID-19 vaccines against the Omicron (B.1.1.529) variant of concern. Cold Spring Harbor Laboratory Press. 2021.12.14.21267615. doi: 10.1101/2021.12.14.21267615.

Protocol Selection of Advanced Imaging Exams using Multi-steps Deep Learning Models

C. Cronister, T. B. Issa, D. Uminsky, R. W. Filice

Abstract—The radiology workflow is a time consuming process that involves significant administrative procedures. One of the administratively intensive tasks that pulls the physician away from patient facing work is the daily exercise of matching the scheduled exams with specified protocol for each exam. We present protocol selection networks (PSNet), a multi-step deep learning model that automatically recommend to physician specified protocols for the top 50 most common protocols at *Anonymize Hospital*. PSNets has the potential to lower the burden for physician at *Anonymize Hospital* to specify protocols by up to 70%. Yet PSNet achieves human expert accuracy in protocol selection with a weighted average top one recall of 90% and precision of 90% across the all protocols. Furthermore, we studied the interpretability of our machine learning system and show that the ‘reason’ and the ‘appointment type’ are the most important variables in the recommendation decision. We also present the results of a pilot hospital deployment study where the first-step model show a top one accuracy of 76% and top two accuracy of 87%. The multi-step model achieve a maximum top one accuracy of 83% and an average of 63%, a maximum top two accuracy of 93% and an average of 77%.

Keywords—Protocol, Machine Learning, Radiology, Deep Learning, Medicine and AI, Deployment, Data Wrangling.

I. INTRODUCTION

In recent decades, the use of medical imaging has drastically increased and this has put a burden on radiology departments everywhere [1]. In the US alone 80 million CT scans are performed annually [2]. Technological advances such as electronic health records (EHRs) have been introduced to help increase efficiency and decrease time physicians spend on seemingly trivial tasks [3]. Even with these improvements to workflow, radiologists still report high levels of fatigue – especially with the added burden of the pandemic [4]. Additionally, 59% of doctors report that their fatigue is due to the many bureaucratic tasks that require their oversight [5], [6]. The saturation of the radiologist workflow with bureaucratic tasks and interruptions may lead to shorter time reviewing an examination which has been shown to decrease the accuracy of a diagnosis in some cases [7]. What’s more, this higher workload is leading to longer days, which turn into nights, and this is leading to more errors in diagnoses [8]. Recent developments in machine learning(ML), specifically deep learning(DL), have been used to ease the burden on radiologists through ML and DL based automation of (or some parts) of the radiology workflow. Rajpurkar et al in [9] built a DL model that autonomously detects thoracic diseases from a chest radiograph. These AI developments have commonly focused on pixel-based diagnostic or triage tasks. However, these methods, or similar methods, can be applied

to other types of data from the radiology field to alleviate and automate mundane tasks. One such example is protocoling an upcoming exam. At *Anonymize Hospital*, from 2015 to 2021, Physicians entered on average 2084 protocols per month (see figure 1). At 5 seconds per protocol, each physician expand about 3 hours per month on protocoling. In this paper, we propose a multi-stage deep learning architecture to build a highly efficient and accurate machine learning model that helps hospitals automate protocol prescription. A crucial step toward this objective is understanding the radiology workflow of exam protocoling.

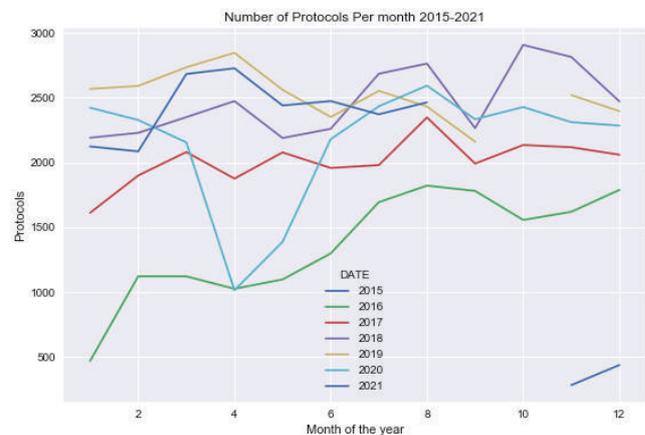


Fig. 1. Protocoling at *Anonymize Hospital* 2015-2021.

Radiologist Workflow

The workflow for a patient’s radiology exam protocol prescription and interpretation at *Anonymize Hospital* is split into the following steps (Figure 2):

- 1) Patient sees a physician or otherwise learns they need some type of advanced imaging exam.
- 2) Patient schedules their exam through central scheduling.
- 3) Patient and exam information is retrieved via Health Level Seven (HL7) scheduling messages, Fast Healthcare Interoperability Resources (FHIR) queries, and other sources. This information is displayed in a custom web application for protocoling.
- 4) Radiologists look at pending exams and enter a protocol for each exam that specifies exactly what should be done. Usually these are selected from a preset list of protocols, but they can enter custom information as needed.
- 5) When the patient arrives, the technologist reviews this protocoling information and performs the exam as prescribed.

Tahir Bachar Issa is with the San Jose State University, United States (e-mail: tahirbachar.issa@sjsu.edu).

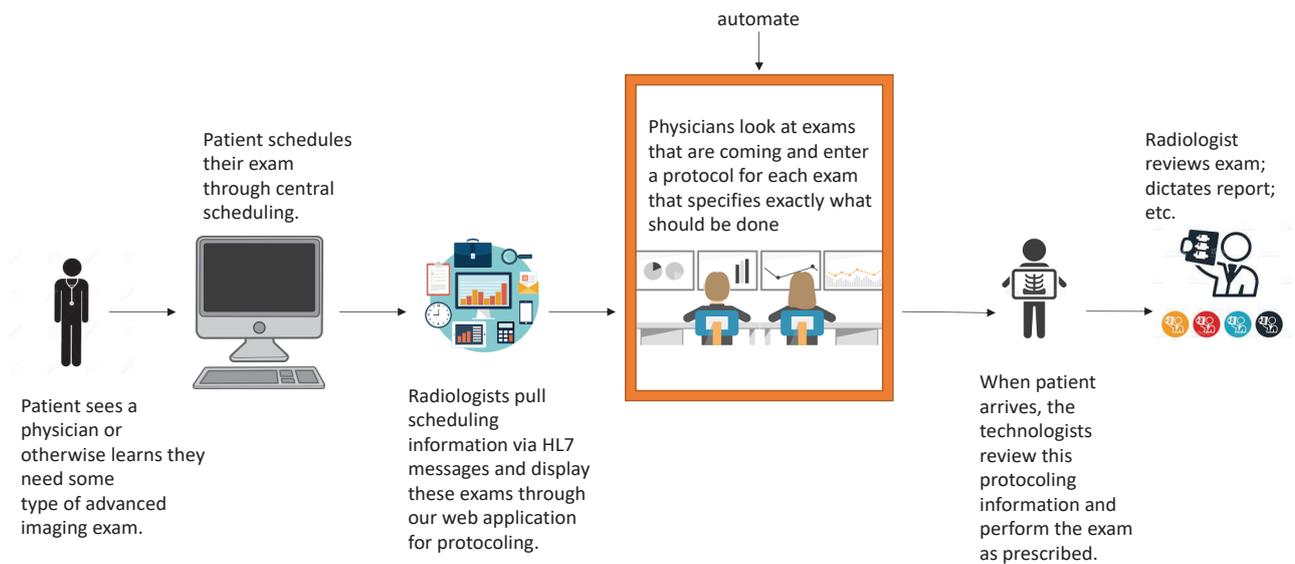


Fig. 2. A comprehensive look at the workflow for a radiology patient and physician with highlighted point of improvement.

6) A radiologist reviews the exam and dictates the report.

This is a time consuming process that involves the administratively intensive task of protocoling in steps 3 and 4 that pull the radiologist away from direct patient care. Many recent machine learning studies, especially those focused on deep learning algorithms, have tried to automate, or assist, the last interpretive step of the workflow [10]. In this paper, we focus on building machine learning models that assist the radiology department in automating the fourth step of the workflow.

Previous approaches applying Machine Learning to Protocoling

As noted before, we focus our AI development efforts on protocoling, the fourth step of the workflow and is an under-explored area for AI practitioners, relative to the efforts spent on image processing. We review here some of the recent approaches to develop tools to automate the process.

A natural language processing (NLP) classifier to parse patient data and then predict protocols is developed in Kalra et al. [11]. The model focused on an initial data set containing 108 distinct protocols, where our initial data had 6657 distinct protocols as a result of the ability to individuate the protocol entry in our system. For their study, they leveraged the full language corpus of rich patient history and EMR to train on and achieved 95% accuracy on 69% of protocols in automation mode with the remaining 31% of cases, seeing 92% accuracy. The use of patient history and full EMR presents challenges for transferability and production. Our model focused on a more limited source of data that attempts to mimic the normal workflow for protocoling patients.

López-Úbeda et al. [12] developed a native Spanish based NLP model that also leveraged the full patient history corpus. They further split their data into CT and MRI protocols but achieved an overall lower accuracy than our work presented here. Brown et al. compares multiple machine learning techniques to find the most accurate model to predict MRI protocols in neuroradiology [13]. This multi-model approach was able to reach 95% accuracy, it was predicting from a smaller subset of 41 protocols drawn from the subfield of neuroradiology. Bhat et al. [14] implement a graph-based semi-supervised learning technique, where the nodes of the graph are concepts from a pre-selected ontology of healthcare terms. This is one of the few references that present a *deployment* study as well. This model is limited to just seven protocols and achieved an overall accuracy of 60% in deployment. We are able to achieve a 63% deployment accuracy on 50 fifty protocols.

Lee et al. uses a convolutional neural network classifier built on short-text classification to predict musculoskeletal MRI protocols [15]. This model had access to high signal features such as the type of contrast used and the protocols were split into routine or tumor/infection protocols which resolved many of the nuanced differences between protocols. Additionally, the model also only predicts large categories within the musculoskeletal MRI protocols – i.e. shoulder, pelvic bone, etc. Since this model predicts a larger category, the model requires refinement from a physician as they must ultimately prescribe the specific protocol within the larger predicted category.

Our 2 steps approach leverages multiple neural networks models to accurately predict protocols across the entire ra-

diology protocol system. Moreover we intentionally take the approach of limiting the data inputs to patient and exam information in the way it is actually stored in practice at *Anonymize Hospital*; including only existing scheduling information and not the full patient EMR. Finally, we are able to report results of a pilot deployment at *Anonymize Hospital*, that are competitive or exceeds the few other reported deployment studies.

