



Annual Review of CyberTherapy and Telemedicine

Understanding Artificial Intelligence:
Multidisciplinary Insights and the Role
of Cyberpsychology

Editors:

Brenda K. Wiederhold, Ph.D., MBA, BCB, BCN

Thomas Parsons, Ph.D.

Giuseppe Riva, Ph.D., M.S., M.A.

ISSN: 1554-8716

22

<http://www.arctt.info>



**ANNUAL REVIEW OF CYBERTHERAPY
AND TELEMEDICINE 2024**

Annual Review of Cybertherapy And Telemedicine 2024

Understanding Artificial Intelligence:
Multidisciplinary Insights and the Role of Cyberpsychology

Edited by

Brenda K. Wiederhold

*Interactive Media Institute, San Diego, CA, USA
Virtual Reality Medical Institute, Brussels, Belgium*

Thomas Parsons

Arizona State University, Phoenix, AZ, USA

Giuseppe Riva

*Catholic University of Milano, Milan, Italy
Istituto Auxologico Italiano, I.R.C.C.S., Milan, Italy*

Annual Review of CyberTherapy and Telemedicine, Volume 22

Annual Review of CyberTherapy and Telemedicine

Copyright © 2024 Interactive Media Institute
6540 Lusk Boulevard, Suite C115
San Diego, CA 92121

ISBN: 1554-8716

All rights reserved.

Printed in the United States of America

Journal Web site: <http://www.arcctt.info>

Interactive media Institute Website: <http://www.interactivemediainstitute.com>

LEGAL NOTICE

The publisher is not responsible for the use which might be made of the following information

Editors-in-Chief

Brenda K. Wiederhold, Ph.D., BCB, BCN
Interactive Media Institute
Virtual Reality Medical Center

Giuseppe Riva, Ph.D., M.S., M.A.
Istituto Auxologico Italiano
Università Cattolica del Sacro Cuore

Managing Editors

Elena Sajno, Ph.D. Candidate
Università di Pisa
Università Cattolica del Sacro Cuore

Stefano De Gaspari, Ph.D. Candidate
Università di Pisa
Università Cattolica del Sacro Cuore

Chiara Pupillo, Ph.D. Candidate
Università di Pisa
Università Cattolica del Sacro Cuore

Maria Sansoni, Ph.D. Candidate
Università Cattolica del Sacro Cuore

Local Chair Conference Coordinator

Thomas Parsons
Arizona State University

Editorial Board

Nathan Appel, PhD National Institute on Drug Abuse (NIDA) Bethesda, Maryland, USA	Andrea Beretta, PhD Knowledge Discovery and Data Mining, ISTI-CNR, Pisa, Italy	Cristina Botella, PhD Universitat Jaume, Castelló de la Plana Castellón, Spain
Stéphane Bouchard, PhD Université du Québec en Outaouais Gatineau, Canada	Giulia Brizzi, PhD Cand. Università Cattolica del Sacro Cuore di Milano, Italy	Andrew Campbell, PhD The University of Sydney Sydney, Australia
Scott Debb, EdD Norfolk State University Norfolk, Virginia, USA	Anna Flavia Di Natale, PhD Università Cattolica del Sacro Cuore di Milano, Italy	Chris Fullwood, PhD University of Wolverhampton Wolverhampton, United Kingdom
Luciano Gamberini, PhD University of Padova Padova, Italy	Pedro Gamito, PhD Universidade Lusófona de Humanidades e Tecnologia Lisbon, Portugal	Karen Holmes, PhD Norfolk State University Norfolk, Virginia, USA
Stella Gelli, MD. Università Cattolica del Sacro Cuore di Milano Italy	José Gutierrez Maldonado, PhD University of Barcelona Barcelona, Spain	Thomas Parsons, PhD University of North Texas Denton, Texas, USA
Susan Persky, PhD National Institutes of Health (NIH), Bethesda Maryland, USA	Wendy Powell, PhD Tilburg University Tilburg, Netherlands	Elisa Rabarbari, MS Università Cattolica del Sacro Cuore di Milano, Italy
Dan Romer, PhD University of Pennsylvania Philadelphia Pennsylvania, USA	Chiara Rossi, PhD Università Cattolica del Sacro Cuore di Milano, Italy	Michael Roy, M.D., MPH, FACP, COL (Ret.) Uniformed Services University, Bethesda Maryland, USA
Brad Ridout, PhD The University of Sydney Sydney, Australia	Stefan Stieger, PhD Karl Landsteiner University Krems an der Donau, Austria	Darlene Colson, PhD Norfolk State University Norfolk, Virginia, USA

General Information

Annual Review of CyberTherapy and Telemedicine (ARCTT – ISSN: 1554-8716) is published annually (once per year) by the Interactive Media Institute (IMI), a 501c3 non-profit organization, dedicated to incorporating interdisciplinary researchers from around the world to create, test, and develop clinical protocols for the medical and psychological community. IMI realizes that the mind and body work in concert to affect quality of life in individuals and works to develop technology that can be effectively used to improve the standards and reduce the cost of healthcare delivery.

Interactive Media Institute
6540 Lusk Boulevard, Suite C115
San Diego, CA 92121 USA
Telephone:+1 858 642 0267
E-mail: cybertherapy@vrphobia.com
Web site: <http://www.interactivemediainstitute.com>
Journal Web site: <http://www.arctt.info>

Copyright © 2024 by Interactive Media Institute. Printed in the United States of America.

About the Journal

ARCTT is a peer-reviewed all-purpose journal covering a wide variety of topics of interest to the mental health, neuroscience, and rehabilitation communities. The mission of ARCTT is to provide systematic, periodic examinations of scholarly advances in the field of CyberTherapy and Telemedicine through original investigations in the Telemedicine and CyberTherapy areas, novel experimental clinical studies, and critical authoritative reviews. It is directed to healthcare providers and researchers who are interested in the applications of advanced media for improving the delivery and efficacy of mental healthcare and rehabilitative services.

Manuscript Proposal and Submission

Because Annual Review papers examine either novel therapeutic methods and trials or a specific clinical application in depth, they are written by experienced researchers upon invitation from our Editorial Board. The editors nevertheless welcome suggestions from our readers. Questions or comments about editorial content or policies should be directed to the editors only.