In this paper we will build a machine learning model that assists hospitals in automating the fourth step of the radiology workflow in an effort to decrease the workload on these radiologists. The main contributions of this paper are as follows:

- 1) We have created a novel pipeline that takes patient and exam information and runs it through our two stage machine learning system and returns the “top 2” protocol procedural category as well as the “top 2” specific protocol predicted. We also return the protocols probabilities and F1 scores so that the physician can use the information to either confirm or reject the protocol. Our model achieves a top one recall of 98% and precision of 98% when predicting the correct procedural category. Our model additionally achieves a weighted average top one recall of 90% and precision of 90% across the 50 protocols.
- 2) The model takes patient and exam information in the way it is actually stored in practice, so there is no need to change the way this data is collected in order to deploy the model. There is no requirement for accessing the full patient EMR in this process.
- 3) We also present a pilot deployment of this model at *Anonymize Hospital*. In this setting our model reaches an average top one and top two accuracy respectively of 63% and 77% on fifty protocols. Previous results deployment studies are sparse and one such study reached a stage for deployment with top one accuracy of 60% on only seven protocols[14].

II. DATA COLLECTION, PREPROCESSING AND FEATURE ENGINEERING

A. The Data

The data used in this paper was recorded over a 6 year period (2014-2020), collected and labeled by expert radiologists at *Anonymize Hospital*. Patient and examination information is pulled from HL7 Scheduling Information Unsolicited (SIU) messages. Sub-specialty and modality information was provided through an existing lookup table. A nightly cron job then pulls this information, and sometimes the front desk uploads a scanned paper order (for example if the patient has a paper prescription or is not in the hospital system). We note here that a full review of EMR patient data is not a normal part of this process and thus we don't include as a part of our training. Using the available data from the HL7 SIU, the radiologists then manually protocol each exam - these are the labels for our model. Inputs are the schedule and the lookup data, and the labels are the protocols that expert radiologists manually entered.

The raw data had 74545 rows with 10000+ unique protocols and 14 features. Assigned protocols are the model labels and the raw features are: reason for visit, appointment type, duration of visit, exact machine used for exam, a text description of the exam, location of the exam, radiology location, patient date of birth, patient sex, referring physician last name, referring physician first name, primary notes from the scheduler, secondary notes from scheduler, section of the body for the exam and type of machine used for exam. The reason, appointment type, primary notes, secondary notes, and section were all coded features. For example, appointment type is a three character code, such as TR1, which corresponds to a CT of the thorax, abdomen and pelvis. The raw data, directly created from the day to day needs of the hospital, required pre-processing in order to be in a useful form for modeling. After the data cleaning process describe below, the cleaned data contains 56840 rows with 6657 unique protocols and 14 features. The population represented in the cleaned data was mostly adults between the ages of 50 and 72 with 49% of the individuals identifying as females and 51% males. 68% of the exams in our training set were MR scans and 32% were CT scans. We then train the model on the top fifty most common protocols with a train and validation split of 80/20 resulting in 45472 and 11368 rows respectively. Our test set is all the data the hospital from the first four months of 2021 and contains 13123 rows of data.

B. Data Preprocessing

Since the data was drawn directly from the workflow process, the raw data set had over 10000 unique protocols. While many of these protocols were valid, many needed to be reviewed and cleaned. Reasons for review and cleaning ranged from misspellings, to notes accidentally recorded as labels, to an extra space after the protocol. We in collaboration with our radiologists team, cleaned and properly consolidate as many of the labels as possible. The cleaned data contains 56840 rows with 6657 unique protocols.

We first lower-case the labels to harmonize differences in capitalization. Next, we explore the data set further and corrected many misspellings. In addition to misspelling and non uniformity of case-use in the protocols, we found many inconsistencies in the order of the letters or words within the protocols such as “Liver 3b” and “3b liver” or “1”, “1 liver”, “liver”, and “1 liver with iv” which are all materially the same protocols, which we corrected or matched in consistent coordination and communication with the attending radiologists.

It also turns out that the “reason” category contained a mix of text-based data and ICD-10 codes. ICD-10 codes are a nearly ubiquitous coding scheme for insurance and billing purposes that maps a disease to a code that is a mix of digits and letters. We chose a consistent form for by transforming each word-based reason into its ICD-10 code. If there were less than 5 instances of a certain raw reason, it was given the label 'other'.

Similar to the reason feature, we also encoded the primary notes feature. Many of these primary notes were codes separated by backslashes, so we removed the backslashes to create

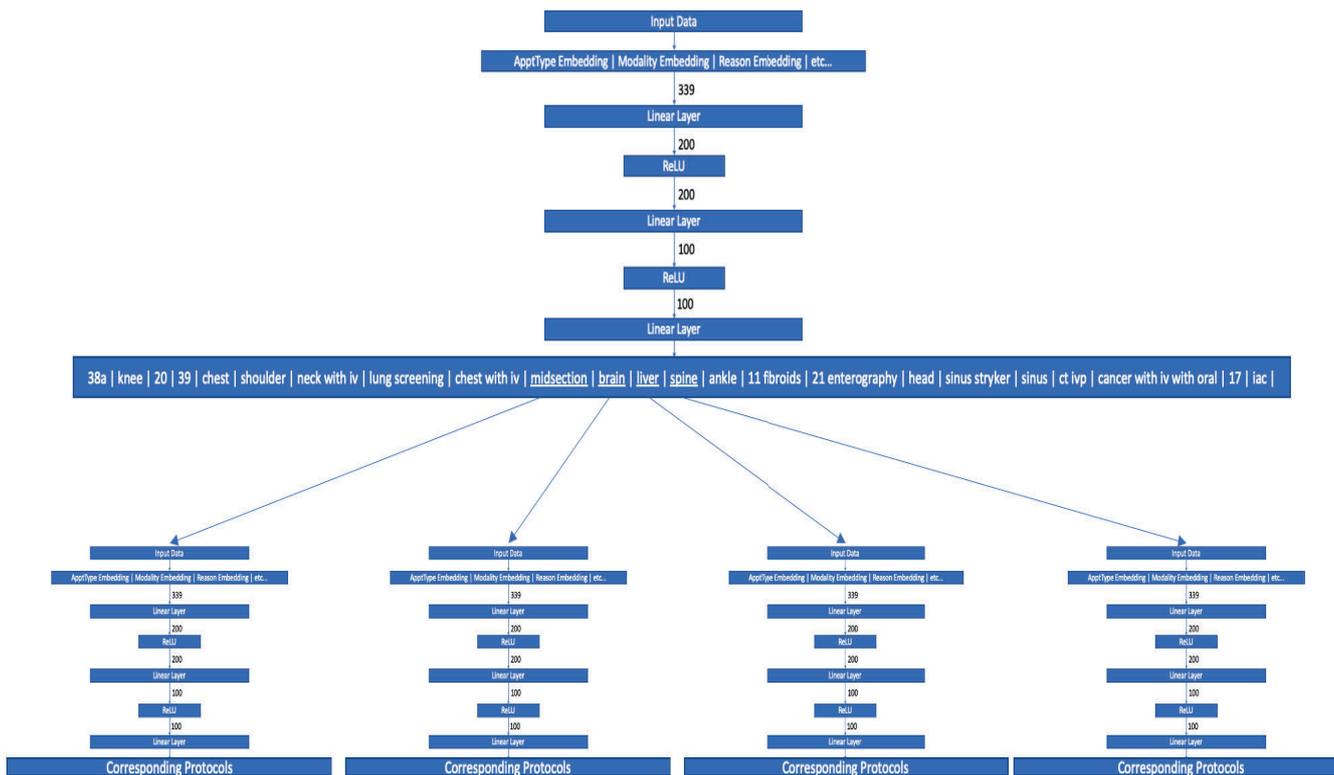


Fig. 3. The data flow through our two level model. A row of patient and examination data enters the large category model and either a unique protocol or category of protocols is predicted. If a category of protocols is predicted, the data is sent to the corresponding specific protocol model to predict the specific protocol.

a list of codes, ordered them alphabetically and then joined them into a string. We did this so that *a/b/c* and *c/a/b* were not treated as different notes when we encoded the column as a categorical variable. Again, if there were less than 5 instances of a certain raw primary note, it was given the label 'other'.

Finally, the description feature was text-based, so we applied a tfidf vectorizer to create a corresponding matrix of continuous features. This brings us to 33 features when we take the original 14, drop the raw description and add the 20 tfidf features from the description feature. We chose 20 features as we did not want to overfit to specific tokens in the description feature. The tfidf features can be seen in table I below.

III. METHODOLOGY

In the process of assigning the correct protocol we identified certain categories of protocols that were more challenging to assign correctly than others. We noted that the protocols associated with brain, midsection, liver and spine were the most challenging to identify correctly. In the first step of our modeling, we build a deep neural network that predicts either a exact protocol or one of the 4 “challenge areas” categories mentioned above. Next, if our first step model predicts one of the specific protocols, we take this as our final prediction. If the first step predicts one of the 4 “challenge” areas, then the data is passed through a second deep neural network as seen in figure 3 that predict a unique final protocol. The second step models are specific protocol models, which are built on each of

Vocab	Feature
abd	Feature 0
abdomen	Feature 1
and	Feature 2
brain	Feature 3
cervical	Feature 4
contrast	Feature 5
ct	Feature 6
extrem	Feature 7
lumbar	Feature 8
mrcp	Feature 9
mri	Feature 10
neck	Feature 11
pelvis	Feature 12
prostate	Feature 13
rt	Feature 14
spine	Feature 15
thorax	Feature 16
with	Feature 17
without	Feature 18
without	Feature 19

TABLE I
THE VOCABULARY CREATED FROM THE TFIDF VECTORIZER PASSED OVER THE DESCRIPTION FEATURE.

the “challenge” categories: brain, midsection liver and spine. We proceed in this paper by naming the first step model as our “large category” model and refer to models in our second step that predict a single category of protocols as a “specific protocol” model.

TABLE II

THE PRECISION AND RECALL SCORES FOR EACH MODEL. THE PRECISION AND RECALL SCORES ARE REPORTED AS WEIGHTED AVERAGES.

	Neural Network
All Labels	
Validation Precision	0.96
Validation Recall	0.96
Test Precision	0.98
Test Recall	0.98
Brain	
Validation Precision	0.85
Validation Recall	0.85
Test Precision	0.85
Test Recall	0.85
Midsection	
Validation Precision	0.95
Validation Recall	0.95
Test Precision	0.95
Test Recall	0.94
Liver	
Validation Precision	0.93
Validation Recall	0.92
Test Precision	0.93
Test Recall	0.92
Spine	
Validation Precision	0.92
Validation Recall	0.92
Test Precision	0.93
Test Recall	0.93

The deep neural network architecture we used for both the first and second step of our model is fast.ai tabular model [16]. The tabular model is a feed-forward deep neural network where the categorical features are passed through a learned embedding. From there, the model is composed of three linear layers with a ReLU at the end of each of the inner layers. We used fast.ai’s built-in learning rate finder and their slice method to choose the best learning rate range for each of our models and regularization techniques such drop out, BatchNorm and weigh decay to combat model over-fitting.

More details on this deep learning model are given in the fast.ai documentation [16].

A. Deployment process

The deployment study we conducted of our model was done as follows:

- 1) The scheduling information pulled from HL7 SIU is passed through the multi stage deep learning model.
- 2) The model will return the top two protocols, their probabilities, their F1 scores and an overall confidence metric.
- 3) A physician will then either confirm or reject the predictions.
- 4) The results are then recorded and reported.

This highlights that the goal of our model is physician assistance rather than complete automation – as we believe qualified eyes should make the final decision on the prescribed protocol.

IV. RESULTS

A. Metrics and Baseline Model

We evaluate our model with precision and recall with an emphasis on precision as it was determined that it was more critical that protocols we predict are correct. We take a random split 80% of the data between 2015 and 2020 as our training data, the remaining 20% as our validation set and data from the first four months of 2021 as our test set.

B. Model Performance on Validation and Test Set

Our two-step model performed quite well on the validation and test sets, as shown in table II. The large category protocol neural network has the same weighted average precision and recall scores of 98% on both the test and validation sets.

The specific protocol neural network models for brain, liver, midsection and spine, performed quite well on the test and validation sets as well. We see consistent numbers between the test and validation sets for each of our models signifying they will generalize quite well, as seen in table II.

C. Model Interpretation

We study feature importance in our deep neural networks by using the permutation feature importance technique, which consist of shuffling a feature in our data set and observing the change in metrics with this nonsensical feature (for more details on this technique, we refer the reader to [17]). We found that the appointment type and reason for visit are the most important features for the large category neural network (see figure 5). This supported our a priori notions of which features should be useful to this category model.

When we study the specific protocol neural networks we see some similarities and differences. Most of the specific protocol neural networks focus on the reason for the visit or the sub-specialty of the referring physician as the most impactful feature. However, the feature importance varies across models from a single tfidf feature to appointment type (see figure 5).

V. DEPLOYMENT

The grand challenge of any AI model is it’s actual effectiveness in production. Traditionally this is where significant degradation of performance occurs [18].

To understand our algorithms performance, we piloted a one week deployment of our AI model at *Anonymize Hospital* where it is was integrated into their already existing protocoling system. The data was collected from the hospital’s scheduling system over a seven day period in April 2021. The data was then run through our model and a resident compared the output of our model to the ground truth of a radiologists protocol. We evaluate our model’s performance using both top one and top two accuracy. We define the top one accuracy as the percentage of protocols we correctly identify as compared to the ground truth of the radiologist. Similarly, we define top two accuracy in the same way, however, we take our model’s prediction to be correct if the true protocol is found in the top two results from our model. Our radiologist and his team collected 211 observations

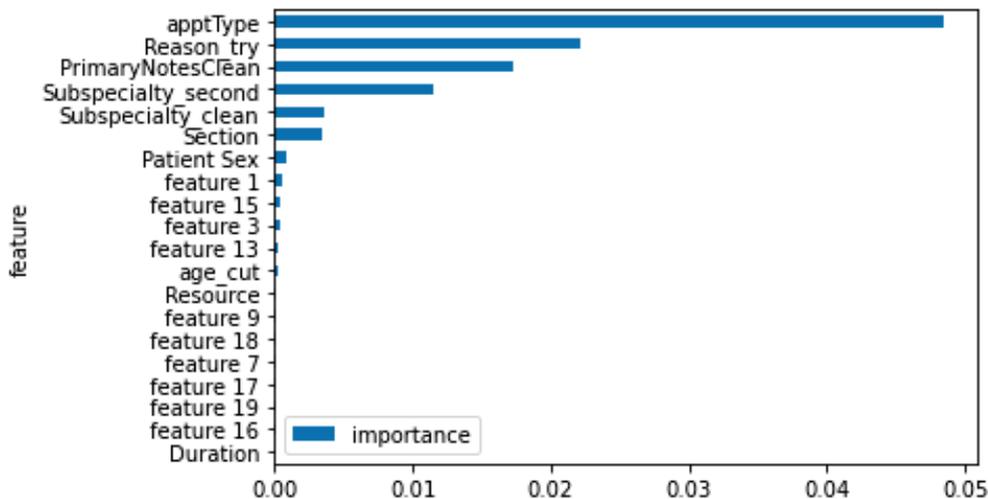


Fig. 4. Feature importance for the large category protocol neural network. We can see that the most important feature is appointment type where the second seems to be reason.

TABLE III
DISTRIBUTION OF PROTOCOLS IN PILOT DEPLOYMENT.

	Number of Records
Abdominal Imaging - CT	34
Chest Imaging - CT	30
Diagnostic Imaging - MR	28
Neuroradiology - CT	33
Neuroradiology - MR	42
Abdominal Imaging - MR	44

TABLE IV
ACCURACY OF MODEL IN PILOT DEPLOYMENT.

	Top One	Top Two
Large Categories of Labels	76%	88%
Abdominal Imaging - CT	73%	91%
Chest Imaging - CT	66%	83%
Diagnostic Imaging - MR	64%	93%
Neuroradiology - CT	45%	52%
Neuroradiology - MR	45%	59%
Abdominal Imaging - MR	82%	89%

from their imaging categories: Abdominal Imaging CT, Chest Imaging - CT, Diagnostic Imaging - MR, Neuroradiology - CT, Neuroradiology - MR, and Abdominal Imaging - MR. The distribution of these observations is as reported in table III. We see strong performance in Abdominal Imaging, Chest Imaging and Diagnostic Imaging, but a large drop in performance in Neuroradiology. We believe this drop in performance is partly due to a small sample size. Additionally, there seems to be a disparity in the distribution of the neuroradiology protocols in the deployment data compared to our test data. Another reason for the drop in performance, is that we encode the reason category based on ICD-10 codes or a mapping to those codes. If there is a reason that is input in an unorthodox manner, it will be put into an ‘other’ category and we will get less information out of it. In the future we would hope to have a stronger mapping of reason to ICD-10 code to get rid of this issue. The accuracy of the large category model and the specific protocol models in the pilot deployment are as seen in table IV.

VI. DISCUSSION

A. A note on Mislabeled Protocols

According to the attending radiologist on our team, many of the mislabeled protocols in our validation set were not harmful differences. For example, if we predicted a liver 1 but the actual protocol as a liver 3b this is not a harmful difference. These two protocols are separate, but both are images of

the liver, 3b is specifically in reference to a liver transplant. Mislabelings similar to this mistake are less worrisome as a physician could easily make this correction. In deployment we see that the only mislabelings for the chest CT section come from predicting chest abdomen pelvis with iv with oral when it is just a chest image. And again, because our goal is assistance rather than automation, this is not a detrimental misclassification.

B. Comparison with other Tabular Deep Learning and ML Methods

Deep learning has made enormous strides in its ability to analyze unstructured data. However, these strides have been noticeably smaller when it comes to the use of deep learning on tabular data sets. Thus, standard machine learning methodologies dominate this area which has led to the push for the creation of many tabular deep learning architectures[17], [19], [20]. We choose to use fast.ai’s tabular data implementation which consists of a feed-forward neural network where our categorical features are passed through a learned embedding and we have a batch norm and dropout between each linear layer[16]. While there are more recent tabular deep learning implementations and more common machine learning techniques, they performed similarly to, or worse than, fast.ai’s tabular implementation on the test set (table V) [21], [22], [23], [24], [25].

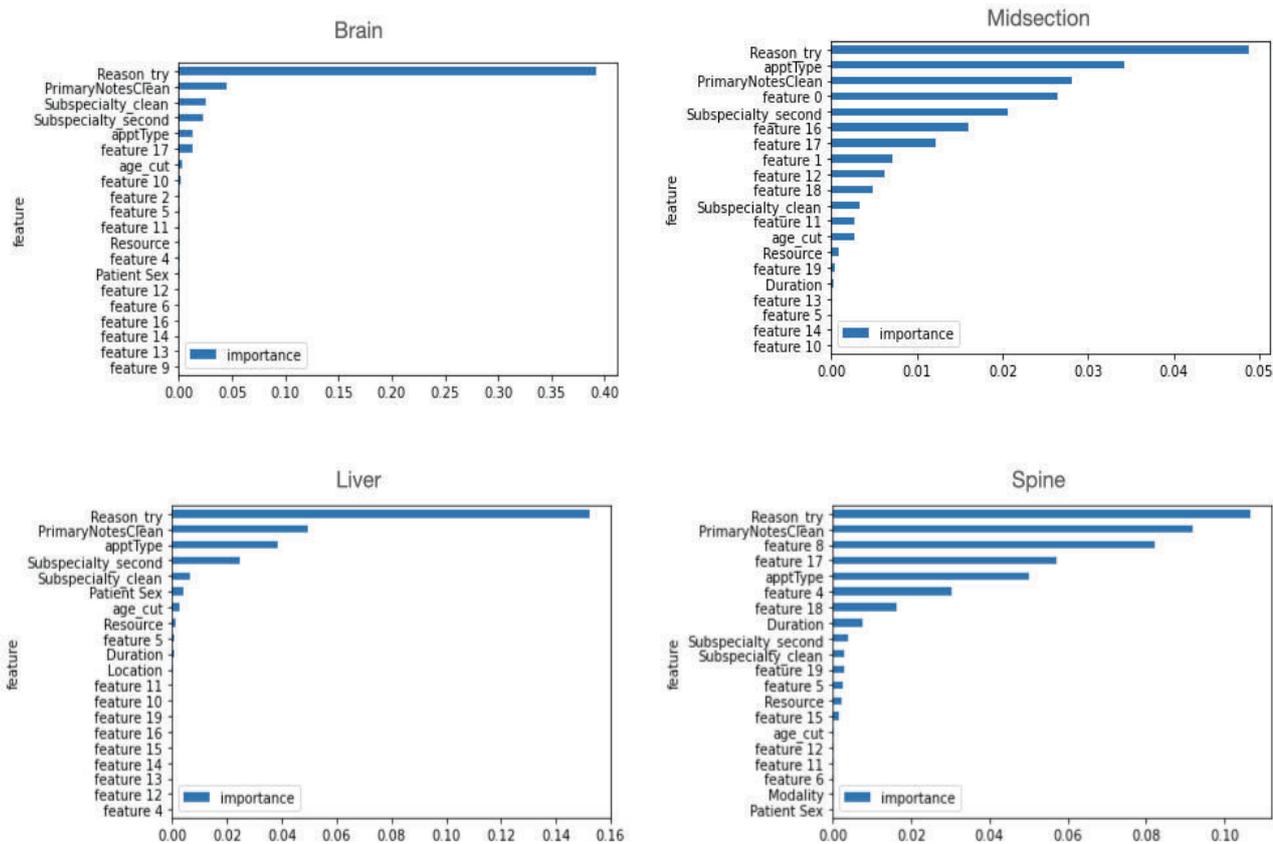


Fig. 5. Feature importance for the specific protocol neural networks. We can see that the most important features are Reason for all of the specific protocol models where the second seems to be primary notes, except for the midsection model where we see appointment type as the second most important. Here, feature 0-19 are the tfidf features as reported in table I

TABLE V

THE PRECISION SCORES ON THE TEST SET FOR OUR LARGE CATEGORY MODEL VERSUS THE STATE OF THE ART ARCHITECTURE AND STANDARD MACHINE LEARNING METHODS. THE SCORES ARE REPORTED AS WEIGHTED AVERAGES.