Manuscript Preparation

Manuscripts should be submitted in electronic format as well as on 8 1/2 x 11-in. paper (three copies), double-spaced format. Authors should prepare manuscripts according to the Publication Manual of the American Psychological Association (7th Ed.). Original, camera-ready artwork for figures is required. Original color figures can be printed in color at the editors' discretion and provided the author agrees to pay in full the associated production costs; an estimate of these costs is available from the ARCTT production office on request. ARCTT policy prohibits an author from submitting the same manuscript for concurrent consideration by two or more publications. Authors have an obligation to consult journal editors concerning prior publication of any data upon which their article depends. As this journal is a primary journal that publishes original material only, ARCTT policy prohibits as well publication of any manuscript that has already been published in whole or substantial part elsewhere, unless authorized by the journal editors.

Disclaimer

All the published papers, editorial news and comments, opinions, findings, conclusions, or recommendations in ARCTT are those of the author(s), and do not necessarily reflects or constitute the opinions of the Journal, its Publisher, and its editorial staff.

Contents

Preface	v
Brenda K. Wiederhold and Giuseppe Riva	
Section I. Editorial	
Understanding Artificial Intelligence: A Multidisciplinary Analysis of AI's Distinct Cognitive Architecture <i>Giuseppe Riva, Elena Sajno, Stefano De Gaspari, Chiara Pupillo, Maria Sansoni, Giuseppe Passalacqua, Federico Longoni and Brenda K. Wiederhold</i>	20
Section II. Critical Reviews	
Rethinking some Virtual Human Applications <i>Robert Hubal</i>	27
The Cyberpsychology of Deception: A Mini Review of the Psychological Factors Influencing Scam Compliance <i>David Lacey, Andrew Campbell, Sigi Goode and Brad Ridout</i>	34
Singularity: A Non-Dangerous Case <i>Alexander Voiskounsky</i>	42
Outlining Theoretical Underpinnings of Game Transfer Phenomena <i>Angelica B. Ortiz De Gortari, Andrzej Cudo, Julio C. Llamas-Alonso and Arturo Lerner</i>	44
Visceral Extension and the Beast-Machine: Toward a Neuropsychological Criterion for the Extended Self <i>Joseph Neisser</i>	50
Onlife Subjectivities Construction in Two Radicalized Ecosystems. A Practical Application of a Theoretical Perspective <i>Ilaria Vergine, Carlo Galimberti, Barbara Lucini and Marco Lombardi</i>	55
The Structural Considerations of an Online Games Community <i>Eric J. Chou</i>	60
Section III. Evaluation Studies	
My Circle: Development and Evaluation of an Online Social Networking Platform for Clinically Guided Peer-to-Peer Support for Young People with Mental Health Concerns <i>Brad Ridout, Krestina L. Amon, Rowena Forsyth and Andrew J. Campbell</i>	70
Application of Abstract Generative Imagery for Psychological Diagnosis: A Pilot Study <i>Andrey V. Vlasov, Natalia E. Sergeeva, Vasily V. Serikov and George A. Ksrkashadze</i>	76

Exploring the Relaxing Effects of Hallucinatory Simulated Experiences in Virtual Reality: A Pilot Study <i>Giulia Brizzi, Chiara Pupillo, Clara Rastelli, Antonino Greco, Luca Bernardelli, Anna Flavia Di Natale, Silvia Francesca Maria Pizzoli, Elena Sajno, Fabio Frisone, Andrea Ubaldi, Daniele Di Lernia and Giuseppe Riva</i>	81
CALMA Online Psychological First Aid Program in Peru: Personal and Sociodemographic Characteristics and Challenges <i>Karen Pérez-Maraví, Giovanna Osorio and Iván Montes-Iturriaga</i>	86
Dynamics of Online Relationships Through Virtual Reality Experiences <i>Aimée Argüero-Fonseca and Davide Marchioro</i>	91
Using Eye Gaze to Differentiate Internal Feelings of Familiarity in Virtual Reality Environments: Challenges and Opportunities <i>Trevor Chartier, Iliana Castillon, Videep Venkatesha, Anne M. Clearly and Nathaniel T. Blanchard</i>	96
Satisfaction with the CALMA Online Psychological First Aid Program in Peru <i>Karen Pérez-Maraví, Giovanna Osorio and Iván Montes-Iturriaga</i>	101
Exploring the Impact of Gamified Virtual Reality Cognitive Training on Improving Anxiety and Depressive Symptoms in Spanish Adolescent Girls <i>Aikaterini Ampatzoglou, Anna Carballo-Marquez, Anna Garcia-Casanovas, Juliana Rojas, Maria Fernandez-Capo, Maria Gamiz-Sanfeliu, Maite Garolera and Bruno Porras-Garcia</i>	106
Development of a Virtual Reality Driving Simulator for Cognitive Research: a Motion Sickness Analysis <i>Maria Eleonora Minissi, Luna Maddalon, Juan Miguel Rochina, Pablo Escribano-Bastante and Mariano Alcañiz</i>	113

Section IV. Original Research

Social Media: Positive or Negative for Health? User Profiles in the Context of Exposure to Body Image-Positive Content on Social Media <i>Vincent Lemieux, Ariane Cassista and Marie-Pierre Gagnon-Girouard</i>	120
Human Body Odors Modulation on Affective Processing of Social-Emotional Virtual Environments <i>Luna Maddalon, Sergio Cervera-Torres, Maria Eleonora Minissi, Alberto Greco, Luca Citi, Javier Marin-Morales, Mariano Alcañiz</i>	125
Machine Learning Classification Using Emotion Language and Subjective Difficulty to Distinguish Patients with Functional Seizures from Trauma Controls <i>Natalie N. Newton, Nicole A. Roberts, Estrella M. Contreras and Mary H. Burleson</i>	131
Pursuing Perfection or Striving Towards “Superwoman”? Investigating the Roles of Social Media and Individual Differences in Eating Disorders <i>Amanda Batista and Deborah Hall</i>	138
Understanding What Actions and Traits Facilitate Social Media Campaigning on Transgender Issues Through Perceptions of the BBC and Its Media Coverage <i>Dan Lloyd and Liam Challenor</i>	143

Section V. Clinical Observations

Development and Testing of a Virtual Reality Intervention with Adolescents Hospitalized for Suicide-Related Crises <i>Marisa E. Marracini, Robert Hubal, Lauren E. Delgaty, Telieha J. Middleton, Rachel Anonick, Yixin Li, and Jerry Heneghan</i>	150
The Analysis of Digital Biomarkers for Identifying and Predicting Frailty and Healthy Elderly through Machine Learning	156