	Our Model	edRVFL	XGBoost
Test Precision	96%	82%	96%
Test Recall	96%	81%	96%

C. Future Work

We see that our model is able to predict the top 50 common protocols quite well. However and unsurprisingly, there is performance degradation on the less frequently occurring protocols in the data. In the future we would like to include more of these important protocols in an effort to make our model more general. Similarly, our model struggles a bit on protocols with less support. To combat this, we would want

to add more data to our training set to better learn these less common protocols. Furthermore, we would like to add more features to the tfidf vectorizer of the description feature to see if this would allow us to get more useful information out of this feature. Finally, we would also be interested in examining current NLP techniques such as BERT or even a simple convolutional neural network[26].

VII. CONCLUSION

In this work, we examined a model to help radiologists specific protocol exams. We presented a two layer modeling approach composed of a large category protocol neural network followed by multiple specific protocol models to predict the protocol. We found these models were accurate on both the validation and the test set, indicating they would generalize quite well as they achieved a top one precision and recall of 98% on the large categories of protocols. The specific

protocol models report top one precision and recall scores ranging from 84% to 95%. In deployment we see top one accuracy ranging from 64% to 82% on the non-neuroradiology protocols, however we see a drop to 45% on these exams. Similarly, we see top two accuracy ranging from 83% to 93% on the non-neuroradiology protocols, however we see a drop to 52%-59% on these exams. This model has an average top one and top two accuracy respectively of 63% and 77% in its pilot deployment.

This work presents both an algorithm and pilot field test with the aim to have immediately impact on the radiology field. This AI assist tool is already operating at a level of accuracy that can cut down the on administrative tasks for physicians and therefore allow more time spent focusing on tasks such as diagnosing exams and improving the quality of care to all patients.

VIII. ACKNOWLEDGMENTS

We thank all who provided immense feedback on previous versions of this paper and those who helped in the pilot deployment.

REFERENCES

- [1] R. Smith-Bindman, D. L. Miglioretti, and E. B. Larson, "Rising use of diagnostic medical imaging in a large integrated health system," *Health Affairs*, vol. 27(6), p. 1491–1502, 2008.
- [2] H. Health, "Radiation risk from medical imaging," <https://www.health.harvard.edu/cancer/radiation-risk-from-medical-imaging>.
- [3] J. M. Walker, P. Carayon, N. Leveson, J. Paulus, R. A. and Tooker, H. Chin, J. Bothe, and W. F. Stewart, "Ehr safety: the way forward to safe and effective systems," *Journal of the American Medical Informatics Association*, vol. 15(3), p. 272–277, 2008.
- [4] J. Vosshenrich, P. Brantner, J. Cyriac, D. T. Boll, E. M. Merkle, and T. Heye, "Quantifying radiology resident fatigue: Analysis of preliminary reports," *Radiology*, vol. 298(3), p. 632–639, 2021.
- [5] N. Stec, D. Arje, A. R. Moody, E. A. Krupinski, and P. N. Tyrell, "A systematic review of fatigue in radiology: Is it a problem?" *American Journal of Roentgenology (1976)*, vol. 201(4), p. 799–806, 2018.
- [6] C. R. Stehman, Z. Testo, R. S. Gershaw, and A. R. Kellogg, "This article corrects: "burnout, drop out, suicide: Physician loss in emergency medicine, part 1"," *West J Emerg Med*, vol. 20(5), p. 840–841, 2019.
- [7] J. W. Oestmann, R. Greene, D. C. Kushner, L. Bourgouin, P. M. and Linetsky, and H. J. Llewellyn, "Lung lesions: correlation between viewing time and detection," *Radiology*, vol. 166(2), p. 451–453, 1988.
- [8] A. G. Patel, V. J. Pizzitola, C. D. Johnson, N. Zhang, and M. D. Patel, "Radiologists make more errors interpreting off-hours body ct studies during overnight assignments as compared with daytime assignments," *Radiology*, p. 201558, 2020.
- [9] P. Rajpurkar, J. Irvin, R. L. Ball, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. P. Langlotz, B. N. Patel, K. W. Yeom, K. Shpanskaya, F. G. Blankenberg, J. Seekins, T. J. Amrhein, D. A. Mong, S. S. Halabi, E. J. Zucker, and M. P. Lungren, "Deep learning for chest radiograph diagnosis: A retrospective comparison of the chexnext algorithm to practicing radiologists," *PLoS Medicine*, vol. 15(11), p. e1002686–e1002686, 2018.
- [10] M. E., C. M., and C.-C. A. et al, "Deep learning workflow in radiology: a primer," *Insights Imaging 11*, vol. 22, p. 604–610, 2020.
- [11] A. Kalra, A. Chakraborty, B. Fine, and J. Reicher, "Machine learning for automation of radiology protocols for quality and efficiency improvement," *Journal of the American College of Radiology*, vol. 17, no. 9, pp. 1149–1158, 2020, focus on COVID-19: Lessons From SCARD. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S154614402030288X>
- [12] P. López-Úbeda, M. C. Díaz-Galiano, T. Martín-Noguerol, A. Luna, L. A. Ureña-López, and M. T. Martín-Valdivia, "Automatic medical protocol classification using machine learning approaches," *Computer Methods and Programs in Biomedicine*, vol. 200, p. 105939, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0169260721000134>
- [13] B. AD and M. TR, "Using machine learning for sequence-level automated mri protocol selection in neuroradiology," *Computer Methods and Programs in Biomedicine*, vol. 25(5), pp. 568–571, 2021.
- [14] A. Bhat, G. Shih, and R. Zabih, "Automatic selection of radiological protocols using machine learning," in *Proceedings of the 2011 Workshop on Data Mining for Medicine and Healthcare*, ser. DMMH '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 52–55. [Online]. Available: <https://doi.org/10.1145/2023582.2023591>
- [15] L. Y. H., "Efficiency improvement in a busy radiology practice: Determination of musculoskeletal magnetic resonance imaging protocol using deep-learning convolutional neural networks," *Journal of digital imaging*, vol. 31(5), p. 604–610, 2018.
- [16] J. Howard et al., "fastai," <https://github.com/fastai/fastai>, 2018.
- [17] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.
- [18] A. Subbaswamy and S. Saria, "From development to deployment: dataset shift, causality, and shift-stable models in health ai," *Biostatistics*, vol. 21(2), p. 345–352, 2020.
- [19] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.
- [20] F. Evelyn and J. L. Hodges, "Discriminatory analysis. nonparametric discrimination: Consistency properties." *International Statistical Review / Revue Internationale de Statistique*, vol. 57(3), p. 345–352, 1989.
- [21] S. Popov, S. Morozov, and A. Babenko, "Neural oblivious decision ensembles for deep learning on tabular data," in *International Conference on Learning Representations*, 2020. [Online]. Available: <https://openreview.net/forum?id=r1eiu2VtWH>
- [22] G. Ke, J. Zhang, Z. Xu, J. Bian, and T.-Y. Liu, "TabNN: A universal neural network solution for tabular data," 2019. [Online]. Available: <https://openreview.net/forum?id=r1eJssCqY7>
- [23] S. O. Arik and T. Pfister, "Tabnet: Attentive interpretable tabular learning," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 8, pp. 6679–6687, May 2021. [Online]. Available: <https://ojs.aaai.org/index.php/AAAI/article/view/16826>
- [24] Q. Shi, R. Katuwal, P. Suganthan, and M. Tanveer, "Random vector functional link neural network based ensemble deep learning," *Pattern Recognition*, vol. 117, p. 107978, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0031320321001655>
- [25] L. Katzir, G. Elidan, and R. El-Yaniv, "Net-{dnf}: Effective deep modeling of tabular data," in *International Conference on Learning Representations*, 2021. [Online]. Available: <https://openreview.net/forum?id=73WTGs96kko>
- [26] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

The Importance of Fruit Trees for Prescribed Burning in a South American Savanna

Rodrigo M. Falleiro, Joaquim P. L. Parimé, Luciano C. dos Santos, Rodrigo D. da Silva

Abstract—The Cerrado biome is the most biodiverse savanna on the planet. Located in central Brazil, its preservation is seriously threatened by the advance of intensive agriculture and livestock. Conservation Units and Indigenous Lands are increasingly isolated and subject to mega wildfires. Among the characteristics of this savanna, we highlight the high rate of primary biomass production and the reduced occurrence of large grazing animals. In this biome, the predominant fauna is more dependent on the fruits produced by the dicotyledonous species, in relation to other tropical savannas. Fire is a key element in the balance between mono and dicotyledons or between the arboreal and herbaceous strata. Therefore, applying fire regimes that maintain the balance between these strata, without harming fruit production, is essential in the conservation strategies of the Cerrado's biodiversity. Recently, Integrated Fire Management has started to be implemented in Brazilian protected areas. As a result, management with prescribed burns has increasingly replaced strategies based on fire exclusion, which in practice have resulted in large wildfires, with highly negative impacts on fruit and fauna production. In the Indigenous Lands, these fires were carried out respecting traditional knowledge. The indigenous people showed great concern about the effects of fire on fruit plants and important animals. They recommended that the burns be carried out between April and May, as it would result in a greater production of edible fruits ("fruiting burning"). In other tropical savannas in the southern hemisphere, the preferential period tends to be later, in the middle of the dry season, when the grasses are dormant (June to August). However, in the Cerrado, this late period coincides with the flowering and sprouting of several important fruit species. To verify the best burning season, the present work evaluated the effects of fire on flowering and fruit production of the *Byrsonima* sp., *Mouriri pusa*, *Caryocar brasiliense*, *Anacardium occidentale*, *Pouteria ramiflora*, *Hancornia speciosa*, *Byrsonima verbascifolia*, *Anacardium humille* and *Talisia subalbans*. The evaluations were carried out in the field, covering 31 Indigenous Lands that cover 104,241.18 Km², where 3,386 prescribed burns were carried out between 2015 and 2018. The burning periods were divided into early (carried out during the rainy season), modal or "fruiting" (carried out during the transition between seasons) and late (carried out in the middle of the dry season, when the grasses are dormant). The results corroborate the traditional knowledge, demonstrating that the modal burns result in higher rates of reproduction and fruit production. Late burns showed intermediate results, followed by early burns. We conclude that management strategies based mainly on forage production, which are usually applied in savannas populated by grazing ungulates, may not be the best management strategy for South American savannas. The

effects of fire on fruit plants, which have a particular phenological synchronization with the fauna cycle, also need to be observed during the prescription of burns

Keywords—Cerrado Biome, Fire Regimes, Native Fruits, Prescribed Burns.