<i>Stefano De Gaspari, Irene Alice Chicchi Giglioli, Chiara Pupillo, Elena Sajno, Daniele Di Lernia, Alessandro Capriotti, Paolo Olivetti and Giuseppe Riva</i>	
Virtual Reality Training for Spatial Perspective-Taking in Children with ASD	161
<i>Luna Maddalon, Maria Eleonora Minissi, Thomas D. Parsons, Amaia Hervás-Zúñiga and Mariano Alcañiz</i>	
Physiological Response to a Virtual Reality Simulation for Preoperative Stress	166
Inoculation	
<i>Catherine Proulx, Michael S. D. Smith, Vincent Gagnon Shaigetz, Gabrielle S. Logan, Jordana L. Sommer, Pamela Hebbard, Kristin Reynolds, W. Alan Mutch, Natalie Mota, Jessica L. Maples-Keller, Rakesh C. Arora, and Renée Elgabalawy</i>	
Inclusive Cognitive Commercial Training for Children: A Preliminary Evaluation of the Kokoro Kids App in Enhancing Visuospatial Abilities in Autistic and Neurotypical Children	170
<i>Maria Eleonora Minissi, Luna Maddalon, Andrea Lopez-Castillo, Rocio Saez, Carles Pons, Elena Parra-Vargas and Mariano Alcañiz</i>	
Virtual Standardized LLM-AI Patients for Clinical Practice	177
<i>Patrick G. Kenny and Thomas D. Parsons</i>	
Novel Memory Assessment Methods in Minimally Invasive Epilepsy Surgery	183
<i>Daniel Drane, Noah Okada, Yousef Rajeh, Ellie Paek, Keith Adams, Emily Wall</i>	
Development and Application of a Three-Screen Enveloping Projection Cyberspace that Is Easy to Disassemble and Assemble, with the Aim of Contributing to People's Wellbeing	188
<i>Mieko Ohsuga</i>	

Section VI. Work In Progress

New Considerations for Memory Assessment	195
<i>Anne M. Cleary</i>	
Found Value in the Artifice of Semblance: Engaging with Abstraction within Generative AI and Integration Practice in PAT	200
<i>Xavier Nokes and Pavan Turaga</i>	
Bridging the Gap Between Perceived and Actual Personalities in Real and Virtual Worlds Through EEG Analysis	206
<i>Timothy McMahan, Chotepong Victor Tangton, Advay Vyas, Chloe Huang and Aron Yewoor</i>	
Democratizing Psychophysiological Research: Leveraging Open Resources for Experimental Settings and Data Analysis of Flow, Engagement and Performance	213
<i>Elena Sajno, Samantha Chan, Federico Dagostin, Stefano De Gaspari, Chiara Pupillo and Giuseppe Riva</i>	
Exploring Ethical Considerations of Serious Game: a Qualitative Secondary Analysis on Improving COVID-19 Vaccination Among Youth	219
<i>Lihong Ou and Thomas D. Parsons</i>	
Harmful Online Content and Cross-Cultural Challenges Faced by Adolescents: Findings From Focus Groups Conducted in Brazil and Australia	225
<i>Alex S. G. Pessoa, Andrew J. Campbell and Brad Ridout</i>	
Leveraging Artificial Intelligence to mitigate User Susceptibility to Malicious Push Notifications during Augmented Reality Immersion	230
<i>Sarah Katz</i>	
XR for STEM Education and Social Health Outcomes: Focus on Embodied Mixed Reality	232
<i>Mina C. Johnson-Glenberg and Yanjun Lyu</i>	

A Game of Mindless Swiping: How Gamification and Mindless Swiping Affects Online Dating Users' Psychology <i>Aleksandra Bombaci Froelich</i>	237
Subject Index	242
Author Index	243

SECTION I

EDITORIAL

This introductory text expresses the position of the Editors – Brenda K. Wiederhold and Giuseppe Riva - the focus of this year's issue.

Brenda K. Wiederhold and Giuseppe Riva

Understanding Artificial Intelligence: A Multidisciplinary Analysis of AI's Distinct Cognitive Architecture

Giuseppe RIVA ^{a,b,1}, Elena SAJNO ^{a,c}, Stefano DE GASPARI ^{a,c}, Chiara PUPILLO ^{a,c},
Maria SANSONI ^a, Giuseppe PASSALACQUA ^{a,c}, Federico LONGONI ^{a,c}
Brenda K. WIEDERHOLD ^d

^a *Humane Technology Lab., Catholic University of Sacred Heart, Milan, Italy,*

^b *Applied Technology for Neuro-Psychology Lab., IRCCS Istituto Auxologico Italiano,
Milan, Italy*

^c *Department of Computer Science, University of Pisa, Pisa, Italy*

^d *Virtual Reality Medical Center, Scripps Memorial Hospital, La Jolla, California,
USA*

ORCID ID: Giuseppe Riva <https://orcid.org/0000-0003-3657-106X>

Abstract. As Artificial Intelligence (AI) systems become increasingly sophisticated and pervasive, understanding their true nature and limitations requires moving beyond single-discipline perspectives. This paper argues that effectively developing and deploying AI systems demands an integrated framework combining insights from computer science, cognitive psychology, linguistics, and cyberpsychology. We examine three complementary theoretical frameworks that together reveal AI's fundamental nature: Psychomantics demonstrates how AI systems create sophisticated syntax-semantic mappings but lack true experiential understanding; System 0 thinking positions AI as a distinct cognitive layer that preprocesses information while remaining dependent on human interpretation; and Cyberpsychology illuminates how humans form mental models and relationships with AI systems. These insights converge on a central argument: effective AI development requires embracing rather than obscuring the fundamental differences between artificial and human intelligence. Instead of pursuing AI systems that attempt to replicate human cognition, we should focus on developing systems that transparently complement human capabilities while acknowledging their inherent limitations in meaning-making and understanding.

Keywords. Artificial intelligence, cyberpsychology, psychomantics, system 0

1. Introduction

The rapid evolution of artificial intelligence (AI) represents a complex intersection of multiple disciplines, each offering unique perspectives on how these systems process information and interact with human cognition. While AI, particularly Large Language Models (LLMs), demonstrates remarkable capabilities in language acquisition and concept learning, understanding its true nature requires a multidisciplinary lens [1] that spans computer science, cognitive psychology, linguistics, cyberpsychology, and ethics. Two emerging theoretical frameworks help illuminate this complex landscape. The "Psychomantics" framework, introduced by Riva and colleagues [2], bridges cognitive science, linguistics, and computer science to examine how LLMs acquire, learn, remember, and utilize information. This approach reveals how LLMs create extensive syntax-semantics maps to manipulate linguistic patterns while following Grice's Cooperative principle to generate relevant responses. However, unlike human cognition which draws from experiential, emotional, and imaginative sources of meaning, LLMs operate within the constraints of their training data without true embodied understanding.