R. M. Falleiro. Author is with the Brazilian Institute of Environment and Renewable Natural Resources, DF 70818-900 BR (phone: +55-61 982077270; e-mail: falleiorm@gmail.com).

J. P. L. Parimé. S. B. Brazilian Institute of Environment and Renewable Natural Resources, RR. 69304-650 BR (phone: +55-95 9999713381; e-mail: Joaquim.Lima@ibama.gov.br).

L. C. Santos T. C. Brazilian Institute of Environment and Renewable Natural Resources, DF 70818-900 BR (phone: +55-61 996043877; e-mail: carregosasantos621@gmail.com).

R. D. Silva. Brazilian Institute of Environment and Renewable Natural Resources, RS 90050-250 BR (phone: +55-51 982694073; rodrigo.silva@ibama.gov.br).

Agroforestry Systems and Practices and Its Adoption in Kilombero Cluster of Sagcot, Tanzania

¹Nnko, L.E.*, ²Kashaigili, J. J, ¹Monela, G. C. and ³Munishi, P. K. T.

¹Department of Forest and Environmental Economics, Sokoine University of Agriculture, P.O. Box 3011, Chuo Kikuu, Morogoro, Tanzania

²Department of Forest Resources Assessment and Management, Sokoine University of Agriculture, P.O. Box 2013 Chuo Kikuu, Morogoro Tanzania.

³Department of Ecosystems and Conservation, Sokoine University of Agriculture, P.O. Box 3010 Chuo Kikuu, Morogoro Tanzania

*Corresponding author, Email: nnkolazaro@yahoo.com

ABSTRACT

Agroforestry systems and practices are perceived to improve livelihood and sustainable management of natural resources. However, their adoption in various regions differs, with the biophysical conditions and societal characteristics. This Study was conducted in Kilombero District to investigate the factors influencing the adoption of different agroforestry systems and practices in agro-ecosystems and farming systems. A Household survey, Key informant interviews and focus group discussion was used for data collection in three villages. Descriptive statistics and multinomial logistic regression in SPSS were applied for analysis. Results shows that Igima and Ngajengwa villages had home garden practices dominated, as revealed by 63.3% and 66.7%, respectively, while Mbingu village had mixed intercropping practice with 56.67%. Agrosilvopasture systems were dominant in Igima and Ngajengwa villages with 56.7% and 66.7%, respectively, while in Mbingu village, the dominant system was agrosilviculture with 66.7%. The results from multinomial logistic regression show that show that different explanatory variable was statistical significance as predictors of the adoption of agroforestry systems and practices. Residence type and sex were the most dominant factor influencing the adoption of agroforestry systems. Duration of stay in the village, availability of extension education, residence, and sex were the dominant factor influencing the adoption of agroforestry practices. The most important and statistically significant factors among these were residence type and sex. The study concludes that agroforestry will be more successful if the local priorities which include social-economic need characteristics of the society will be considered in designing systems and practices. The socio-economic need of the community should be addressed in the process of expanding the adoption of agroforestry systems and practices.

Keywords: Agroforestry adoption, agroforestry systems, agroforestry practices, agroforestry, Kilombero.

INTRODUCTION

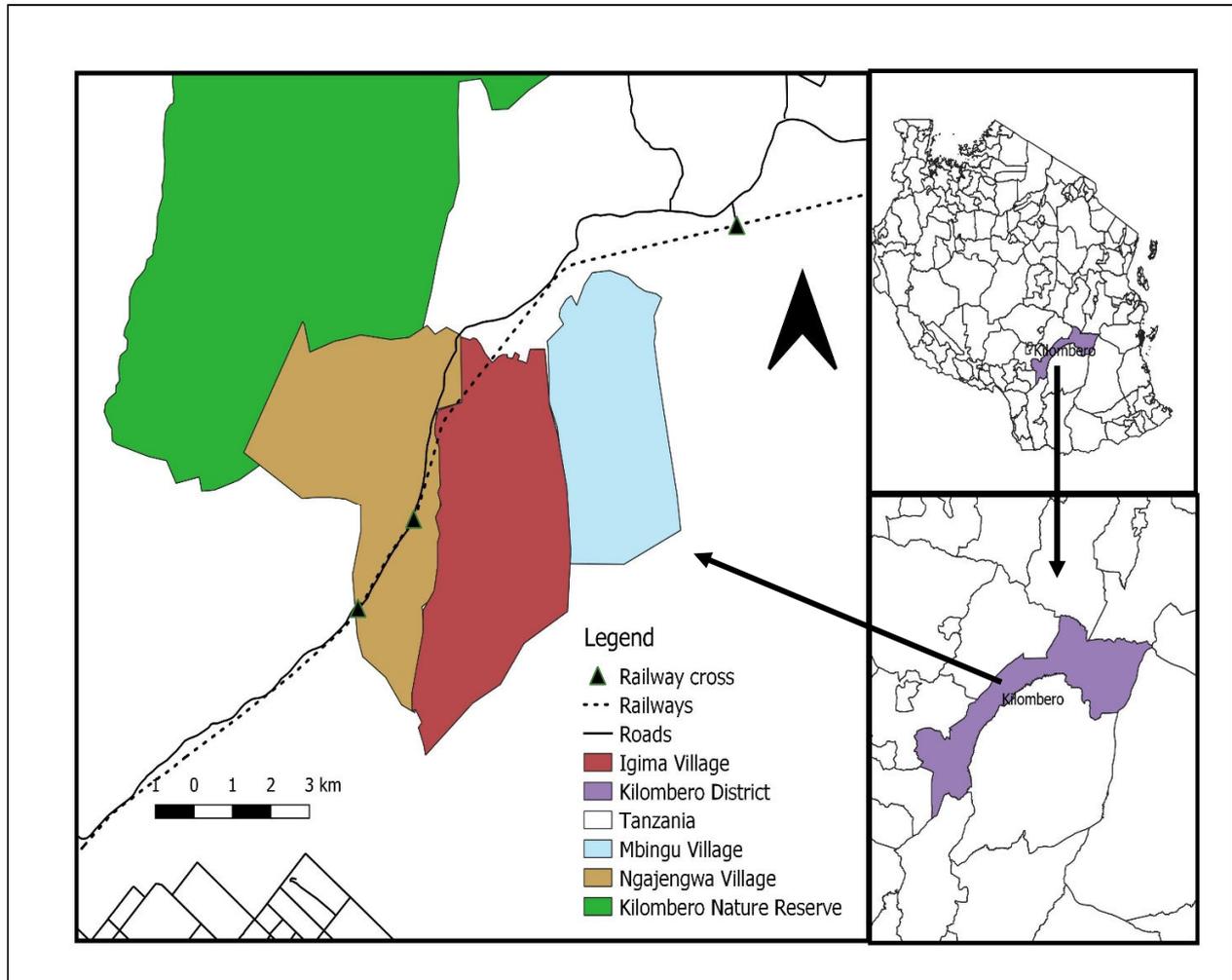
Agroforestry systems and practices have been used interchangeably in various articles. However, Nair (1989) pointed out that agroforestry systems are described based on the components (animals, crops, and trees) present in the agroforestry farmland. Agroforestry practices refer to the arrangement of the components present in the agroforestry systems in time and space. Integration of trees on farmland increases socio-economic and environmental benefits for land uses at all levels (Incrat 2006). Since agroforestry is now emerging as a promising land-use option and climate-smart agriculture, its productivity depends on the systems and practices that involve components available and their arrangements (Syampungani et al. 2010). In order to overcome the reduction of arable land, assure food security, and improve livelihoods, it is important to consider the systems and practices since their productivity and management differ. (Vincent et al. 2012). Sanchez et al. (1997) pointed out that agroforestry systems and practices gave an alternative solution to poor smallholder farmers who would otherwise have a reduction in crop yield.

Eneji et al. (2004) reported that the aim of agroforestry systems and practices is to optimize the positive outcome in order to obtain a diversified and more sustainable production system from limited resources than other systems of land use. However, the potential benefits of improved livelihoods and long-term environmental management will not be realized unless farmers engage in agroforestry on a large scale (Magugu et al. 2018).

Southern Agricultural Growth Corridor of Tanzania (SAGCOT) is a public-private partnership established in 2010 with the objective of improving agricultural productivity, food security, reducing poverty, and ensuring environmental stability in all the areas that the corridor covers (Milder et al. 2012). Kilombero District is among the SAGCOT Clusters where various intervention such as conservation farming through agroforestry has been implemented. Though SAGCOT has been supporting different interventions, including capacity building on agroforestry and conservation farming, to benefit ecosystems, biodiversity, and climate, there have been challenges to the adoption of agroforestry (Msofe et al. 2018). In this context, it is vital to identify the socio-economic factors affecting the adoption of agroforestry systems and practices. Identifying the socio-economic factors will ascertain the opportunities for the development of agroforestry systems and practices. Sinclair and Walker (1999) reported that quantitative and predictive understanding of agroforestry systems and practices enables easy adoption. Developing strategies and encouraging farmers to plant trees on their farmland can be done only if the characteristics of the farmland and farmers in relation to tree growing exist (Irshand et al. 2011). This study had two specific objectives, to identify the agroforestry systems and practices in the study area and to determine the factors influencing the adoption of agroforestry systems and practices.

MATERIALS AND METHODS

Description of the study area



The study was conducted in Kilombero District, which is located in the Morogoro Region between latitudes of $8^{\circ}15'0''$ South and longitudes $36^{\circ}25'0''$ East, with elevation ranging from 262 m to 550 m above mean sea level. Administratively, the Kilombero District has five divisions, 19 wards, and 46 villages. The district is bordered by Kilosa District in the north, the south east by Ulanga District, the south west by the Iringa Region and in the west by the Lindi Region (URT 2007). According to the 2012 census, the population of Kilombero was 407 880, with 202 789 males and 205 091 females (URT 2013). This area is currently experiencing a doubling of the human population over the years (Madulu, 2004). The large migration of farmers due to fertile land and livestock keepers due to the presence of animal fodder is the primary cause of population growth.