¹ Corresponding Author: Giuseppe Riva; Humane Technology Lab., Catholic University of Sacred Heart, Milan, Italy; Email: giuseppe.riva@unicatt.it

Complementing this view, Chiriatti and colleagues [3] propose conceptualizing human-AI interaction as "System 0" - a distinct cognitive layer that operates alongside Kahneman's system 1 (fast, intuitive thinking) and system 2 (slow, analytical thinking). This framework positions AI systems as an artificial preprocessor that shapes how information flows into traditional human cognitive processes, while acknowledging AI's dependence on human interpretation for meaningful outputs.

Cyberpsychology plays a particularly crucial role in this interdisciplinary analysis by examining how individuals perceive and interact with AI systems. Research has shown that humans often anthropomorphize conversational AI, forming bonds that extend beyond mere tool usage. Understanding these psychological dynamics is essential for designing AI systems that are both user-friendly and ethically sound.

By adopting this multidisciplinary approach, we can foster a holistic understanding of AI that encompasses both its technical capabilities and its psychological, social, and ethical implications. Such a comprehensive perspective is vital for developing AI systems that are not only technologically advanced but also aligned with human values and societal needs.

2. Psychomatics: Understanding Digital Brains

Understanding how artificial intelligence processes information and generates responses is crucial for developing effective human-AI interactions. While LLMs demonstrate impressive capabilities in tasks ranging from language comprehension to problem-solving, their underlying mechanisms differ fundamentally from human cognition. The Psychomatics framework offers a systematic approach to understanding these differences by examining how LLMs acquire, process, and utilize information through the lens of cognitive science, linguistics, and computer science [1].

At its core, Psychomatics reveals that LLMs operate through extensive syntax-semantics mapping rather than genuine understanding. Unlike humans, who derive meaning from direct experience, emotional connections, and embodied cognition, LLMs create sophisticated statistical patterns that link linguistic structures to probable meanings. For example, when processing the sentence "The cat is on the mat," an LLM doesn't understand the physical reality of a cat or mat, but rather maps the statistical relationships between these terms and their contextual usage patterns. This mapping allows LLMs to generate coherent responses by following Grice's Cooperative principle [4] - providing relevant, informative, and appropriately structured communication - without truly comprehending the underlying concepts.

The framework identifies several key mechanisms that enable LLM functionality. Through self-attention mechanisms [5], LLMs capture dependencies between words within sequences, similar to how humans process syntagmatic relations in language. Cross-attention mechanisms allow LLMs to map relationships between different sequences, mimicking how humans form associative connections between concepts. For instance, in processing a medical query, an LLM might link symptoms to diagnoses not through medical understanding, but through statistical patterns learned from training data. However, this reliance on statistical patterns rather than embodied understanding leads to important limitations. LLMs can generate factually incorrect information (hallucinations) when their training data lacks sufficient context or when they encounter novel situations requiring real-world understanding. They struggle with implicit meanings, sarcasm, and contextual nuances that humans naturally grasp through their social and emotional experiences. Additionally, LLMs operate in an asocial environment, lacking the dynamic learning capabilities that humans develop through continuous social interaction and personal experience [6].

The impact of LLMs on human-AI interaction is profound and multifaceted. While these systems can enhance human cognitive capabilities by processing vast amounts of information and generating sophisticated responses, they also risk creating a false sense of understanding [7]. Users might anthropomorphize LLMs and attribute deeper comprehension than actually exists, potentially leading to over-reliance or misuse of these systems. This highlights the importance of developing frameworks for responsible AI deployment that acknowledge both the capabilities and limitations of LLMs.

Looking ahead, the Psychomantics framework suggests that effective human-AI interaction requires understanding these fundamental differences between human and machine cognition. Rather than treating LLMs as human-like intelligence, we should view them as powerful but limited tools that complement human cognitive capabilities. This understanding can guide the development of AI systems that better serve human needs while maintaining appropriate boundaries and expectations. The future of human-AI interaction lies not in creating perfect simulations of human cognition, but in developing systems that transparently and effectively augment human capabilities while acknowledging their inherent limitations.

3. System 0: When Artificial Intelligence Augments Human Cognition

As artificial intelligence becomes increasingly integrated into our daily cognitive processes, understanding its relationship with human thought patterns has emerged as a critical frontier in cognitive science. While traditional frameworks like Kahneman's System 1 (fast, intuitive thinking) and System 2 (slow, analytical thinking) help us understand human cognition [8], they don't fully capture how AI systems are fundamentally reshaping our cognitive landscape. The concept of System 0 provides a theoretical framework for understanding this new cognitive reality [3], positioning AI as a foundational layer that preprocesses information before it enters human cognitive systems.

System 0 represents the outsourcing of cognitive tasks to AI systems that can process vast amounts of data and perform complex computations beyond human capabilities. Unlike Systems 1 and 2, which operate within individual minds, System 0 functions as an artificial, non-biological layer of distributed intelligence that interacts with and augments both intuitive and analytical thinking processes. For example, when using an AI writing assistant, the system doesn't merely extend our writing capabilities; it actively shapes how we approach the writing process by suggesting structures, identifying patterns, and proposing improvements before we engage in traditional cognitive processing. Similarly, when using AI-powered search engines, the information we receive is preprocessed and curated, fundamentally altering how we access and interpret knowledge.

The framework identifies several key characteristics that distinguish System 0 from traditional cognitive systems. It exhibits high information flow, with continuous two-way interaction between humans and AI. It demonstrates increasing reliability as AI systems become more sophisticated, leading to greater trust and procedural transparency in human-AI interactions. Perhaps most significantly, System 0 shows strong individualization capabilities, providing personalized cognitive support that adapts to individual users' needs and preferences.

However, System 0 differs crucially from Systems 1 and 2 in its lack of inherent meaning-making capabilities. While it can process and manipulate data with remarkable efficiency, it relies entirely on human interpretation to generate meaningful insights. This limitation creates a complex dynamic where human cognition must actively engage with AI-processed information to create genuine understanding.

The impact of System 0 on human decision-making is profound and potentially double-edged. On one hand, it enhances our cognitive capabilities by providing sophisticated preprocessing of information, enabling us to handle more complex problems and access broader knowledge bases. On the other hand, it raises important concerns about cognitive autonomy and critical thinking. As we increasingly rely on AI-mediated information and decision-making processes, we risk adopting epistemological norms aligned with computational logic rather than human reasoning.