The climate in the study area is marked by wet and dry seasons, which are further categorized into four sub-seasons: the hot wet season from December to March, the cool wet season from April to June, the cool dry season from July to August, and the hot dry season from September to November. The area receives between 1200 and 1800 mm of rainfall per year, and temperatures range from 26°C to 32°C (Balama et al. 2016). Generally, land use is categorized as village land, reserved land, and general land as defined in the Village Land Act 1999 (URT 2017).

The main economic activities in the area include cash crop cultivation, food crop cultivation, petty trading, and fishing in the Kilombero River (URT 2007). Overall, cereals from the coast, such as rice, millet, and maize, are widely grown. Also, vegetables such as sweet potatoes, yams, ground-nuts, melons, pumpkins, cucumbers, and many other excellent food crops are grown. Tobacco is grown abundantly, sugar-cane, the castor oil plant, cocoa, and cotton are also cultivated (Bergius et al. 2020).

Methods

Sampling procedure

Three villages were purposely selected due to the presence of agroforestry farmers. A random sampling procedure was adopted for selecting households with agroforestry systems and practices. Village registers were used as a sampling frame. The sampling unit for this study was the individuals chosen from the population as respondents to represent others and the information obtained was used to describe the characteristics of the entire population (Bryman 2004, Nkonoki 2015).

Sample size determination

Nachimias and Nachimias (1996) pointed out that sample size is the most important determinant of any survey estimates. Studies by Bailey (1998), Saunders *et al.* (2007), Mbeyale (2009) and Mtongani *et al.* (2014), indicated that a sample of 30 units is sufficient, irrespectively of the population size for field work data collection and analysis. Therefore, a total of 90 respondents were sampled from the three villages (Igima, Mbingu and Ngajengwa) for interview. The sample size was considered sufficient to generate the statistical inferences required for making study conclusions.

Data collection

A house hold survey, a focus group discussion, and a key informant interview were used to collect data. The Questionnaire was the main instrument for collecting data from the agroforestry farmers on agroforestry practices and determinant factors for adoption. Focus group discussion enabled to get an insight on status of agroforestry in the villages. Key informant interviews enabled us to get clarification on particular issues raised during focus group discussion and household interviews. Key informants involved were a forest officer from Kilombero Nature Reserve, District Forest Officer and a Ward agriculture extension officer.

Data analysis

Information from focus group discussions and key informant interviews was analyzed using content analysis, whereby raw data was broken down to generate meaningful units of information. Information from the household survey was coded and assigned variables in the Statistical Package for Social Science (SPSS). Descriptive statistics such as percentages and frequencies were used to profile the agroforestry systems and practices of the respondents. In addition, multinomial logistic regression was used to identify factors determining the choice of

adopting agroforestry systems and practices. This model is suitable for determining adoption when the dependent variables have more than two categories (Ojo *et al.* 2013, Obadim *et al.* 2020).

Let π_j denotes the multinomial of an observation falling in the j^{th} category to find the relationship between the probability and the p - explanatory variables

$$X_1 + X_2 + X_3 \dots \dots \dots X_p \dots \dots \dots \text{Equation 1}$$

The multiple logistic regression is given by:

$$\log \left[\frac{\pi_j(x_1)}{\pi_k(x_1)} \right] = a_{oi} + \beta_{1j}x_{1j} + \beta_{2j}x_{2j} + \dots \dots \dots \beta_{pj}x_{pj} \dots \dots \dots \text{Equation 2}$$

Where $J= 1, 2, \dots, (k - 1), I = 1, 2, \dots, p$

K stands for number of response or dependent categories where for this study dependent categories for agroforestry systems are agrosilvopasture, agrosilviculture and silvopasture. For agroforestry practices dependent categories are home garden, mixed intercropping, parkland and boundary.

P = Number of explanatory variables included in the model.

When estimating the model, the coefficient of the reference group is normalized to zero (Rahij and Fakayode 2009, Ojo *et al.* 2013). This is because the probability of the choice must sum up to unity. Hence, for the choice of three categories (agrosilvopasture, agrosilviculture, and silvopasture), only two sets of parameters were identified and estimated. For four categories (home garden, mixed intercropping, parkland, and boundary), only three distinct sets of parameters were identified and estimated. In this study, the reference category for agroforestry systems was agrosilvopasture, and for agroforestry practices, the reference category was home garden. The natural logarithm for the odd ratio for equation 1 and 2 gives the estimation equation below.

$$\log \pi_j(x_i) = \frac{\exp(a_{oi} + \beta_{1j}x_{1i} + \beta_{2j}x_{2i} + \dots \dots \dots + \beta_{pj}x_{pi})}{1 + \sum_{j=1}^{k-1} \exp(a_{oi} + \beta_{1j}x_{1i} + \beta_{2j}x_{2i} + \dots \dots \dots + \beta_{pj}x_{pi})} \dots \dots \dots \text{Equation 3}$$

$J= 1, 2, \dots, K - 1$ the model parameter is estimated by the method of multinomial logit.

The independent variable included

X_1 Farming experience (Years), X_2 House hold income (Tanzania shillings), X_3 Duration of stay in the village (years), X_4 Residence type (1 native 2 other wise), X_5 Education level, X_6 Extension services, X_7 House hold size, X_8 Sex (1 male 2 other wise). These independent

variables were selected on basis of other agroforestry adoption studies conducted in different tropical counties (Rahij and Fakayode 2009, Ojo *et al.* 2013 and Obadimu *et al.* 2020).

RESULTS

Agroforestry systems and practices in Kilombero

A total of three agroforestry systems were found in this study which are agrosilvopasture, agrosilviculture and silvopasture systems. The agrosilvopasture system was highly practiced in Igima and Ngajengwa villages, with 56.6% and 67.7%. In Mbingu village, the most dominant system was the agrosilviculture system with 66.7% (Figure 2). Four agroforestry practices were identified in the study area. The practices found were home garden, parkland, boundary and mixed intercropping. Home garden was highly practiced in Igima and Ngajengwa with 63.3% and 66.7% respectively, while mixed intercropping was highly practiced in Mbingu villages with 56.7% (Figure 3).

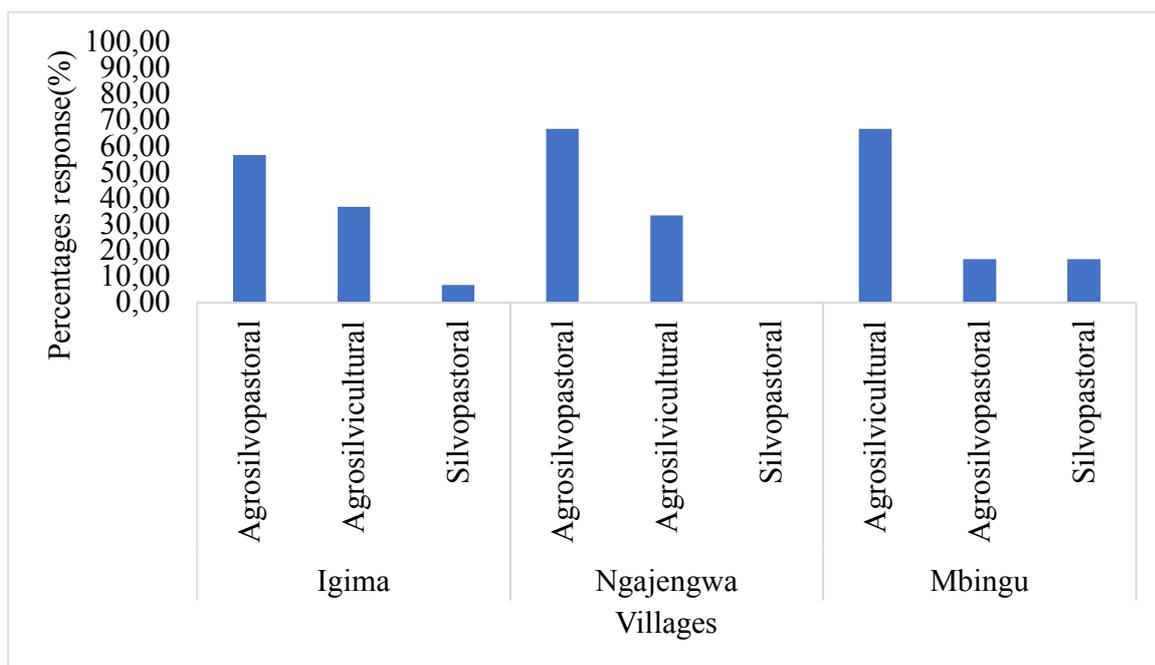


Figure 1: Agroforestry systems in Kilombero District

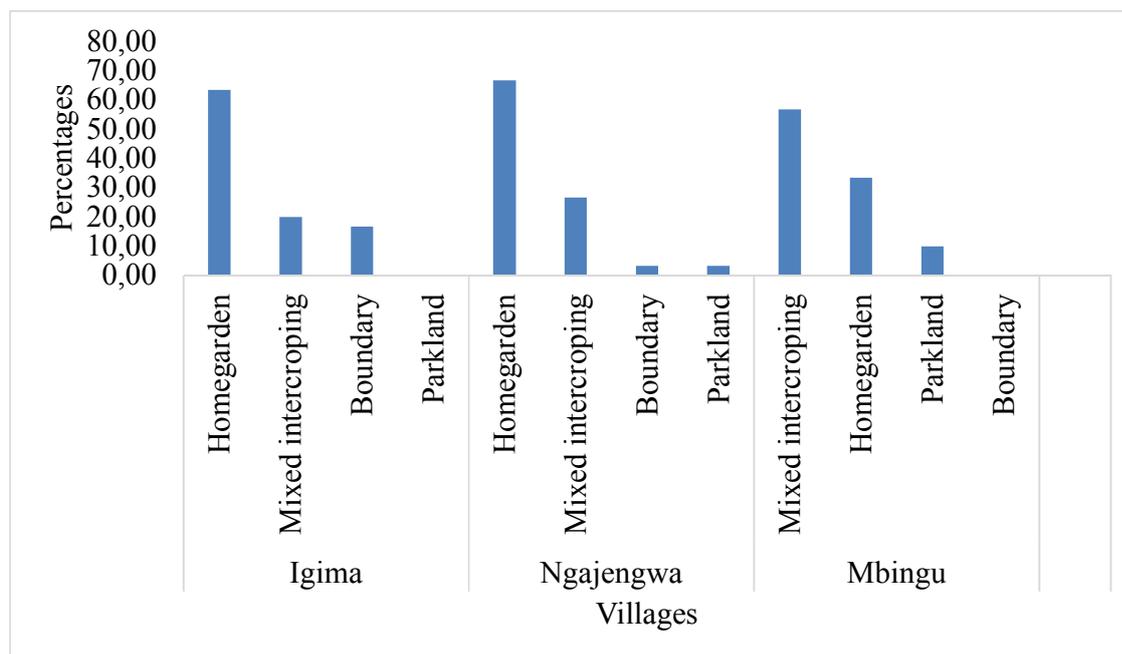


Figure 2: Agroforestry practices in Kilombero

Adoption of agroforestry systems

Multinomial logistic regression analysis indicated that male and native residences had a statistically significant difference in the adoption of agrosilviculture with reference to agrosilvopasture ($P < 0.05$). Native residences negatively influenced the adoption of agrosilviculture with reference to agrosilvopasture, and the rate of adoption decreased at an odd ratio of 0.14. On the other hand, males were found less likely to adopt agrosilviculture with reference to agrosilvopasture, at an odd ratio of 0.344 (Table 1). Furthermore, other variables had no statistic significant difference in the adoption of silvopasture with reference to agrosilvopasture. Table 2 is illustrative.