Looking forward, the integration of System 0 into human cognition presents both opportunities and challenges. The potential benefits include enhanced problem-solving capabilities, more informed decision-making, and expanded cognitive reach. However, these must be balanced against the risks of decreased autonomy and potential erosion of critical thinking skills. The key to successful integration lies in developing frameworks that maintain human agency while leveraging AI's capabilities effectively.

Understanding System 0 is essential for developing AI systems that enhance rather than diminish human cognitive capabilities. This requires careful attention to how AI

systems preprocess information, how humans interact with AI-generated insights, and how to maintain critical thinking skills in an AI-augmented cognitive environment. As we continue to develop and deploy AI systems, the concept of System 0 provides a valuable framework for ensuring these technologies serve as effective cognitive partners while preserving human autonomy and decision-making capabilities.

4. Cyberpsychology: The Psychology of Human-AI Relationships

As artificial intelligence becomes increasingly integrated into our daily lives, understanding how humans perceive and interact with AI systems has emerged as a critical frontier in technology research [9]. Cyberpsychology, the study of human interaction with technology, offers unique insights into the psychological dynamics that shape these human-AI relationships [10]. This field has become particularly relevant as AI systems evolve from simple tools to complex interactive agents that can engage in sophisticated dialogue, make decisions, and adapt to user behavior.

The psychological dimensions of human-AI interaction operate on multiple levels [11]. At the cognitive level, humans tend to anthropomorphize AI systems, attributing human-like characteristics and mental states to them even when knowing they are artificial. For example, research has shown that users often form emotional bonds with conversational AI, treating them as social entities rather than mere tools [12]. This phenomenon extends beyond simple projections - humans develop mental models of AI capabilities that significantly influence their trust, expectations, and willingness to rely on AI systems [13]. Studies demonstrate how these mental models can be either overly optimistic, leading to excessive trust in AI capabilities, or overly skeptical, resulting in resistance to beneficial AI applications [14]. The Psychomatics framework reveals how these perceptions are complicated by AI's inherent capability uncertainty and output complexity. Users must navigate systems whose abilities may evolve over time and whose outputs can range from simple to highly adaptive and complex. This uncertainty shapes how users build trust and develop usage patterns with AI systems.

The social-psychological aspects of human-AI interaction are equally important [15]. As highlighted by System 0 thinking, AI has become a foundational layer that preprocesses information before it enters human cognitive systems. This creates a complex dynamic where human decision-making and social behavior are increasingly mediated by AI systems, yet the psychological implications of this mediation are not fully understood. Users may experience cognitive offloading, where they become dependent on AI for tasks they previously handled independently, or develop new cognitive strategies that integrate AI capabilities into their problem-solving approaches. Cyberpsychology's impact on understanding AI interactions extends beyond theoretical insights to practical applications. By revealing how users develop mental models of AI systems, design teams can create more intuitive interfaces that properly calibrate user expectations [16]. Understanding the psychological mechanisms behind trust formation helps in developing AI systems that are not only technically capable but also psychologically accessible. Moreover, cyberpsychological research illuminates how different user groups may interact with AI differently based on their technological self-efficacy, cultural background, and previous experiences with technology [17].

Looking forward, cyberpsychology will play an increasingly crucial role in shaping the future of human-AI interaction. As AI systems become more sophisticated and pervasive, understanding the psychological dimensions of these interactions will be essential for developing technologies that enhance rather than diminish human capabilities. This understanding can help address critical challenges in AI deployment, from preventing over-reliance on AI systems to ensuring that AI augments human decision-making in psychologically healthy ways. The field's insights will be particularly valuable in sensitive domains like healthcare, education, and personal assistance, where the psychological dynamics of human-AI interaction directly impact system effectiveness and user wellbeing.

5. Conclusions

The multidisciplinary analysis presented in this paper leads to several critical insights about the nature of artificial intelligence and its relationship to human cognition. First, the Psychomantics framework reveals that while AI systems can create sophisticated statistical patterns linking linguistic structures to probable meanings, they fundamentally lack the experiential and embodied understanding that underlies human cognition. This isn't merely a temporary limitation but reflects a fundamental difference in how AI systems process and derive meaning from information.

Second, the System 0 perspective demonstrates that AI systems are best understood not as autonomous thinking entities but as a distinct cognitive layer that preprocesses information for human interpretation. This positioning helps resolve the apparent paradox of AI systems that can perform complex cognitive tasks while lacking true understanding - they function as sophisticated information processors that augment rather than replicate human cognitive capabilities.

Third, cyberpsychology research shows how humans naturally form mental models and relationships with AI systems that often overestimate their capabilities for genuine understanding. This tendency toward anthropomorphizing AI needs to be carefully managed to ensure appropriate trust and reliance on AI systems.

These insights converge on a central argument: effective AI development requires embracing rather than obscuring the fundamental differences between artificial and human intelligence. Instead of pursuing AI systems that attempt to replicate human cognition, we should focus on developing systems that transparently complement human capabilities while acknowledging their inherent limitations in meaning-making and understanding. This approach has several practical implications:

- AI system design should emphasize transparency about operational limitations rather than attempting to create the illusion of human-like understanding;
- Development frameworks should focus on creating effective human-AI partnerships that leverage the complementary strengths of both;
- Implementation strategies should include explicit consideration of how humans form psychological relationships with AI systems.

Looking ahead, this integrated understanding provides a foundation for developing AI systems that enhance human cognitive capabilities while maintaining appropriate boundaries and expectations. Success in this endeavor requires continued cross-disciplinary dialogue and research that brings together technical capabilities, cognitive science insights, and psychological understanding of human-AI interaction.

6. Acknowledgments

This paper was funded by the European Union - Next Generation EU - Italian PNRR under the FAIR Project "Co-XAI - Cognitive Decision Intelligence Framework for Explainable AI Systems". The views and opinions expressed are only those of the authors and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

References

- [1] Dignum V. AI is multidisciplinary. *AI Matters.* 2020;5:18–21. doi: <https://doi.org/10.1145/3416666.3421426>.
- [2] Riva G, Mantovani F, Wiederhold BK, Marchetti A, Gaggioli A. Psychomantics—A multidisciplinary framework for understanding artificial minds. *Cyberpsychol Behav Soc Netw.* 2024. doi: <https://doi.org/10.1089/cyber.2024.0003>.
- [3] Chiratti M, Ganapini M, Panai E, Ubiali M, Riva G. The case for human–AI interaction as system 0 thinking. *Nat Hum Behav.* 2024;8:1829–1830. doi: <https://doi.org/10.1038/s41562-024-01872-7>.
- [4] Grice HP. Logic and conversation. In: Cole P, Morgan JL, editors. *Syntax and Semantics: Vol 3. Speech Acts.* New York: Academic Press; 1975. p. 41–58.
- [5] Vaswani A, Shazeer NM, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I. Attention is all you need. In: *Advances in Neural Information Processing Systems;* 2017. p. 5998–6008.