Table 1: Determinant of adoption of agrosilviculture with reference to agrosilvopasture

Parameter Estimates	B	Std. Error	Sig.	Exp(B)
Agrosilviculture system				
Farming experience	-0.048	0.03	0.115	0.953
Household income	0	0	0.313	1
Time to stay	0.05	0.026	0.059	1.051
Household size	-0.066	0.117	0.574	0.937
Education	-0.108	0.256	0.675	0.898
[Extension services=Yes]	0.84	0.563	0.136	2.316
Residence (Native)	-1.963	0.729	0.007*	0.14

Sex (Male)	-1.068	0.528	0.043*	0.344
------------	--------	-------	--------	-------

a The reference category is: Agrosilvopastore. * P<0.05

Table 2: Determinant of adoption of silvopasture with reference to agrosilvopasture

Parameter Estimates	B	Std. Error	Sig.	Exp(B)
Silvopastural systems				
Farming experience	-0.009	0.048	0.843	0.991
Household income	0	0	0.965	1
Time to stay	0.074	0.043	0.083	1.077
Household size	0.288	0.289	0.319	1.333
Education	0.692	0.405	0.087	1.998
Extension services (Yes)	-0.546	1.155	0.636	0.579
Residence (Native)	-1.364	1.214	0.261	0.256
Sex (Male)	-0.521	1.031	0.614	0.594

a The reference category is: Agrosilvopastoral. * P<0.05

Adoption of agroforestry practices

On the other hand, results from multinomial logistic regression show that the duration of stay in the village, residence type, extension education, and sex ($P < 0.05$) were the factors that influenced the adoption of agroforestry practices with a statistically significant difference. A unit increase in the duration of stay in a village leads to increased adoption of mixed intercropping at an odd ratio of 1.064 (see table 3). Furthermore, the availability of extension education, information, and awareness to farmers increased the adoption of agroforestry practices at an odd ratio of 4.052 (see table 3). Native residences less likely to influenced the adoption of mixed intercropping with reference to home gardens. This indicates that native residences are likely to adopt home garden practices at an odd ratio of 0.068 (see table 3). Furthermore, males were less likely found to adopt mixed intercropping with reference to home gardens at an odd ratio of 0.167. (See Table 3). Boundary and parkland agroforestry practices were influenced by explanatory variables, but there was no statistically significant difference with reference to the home garden (see Table 4 and Table 5).

Table 3: Determinant of adoption of mixed intercropping with reference to home garden practices

Parameter Estimates	B	Std. Error	Sig.	Exp(B)
Mixed intercropping				
Farming experience	-0.048	0.036	0.175	0.953
Household income	0	0	0.303	1
Time to stay	0.062	0.031	0.041*	1.064
Household size	-0.139	0.14	0.32	0.87
Education	-0.44	0.301	0.144	0.644

Extension services (Yes)	1.399	0.675	0.038*	4.052
Residence (Native)	-2.684	0.908	0.003*	0.068
Sex (Male)	-1.79	0.65	0.006*	0.167

a The reference category is: Home garden. * P<0.05

Table 4: Determinant of adoption of boundary with reference to home garden practices

Parameter Estimates				
Boundary practice	B	Std. Error	Sig.	Exp(B)
Farming experience	-0.064	0.053	0.229	0.938
Household income	0	0	0.125	1
Time to stay	-0.086	0.069	0.213	0.917
Household size	-0.123	0.376	0.744	0.884
Education	0.474	0.655	0.469	1.607
Extension services (Yes)	0.715	1.388	0.606	2.045
Residence (Native)	2.951	2.495	0.237	19.134
Sex (Male)	3.004	2.128	0.158	20.164

a The reference category is: Home garden. * P<0.05

Table 5: Determinant of adoption of parkland with reference to home garden practices

Parameter Estimates				
Parkland practice	B	Std. Error	Sig.	Exp(B)
Farming experience	-0.18	0.138	0.192	0.835
Household income	0	0	0.077	1
Time to stay	0.141	0.093	0.13	1.151
Household size	0.018	0.256	0.944	1.018
Education	-18.36	0	.	1.06E-08
Extension services (Yes)	3.027	1.759	0.085	20.64
Residence (Native)	-24.553	0	.	2.18E-11
Sex (Male)	-0.722	1.7	0.671	0.486

a The reference category is: Home garden. * P<0.05

DISCUSSION

Agroforestry systems and practices in Kilombero

Through observation and interviews, two villages were found with more home garden practices and an agrosilvopastoral system, leading to the conclusion that the components involved were trees, herbaceous crops, and/or animals at a high percentage. During interviews, respondents said that the availability of animal fodder and small grazing areas has attracted many farmers to engage in home garden practices. Furthermore, home garden practices and agrosilvopasture were highly practiced because they produce supplementary staple crops and also serve as a source of income for many families. The home garden provides a diversity of crops and livestock, which enables the year-round production of different products and reduces production risk. A study conducted by Maroyi (2009) shows that goods obtained from home garden practices are all consumed at home and cannot be sold to other relatives, so they are offered for free to strengthen the relationships in the village and can only be sold when the household has a surplus. Farmers also urge that agrosilvopasture systems optimize the production per unit area while ensuring a sustained yield over time. Growing trees on farms while integrating with livestock also helps to increase income, produce more food, resulting in food security, and protect the environment, according to a key informant. A study by Saleh (2016) indicates that agrosilvopasture contributes significantly to soil improvement through the supply of green manure to the soil. 66.73% of tree species used in the agrosilvopasture system were *Theobroma cacao*, *Mangifera indica*, and *Tectona grandis*, which provide animal fodder, shade, fruits, timber, and are used as a source of food for households. On the other hand, 73.17% of the tree species encountered in home gardens were similar to those found in agrosilvopasture.

In Mbingu village, mixed intercropping was the most dominant agroforestry practice, and this was due to suitable land for cocoa growing as well as a lack of reserved area for grazing compared to the other two villages. Also, in the focus group discussion, respondents revealed that mixed intercropping provides diversification of crops, especially cash crops like cocoa, which in turn provides income to households. Also, during focus group discussion, respondents in Mbingu village revealed that the presence of Cocoa Kamili Company in the village has influenced mixed intercropping and agrosilviculture of cocoa with trees and other crops. Since mixed intercropping dominates in Mbingu Village, the components included were trees and herbaceous crops, which are termed agrosilviculture. A study conducted by Antriyandarti et al. (2013) indicated that agrosilviculture dominates the land suitable for vegetation. Most of the trees intercropped were *Theobroma cacao* and fruit trees such as *Cocos nucifera* and *Mangifera indica*. Robiglio et al. (2013) pointed out that the integration of cocoa trees, other trees, and food crops has been an easily manageable strategy because it is easier to manage cocoa and tree species. Another survey by Sonwa (2004) found that farmers in Southern Cameroon usually use fruit plants to diversify the cocoa plantation.

Silvopasture was observed by a few farmers, and it is not commonly practiced by farmers in the study area. During focus group discussion, it was revealed that silvopasture is difficult to implement due to a shortage of grazing land and climate variability. Other respondents pointed out that operation costs are very high and there is a poor market structure for livestock products. Caradona et al. (2014) pointed out that livestock production depends on climate factors; therefore, changes in the climate have an enormous impact on production.

Adoption of agroforestry systems in Kilombero.

The results indicated that there was a negative correlation between native residences and the adoption of agrosilviculture systems, with reference to agrosilvopasture systems, indicating that native residences were more likely to adopt agrosilvopasture than agrosilviculture. During the survey, respondents pointed out that agrosilvopasture is better than agrosilviculture as it allows the diversification of all three components. In addition, during focus group discussion, respondents pointed out that non-natives are less likely to engage in agrosilviculture systems because many of them spend a short time and then move to an area. A study by Obeng and Weber (2014) reported that non-native farmers are less likely to adopt agrosilvopasture due to their shorter horizons. Gender was another factor which showed the statistical significance difference in the adoption of agrosilviculture with reference to agrosilvopasture. In this study, males were less likely to adopt agrosilviculture, with reference to agrosilvopasture. Respondents pointed out that with the presence of trees and crop integration, livestock is also important, especially for income contribution when there is crop failure. Similar results have been observed in Malawi by Thangata and Alavalapati (2003) and in Kenya by Sanchez (2002), who indicated that female farm headed houses did not adapt to the agroforestry system compared to male farm headed houses because most males prefer trees as the long-term major source of income. Also, during the focus group discussion, a few spouses were available to respond on behalf of the rest. The woman pointed out that men are always the heads of households; as a result, they make decisions on household affairs, such as controlling resource allocation and general land use management. According to Oino and Mugure (2013), male land ownership has put them at the forefront of decision making on land use systems such as the type of agroforestry system to be practiced for the benefit of households. On the other hand, during focus group discussion, males pointed out that females are always involved during the planning of land use, but they cannot change the last decision made by males, and poultry were found to belong to females and livestock to males. Similar results have been observed by Merce (2004), who indicates that women are more involved practically in agroforestry systems than men, but they cannot make final decisions on the utilization of the land and agroforestry products.

Adoption of agroforestry practices

The findings revealed that native residences had a negative correlation with the adoption of mixed intercropping with reference to the home garden. This indicates that native residences are more likely to adopt home gardens compared to mixed intercropping. A study by Irshad et al. (2011) pointed out that native residences have a high chance of succeeding in the implementation of home garden practices as it takes time to establish a permanent settlement. This result is similar to a study by Magugu et al. (2015) who pointed out that native residences are in a good position to attain land tenure and secure enough land for agroforestry since it is a long-term investment.

Duration of stay in the village had a positive correlation with the adoption of mixed intercropping with reference to the home garden. This indicates that as the duration of stay in the village increases, farmers become more interested in adopting mixed intercropping than home gardens. This result is similar to a study by Liniger et al. (2011) who pointed out that the duration of stay influences crop diversification, hence shifting from practicing agroforestry near the home to the farm land. Farmers with access to extension education were likely to adopt

mixed intercropping over the home garden. During the survey, farmers practicing mixed intercropping argued that in their home gardens there was no proper arrangement of crops and trees, which made them less productive. Similarly, a study by Chija (2013) indicates that extension education is the most critical factor that enhances farmers' adoption of particular agroforestry practices in consideration of the product and production. On the other hand, males were found less likely to adopt mixed intercropping with reference to their home garden. During the survey, it was noted that the majority of the households were headed by males, and one of the roles of males is to ensure food security. In that respect, men preferred home gardens because most of the components in the home garden supplement the household's food and income. In the study area, fruit trees were the primary source of food, especially during drought. The fruit trees found in the study area were *Mangifera indica*, *Persea americana*, and *Cocos nucifera*. A similar study on the home garden by Mengistu (2008) confirmed that fruit trees in the home garden have a significant role in the family during environmental crises such as drought.

CONCLUSION

From this study, it is revealed that potential exists in agroforestry systems and practices in terms of product diversification and biodiversity conservation. However, systems and practices were practiced at different levels in the study area. This indicates that farmers have different preferences in establishing agroforestry systems and practices for different purposes, but the major purposes observed were for domestic use. In order to achieve the development of agroforestry systems and practices, smallholder farmers must develop intensive management to yield a quality product from components of priority to meet market demand. Farmers' oriented factors are critical in the adoption of agroforestry systems and practices. Though the model did not provide statistically significant support to accept the influence of most of the explanatory variables on farmers' decisions to adopt agroforestry systems and practices. But the key factors discussed provide an empirical overview of the factors that should be given attention in the adoption of agroforestry practices and systems. Therefore, the study recommends that all the independent variables outlined must be addressed in order to expand adoption of different systems and practices. This study recommends that different stakeholders be involved in implementing the agroforest projects, and these findings should be disseminated to enhance adoption. Furthermore, these findings may be replicated in other parts of tropical countries to improve the adoption of agroforestry systems and practices.

REFERENCES

- Antriyandarti, E., Ferichani, M. and Ani, S. W., 2013. Sustainability of post-eruption socio-economic recovery for the community on Mount Merapi slope through horticulture agribusiness region development in Boyolali District. *Procedia Environmental Sciences* 17:46 - 52.
- Bailey, D. K., 1998. *Method of Social Research*. The free press Collier-Macmillan Publisher, London. 478 p.
- Balama, C., Augustino, S. and Makonda, F. B. S., 2016. Forestry adjacent household voices on their perceptions and adaptation strategies to climate change in Kilombero District Tanzania. *Springer Plus* 5(792): 1 - 21.

- Bergius, M., Benjaminsen, T., Maganga, F. and Buhaug, H., 2020. Green economy, degradation narratives, and land-use conflicts in Tanzania. *World Development* 129: 104 - 850.
- Bryman, A., 2004. *Social Research Methods*. Oxford University Press, Oxford. 591p.
- Caradona, C. A., Ramirez, N. J., Morales, T. M. and Sanchez, J. S., 2014. Contribution of intensive silvipastoral systems to animal performance and adaptation and mitigation of climate change. *Revista colombiana de ciencias pecurias* 27:76 - 94.
- Chija, M., 2013. Adoption status and management of agroforestry systems and practices by communities. A case study of Kasulu district, Kigoma, Tanzania Unpublished Dissertation for Award of Doctor of Philosophy at Sokoine University of Agriculture, Morogoro. Tanzania. pp 110 – 120.
- Eneji, A.E., Irshad, M. and Inanaga, S., 2004. Agroforestry as a tool for combating soil and environmental degradation. *Sand Dune Res* 51: 47-56.
- ICRAF., 2006. International Centre for Research in Agroforestry. World Agroforestry Centre, Southeast Asia web site. www.worldagroforestrycentre.org/sea. Date accessed: 03 June 2011.
- Irshad, M., Khan. A., Inoue, M., Ashraf, M. and Sher, H., 2011. Identifying factors affecting agroforestry system in Swat, Pakistan. *African Journal of Agricultural Research* 6 (11): 2586 - 2593.
- Jeffrey C Milder, Buck, L. E., Hart, A. K., & Scherr, S. J. (2012). Green Growth Investment Framework for SAGCOT. The SAGCOT Green print: Dar es salaam, Tanzania.76p
- Liniger, H., Studer, R.M., Hauert, C. and Gurtner, M., 2011. Sustainable land management in practice Guidelines and best practices for Sub-Saharan Africa. Food and Agriculture Organization of the United Nations (FAO). 240p
- Madulu, N. F., 2004. Assessment of linkages between population dynamics and environmental change in Tanzania. *African Journal of Environmental Assessment and Management* 9: 88 - 102.
- Magugu, J. W., Feng, S., Huang, Q. and Ototo, O. G., 2018. Socio-economic factors affecting agro-forestry practices adoption in Nyando, Kenya. *Journal of Water and Land Development* 39 :83 - 91.
- Maroyi, A., 2009. Traditional home gardens and rural livelihoods in Nhema, Zimbabwe: a sustainable agroforestry system. *International Journal of Sustainable Development and World Ecology* 16(1): 1 - 8.
- Mbeyale, G. E., 2009. The impact of institutional changes on the management of common pool resources in Pangani River Basin. A case study of Eastern Same Kilimanjaro, Tanzania. Dissertation for award of Doctor of Philosophy at University of Dar es Salaam. Tanzania. 307 p.
- Mengistu, F., 2008. Fruit tree species in the wild and in home garden agroforestry: species Composition, Diversity and Utilization in Western Amhara Region, Ethiopia. Unpublished Dissertation for Award of Doctor of Philosophy at Vienna University. Vienna. 100 p.
- Mercer, D., 2004. Adoption of agroforestry innovations in the tropics. *Agroforestry Systems* 61:311 - 328.
- Msofe, N. K. Sheng, L and Lyimo, J.,2018. Land Use Change Trends and Their Driving Forces in the Kilombero Valley Floodplain, Southeastern Tanzania. *Sustainability*.11: (505) 1-25

- Mtongani, W. A., Munishi, P. K. T., More, S. R. and Kashaigili, J. J., 2014. Local knowledge on the influence of land use land cover changes and conservation threats on avian community in the Kilombero Wetland. *Open Journal of Ecology* 4:723 - 731
- Nachimias, C. F. and Nachimias, D., 1996. *Research methods in the social science* 5th Edition. St Martin's Press, New York.38p.
- Nair, P. K. R., 1989. *Agroforestry systems in the tropics*. Kluwer. Academic, Dordrecht. Netherlands. 664p.
- Nkonoki, J. B., 2015. Effects of institutional changes on forest condition: a case of Chenene forest reserve in Bahi district, Tanzania. Unpublished thesis submitted in fulfilment of the requirements for the degree of doctor of philosophy in forestry of the Sokoine University of Agriculture. Morogoro. Tanzania. pp 66 - 67.
- Obadimu, O. O., Oke, O. S., Asunlegan, O. A., Alaje, M. A., Ojo, D. and Olanrewaju, C. M., 2020. Determinant of Agroforestry Practices among Small Holder Farmers in Oyo State. *Nigeria Journal of Applied Science and Environmental Management* 24 (12): 2107 - 2111
- Obeng, E. A. and Weber, M., 2014. Socio-economic factors affecting agroforestry adoption in Ghana. *Ghana Journal of Forestry* 30 (1): 43 - 60.
- Oino, P. and Mugure, A., 2013. Farmer-oriented factors that influence adoption of agroforestry practices in Kenya. Experiences from Nambale District, Busia County *International Journal of Science and Research* 2(4): 450 - 456
- Ojo, M. A., Nmadu, J. N., Tanko, L. and Olaleye, R. S., 2013. Multinomial Logit Analysis of Factors Affecting the Choice of Enterprise among Small-holder Yam and Cassava Farmers in Niger State. *Nigeria journal of agriculture science* 4(1): 7 - 12.
- Rahji, M. A. Y. and Fakayode, S. A., 2009. A Multinomial Logit analysis of Agricultural Credit Rationing by Commercial Banks in Nigeria. *International Research Journal of Finance and Economics*. 24(1): 90 - 100
- Robiglio, V., Guillaume, L. and Paolo, C., 2013. From farmers to loggers: the role of shifting cultivation landscapes in timber production in Cameroon. *Small Scale Forest* 12(1):67 - 68
- Saleh, M. K., 2016. Agrosilviculture as a strategy for environmental conservation and sustainable peace in Sudano Sahelian zone of Northern Nigeria. *Nigeria Journal of Rural Sociology* 16(4):14 - 21.
- Sanchez, P. A., 2002. Soil Fertility and Hunger in Africa. *Science* 295 (5562): 2019 - 2020.
- Sanchez, P. A., Shepherd, K. D., Soule, M. J., Place, F. M., Buresh, R. J., Izac M. N., Mokwunye A. U., Kwesiga F. R., Ndiritu C. G. and Woome P. L., 1997. Soil fertility replenishment in Africa: An investment in natural resource capital. In: Buresh, R. J. Sanchez, P. A. and Calhoun, F. (Editors). *Replenishing soil fertility in Africa*. Special Publication. pp 1- 46.
- Saunders, M., Lewis, P. and Thornhill, A., 2007. *Research Methods for Business Students*. 4th Edition, FT Prentice Hall, Harlow, 624 p.
- Sonwa, D. J., 2004. Biomass management and diversification within cocoa agroforest in the humid forest zone of southern Cameroon. Unpublished Dissertation for award of Doctor of Philosophy at University of Bonn. Bonn. Germany. 112 p
- Syampunani, S., Chirwa, P. W., Akinnifest, F. K. and Ajayi, O. C., 2010. The potential of using agroforestry as a win-win solution to climate change mitigation and adaptation and meeting food security challenges in Southern Africa. *Agricultural Journal* 5:80 - 88.

- Thangata, P. and Alavalapati, J., 2003. Agroforestry adoption in southern Malawi: the case of mixed intercropping of *Gliricidia sepium* and maize. *Agricultural Systems* 78: (1):57 - 71
- United Republic of Tanzania (URT)., 2007. National sample census Security. Food Development, Livestock Office. Presidents Government, Local. Dar es salaam. Tanzania. 322 p.
- United Republic of Tanzania (URT).,2013. Population and Housing census. National Bureau of Statistics Ministry of Finance, Dar es Salaam. 244 p.
- United Republic of Tanzania (URT).,2017. Ramsar Advisory Mission Report Kilombero valley. Government print. Dar es Salaam. Tanzania. 77 p.
- Vincent, B. S. and Mosango, D. M., 2012. Adoption of Agroforestry Systems by Farmers in Masaka District of Uganda. *Ethnobotany Research and Applications* 10:59 – 68