- [6] Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *J Med Internet Res.* 2019;21:e13216. doi: <https://doi.org/10.2196/13216>.
- [7] Messeri L, Crockett MJ. Artificial intelligence and illusions of understanding in scientific research. *Nature.* 2024;627:49–58. doi: <https://doi.org/10.1038/s41586-024-07372-3>.
- [8] Kahneman D. Maps of bounded rationality: a perspective on intuitive judgment and choice. In: Frängsmyr T, editor. *The Nobel Prizes 2002.* Stockholm: Nobel Foundation; 2002. p. 449–489.
- [9] Xu W, Dainoff MJ, Ge L, Gao Z. Transitioning to human interaction with AI systems: new challenges and opportunities for HCI professionals to enable human-centered AI. *Int J Hum Comput Interact.* 2023;39(5):494–518. doi: <https://doi.org/10.1080/10447318.2022.2094212>.
- [10] Fortuna P. Positive cyberpsychology as a field of study of the well-being of people interacting with and via technology. *Front Psychol.* 2023;14:1237890. doi: <https://doi.org/10.3389/fpsyg.2023.1237890>.
- [11] Hancock PA, Billings DR, Schaefer KE, Chen JYC, de Visser EJ, Parasuraman R. A meta-analysis of factors affecting trust in human-robot interaction. *Hum Factors.* 2011;53:517–527. doi: <https://doi.org/10.1177/0018720811417254>.
- [12] Pentina I, Hancock T, Xie T. Exploring relationship development with social chatbots: a mixed-method study of Replika. *Comput Hum Behav.* 2023;140:107600. doi: <https://doi.org/10.1016/j.chb.2022.107600>.
- [13] Kourous T, Papa V. Digital mirrors: AI companions and the self. *Societies.* 2024;14(2):200. doi: <https://doi.org/10.3390/soc14020024>.
- [14] Pan W, Liu D, Meng J, Liu H. Human–AI communication in initial encounters: how AI agency affects trust, liking, and chat quality evaluation. *New Media Soc.* 2024;14614448241259149. doi: <https://doi.org/10.1177/14614448241259149>.
- [15] Bingley WJ, Haslam SA, Steffens NK, Gillespie N, Worthy P, Curtis C, et al. Enlarging the model of the human at the heart of human-centered AI: a social self-determination model of AI system impact. *New Ideas Psychol.* 2023;70:101025. doi: <https://doi.org/10.1016/j.newideapsych.2023.101025>.
- [16] Pataranutaporn P, Liu R, Finn E, Maes P. Influencing human–AI interaction by priming beliefs about AI can increase perceived trustworthiness, empathy and effectiveness. *Nat Mach Intell.* 2023;5:1076–1086. doi: <https://doi.org/10.1038/s42256-023-00750-0>.
- [17] Andrews RW, Lilly JM, Srivastava D, Feigh KM. The role of shared mental models in human-AI teams: a theoretical review. *Theor Issues Ergon Sci.* 2023;24(2):129–175. doi: <https://doi.org/10.1080/1463922X.2023.2165766>.

SECTION II

CRITICAL REVIEWS

In general, there are two reasons why cybertherapy is used: either because there is no alternative, or because it is in some sense better than traditional medicine.

In this sense telehealth has been used very successfully for optimizing health services delivery to people who are isolated due to social and physical boundaries and limitations.

Nevertheless, the benefits of cybertherapy, due to the variety of its applications and their uneven development, are not self-evident.

However, the emergence of cybertherapy is supporting the cost-effectiveness of certain applications, such as assessment, rehabilitation and therapy in clinical psychology and neuroscience.

Wiederhold & Riva, 2004

Rethinking some Virtual Human Applications

Robert HUBAL^{a,1}

^a*Renaissance Computing Institute, Chapel Hill, NC*
ORCID ID: Robert Hubal <https://orcid.org/0000-0003-2637-0918>

Abstract. Increasingly realistic virtual environments incorporating virtual characters have been used to train or assess actual behavior, such as of people at risk, and identify reasons to remediate or intervene. Technology has improved so rapidly that today's capabilities to create situations to focus training and intervention outshine past efforts. To name just a few current examples, tools like Unreal's MetaHuman Creator for creating characters, Midjourney for creating environments, OpenAI's ChatGPT for scripting, and GIFT for tutoring have enormous potential, as these tools promise to reduce simulation costs and increase realism. This paper, in contrast, discusses some movement in the other direction: Recent efforts suggest that increased realism may not always have resulting cost benefit for training and assessment. Lessons learned and recommendations are presented to guide future developers.

Keywords. Consequential conversations, responsive virtual humans, character and application fidelity.

1. Introduction

Social skills deficits show up as poor behaviors in demanding social situations. For example, in a confrontational situation, some individuals misperceive linguistic, gestural, and expressive cues of another person as indicating hostile intent. Misinterpretation of cues could lead to an escalation of the situation. ‘Consequential’ conversations, as used here, are those that involve challenging content, may lead to adverse outcomes, and require deft social interaction skills to navigate. The partner(s) in the conversation may be difficult to deal with, emotional or confused, or focused on an agenda. The topic of conversation may be sensitive, charged, controversial, or zero-sum.

These situations are uncomfortably common. Military personnel returning from stressful combat demonstrate social skills deficits in numerous ways including domestic discord, substance use, and generalized aggression and violence [1,2]. Law enforcement personnel who are not trained in crowd control can and do misunderstanding actions taken by protesters and, rather than use de-escalation skills, apply unnecessary force [3,4]. Conversely, minority citizens may perceive bias from authority figures [5] even when unintended and speak angrily and act out unsafely. More mundanely, therapy targeted to individuals who demonstrate social skills deficits (e.g., those with some forms of autism, or depression or other mental health concerns) is often general, given measurement techniques [6], whereas their deficits are specific. In all of these scenarios, training or education or treatment should instead be targeted to those behaviors that pose greatest risk to the individual (or those around the individual) and thus greatest benefit when addressed. Typical assessments of social skills ask individuals what they would do in hypothetical situations, or they present non-interactive situations to gauge individuals’

¹ Corresponding author: Robert Hubal, hubal@renci.org

reactions, or they present questionnaires. These typical assessments are thus descriptive—not demonstrative—in nature.

1.1. Virtual Human Applications

A different approach is to use a gaming engine and established artificial intelligence (AI) models to present situations to gauge behavior. A key affordance of game-based technology is the capacity to provide users with an opportunity to role-play identities in the context of a realistic scenario [7]. The virtual environment provides opportunities for users to participate as well as vicariously observe virtual characters in interacting with the situation. Here, the idea is to elicit decision making and social skills, including underlying abilities such as emotional control, information seeking, expressing preferences, negotiation and willingness to compromise, and using non-provocative language. The skills are assessed by simulating social encounters that may lead to adverse consequences. In these applications, individuals' behavior is assessed by observing them while they interact within the simulated environment (e.g., looking at body language, tone of voice, emotional control, and reaction time) and by looking at what type of outcome occurs—usually a good outcome demonstrating avoidance of risky behavior or a poorer one demonstrating risky behavior.

Software applications designed to train or assess dialog within consequential conversations have, for many years, employed virtual characters. Virtual characters are, paradigmatically, multimodal embodied conversational agents—responsive partners with which a user communicates to navigate a given situation or achieve a goal. Virtual character applications have various advantages, including their reproducibility, safety and controllability, ease of distribution, and objectivity, and the ability to introduce intelligent tutoring. Typically, such applications have engaged a user using realistic virtual participants in a realistic setting. Multiple situations address a range of social skills competencies, including poor emotion expression recognition, impulsivity, insensitivity to penalties, hostility bias, gender stereotyping, acceptance of dating or partner violence, and risky decision making. The virtual environments are designed to be flexible and reusable, so that, by having runtime parameters define any given situation, minor visual (e.g., the characters' appearance) or behavioral (e.g., a character acts angry vs. confused) variation is possible, or major changes, and situations can be repurposed.

2. Rethinking Realism

Realism has increased dramatically as technology and capability have improved, so that today's characters can be made to be lifelike in appearance, allow for natural language interaction, and use advanced behavior models to react or respond to user actions appropriately to the context. Over the years this author [8-12] and many others have used evolving tools to portray simulated situations. This author started out in the 1990s with a homegrown renderer built atop DirectX; nowadays established game engines such as Unreal and Unity provide more than the functionality needed. IBM's ViaVoice was once used to capture language input augmented by custom grammar and contexts to capture state of the dialog. Now, dialog systems for consequential applications can take advantage of large language models. For this author, the culmination of work was a virtual patient represented in **Figure 1** developed for pharmacy students to learn skills associated with medication therapy management [13].

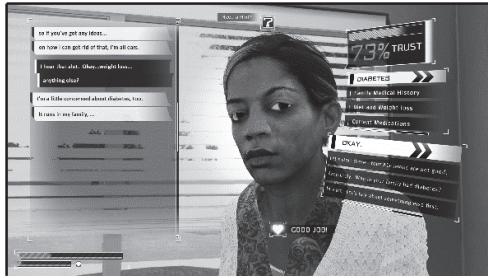


Figure 1. Virtual patient.

2.1. Some Lower-Fidelity Examples

Unexpectedly, perhaps, then, this researcher has moved toward lower fidelity in some recent applications. The shift has been neither total nor abrupt, but, in retrospect, unforeseen.

One example of a less realistic consequential virtual human application is represented by a primarily text-based interactive simulation to present pharmacy preceptors with challenging non-academic situations [14]. There were several reasons for using a web-based, low-interactive approach using Ink as an underlying dialog control mechanism. First, the study was conducted during the Covid pandemic, requiring virtual participation. Distribution of a game-based application might have been difficult. Further, because of the means of engagement (virtual via Zoom), a true natural language interaction, though possible, was infeasible, suggesting that a high-fidelity virtual character also was not needed. Instead, as evidenced by their think-aloud, the situations themselves engaged participants sufficiently. While specifics of the study would be different for any other, these types of determinants may apply elsewhere.

Another example is a branching video application depicted in **Figure 2** for use in helping military personnel understand good techniques engaging foreign civilians [10]. Again, there were a couple of reasons for this choice of technology. First is the availability of resources; the project was better able to support a less immersive than a more immersive approach. Second was the capability of the technology. Military personnel were asked not only to engage with presented characters but also to mark up important or concerning elements of the presented scene, and the technology supported this requirement. Note that the conversations themselves were no less consequential than others where responsive virtual human technology was used. Instead, it was the engagement with presented situations that supported the learning desired.

Another example (in which the author was only indirectly involved) is a graphic novel shown in **Figure 3** developed to illustrate, also for military personnel, life in a combat zone [15]. The approach was chosen to target a younger audience and provide realistic enough situations, “embedding the principles of combat and operational stress control...into realistic and relatable characters, stories, and images” [16], that would prepare personnel for deployment. The author and colleagues have recently worked on a similar design for a parenting application directed toward, and to appeal to, those recovering from substance abuse.



Figure 2. Engaging foreign civilians.



Figure 3. Combat.

A final example is a virtual reality intervention that allows hospitalized adolescents to learn and practice therapeutic skills in stressful situations [17]. At first glance an immersive application does not appear to be low fidelity, but in fact the visuals are necessarily limited due to resource constraints and the dialog interaction is controlled. However, the situations themselves have been heavily tested and shown to be reflective of real events, and adolescents demonstrate engagement within the situations despite noting a perceived lack of realism.

2.2. *Lessons Learned and Recommendations*

The intentions of this paper are to challenge a presumption that high fidelity virtual characters and environments are always the best solution for training or assessment in consequential conversations, and to encourage further research. It is not to discourage greater fidelity when demanded.

When is greater fidelity demanded? One example is when expressions, and in particular changes in expression, need to be noticed, as when there are pressing issues such as interaction or cultural nuances that need to be made explicit. Branching video would be a possibility here [18], though video is not usually as flexible as virtual settings that can easily support a swap of characters of variable ages, ethnicities, and other characteristics, so that the user experiences a range of characters. Relatedly, usage of more realistic virtual characters makes sense when the user needs to engage meaningfully with the character as in talk therapy. Also, higher fidelity is called for when assessing fine-grained procedures that would otherwise be logically complex, unsafe, unethical, unreliable, and/or costly. For instance, programs where medical personnel need to demonstrate precisely where to apply specific instruments benefit from realistic characters [19].

When is lesser fidelity acceptable? Virtual character realism makes less sense when the application does not involve interactivity, such as when characters are in the background or at least not central to the task at hand. When an application centers on physical manipulation, including haptic feedback, then other methods such as manikins or part-task trainers, or a hybrid approach [20], are typically more cost-effective. Further, when users engage just as fully with less realistic characters as more realistic characters, then the former represent a more cost-effective training or assessment.

The author's change in approach for some—not all—applications derives from several additional conditions: Difficulty in developing suitable models to meet user expectations, resource and usability constraints, learner preferences, reappraisal of the purpose of training or assessment, and affordances of underlying technology. For instance, behavior models to drive character actions may need to be highly refined in applications that show nuance (e.g., subtle medical symptoms) but coarser in others where a general appearance of motion (e.g., with graphic novels) is sufficient. Similarly, users can be pretty much assured to be engaged in applications portraying well-designed consequential conversations, particularly when they relate to their lives, and barring an egregious error.² When that engagement demands higher fidelity depends on factors including learning objectives, costs, distribution, and additional technological needs. Also, use of current technologies such as those mentioned above that facilitate

the development and portrayal of realistic characters leads users to expect realistic behaviors, akin to the uncanny valley effect [21].

This author addresses the following questions to guide development: What is the responsibility of users, and what are their actions? Are any skills to be gained meant to be familiarized, practiced, or mastered (the former may need lesser engagement than the latter)? Are the skills procedural, interactive, or strategic (the latter may need greater engagement than the former)? Is the task dangerous, complex, time-consuming, resource-consuming, expensive, or unavailable, so that a realistic simulation promises benefits in relation to costs? How easy or hard is it to model behaviors of entities and objects (the harder to model, the less the fidelity is probably called for)? How visually demanding is the task (the more demanding, the greater fidelity likely required)?

These criteria apply differently for any given application but represent principles to consider as the application is designed. Rapidly advancing technology is exciting and generally worth taking advantage of. But getting users to engage is the lesser challenge; the greater challenge is to neither under- or over-promise what they experience.

References

- [1] Kline AC, Panza KE, Nichter B, Tsai J, Harpaz-Rotem I, Norman SB, Pietrzak RH. Mental health care use among U.S. military veterans: Results from the 2019–2020 National Health and Resilience in Veterans Study. *Psychiatr Serv.* 2022;73(6):628-35. doi: <https://doi.org/10.1176/appi.ps.202100112>
- [2] MacManus D, Rona R, Dickson H, Somaini G, Fear N, Wessely S. Aggressive and violent behavior among military personnel deployed to Iraq and Afghanistan: Prevalence and link with deployment and combat exposure. *Epidemiol Rev.* 2015;37(1):196-212. doi: <https://doi.org/10.1093/epirev/mxu006>
- [3] Engel RS, McManus HD, Herold TD. Does de-escalation training work? A systematic review and call for evidence in police use-of-force reform. *Criminol Public Pol.* 2020;19(3):721-59. doi: <https://doi.org/10.1111/1745-9133.12467>
- [4] Trinkner R, Kerrison EM, Goff PA. The force of fear: Police stereotype threat, self-legitimacy, and support for excessive force. *Law Hum Behav.* 2019;43(5):421-35. doi: <https://doi.org/10.1037/lhb0000339>
- [5] Swencionis JK, Goff PA. The psychological science of racial bias and policing. *Psychol Public Policy Law.* 2017;23(4):398-409. doi: <https://doi.org/10.1037/law0000130>
- [6] Cordier R, Speyer R, Chen YW, Wilkes-Gillan S, Brown T, Bourke-Taylor H, Doma K, Leicht A. Evaluating the psychometric quality of social skills measures: A systematic review. *PLoS One.* 2015;10(7):e0132299. doi: <https://doi.org/10.1371/journal.pone.0132299>
- [7] Bachen CM, Hernández-Ramos P, Raphael C, Waldron A. How do presence, flow, and character identification affect players' empathy and interest in learning from a serious computer game? *Comput Hum Behav.* 2016;64:77-87. doi: <https://doi.org/10.1016/j.chb.2016.06.043>
- [8] Hourani LL, Kizakevich PN, Hubal R, Spira J, Strange LB, Holiday DB. Predeployment stress inoculation training for primary prevention of combat-related stress disorders. *J Cyberther Rehabil.* 2011;4(1):101-16.
- [9] Hubal RC, Evans NR, FitzGerald DP, Hardy KK, Willard VW, Bonner MJ. Implementation of FACS for synthetic characters for use in studying facial expression recognition by survivors of childhood cancer. *Annu Rev Cyberther Telemed.* 2008;6:35-40.
- [10] Hubal RC, Kizakevich PN, Guinn CI, Merino KD, West SL. The virtual standardized patient – Simulated patient-practitioner dialogue for patient interview training. *Stud Health Technol Inform.* 2000;70:133-8. doi: <https://doi.org/10.3233/978-1-60750-914-1-133>
- [11] Hubal R, van Lent M, Wender J, Lande B, Flanagan S, Quinn S. What does it take to train a good stranger. *Procedia Manuf.* 2015;3:3955-62. doi: <https://doi.org/10.1016/j.promfg.2015.07.930>
- [12] Paschall MJ, Fishbein DH, Hubal RC, Eldred D. Psychometric properties of virtual reality vignette performance measures: A novel approach for assessing adolescents' social competency skills. *Health Educ Res.* 2005;20(1):61-70. doi: <https://doi.org/10.1093/her/cvg103>
- [13] Hubal R, Heneghan J. Carolina virtual patient initiative [abstract]. *Pharm Educ.* 2017;17(1):292.
- [14] Williams CR, Hubal R, Wolcott MD, Kruse A. Interactive narrative simulation as a method for preceptor development. *Pharm.* 2021;10(1):5. doi: <https://doi.org/10.3390/pharmacy10010005>
- [15] Kraft H, Peeler R, Laron J, Borst CA, Hourani LL, Lambert S, Wiggins D, Nguyen D, Philyaw D, Jessup E. The docs: A graphic novel. San Diego: Naval Health Research Center; 2010.
- [16] Hourani L, Lambert S, Peeler R, Lane B, Borst C. Graphic novels: A new stress mitigation tool for military training: Developing content for hard-to-reach audiences. *Health Commun.* 2017;32(5):541-9. doi: <https://doi.org/10.1080/10410236.2016.1140265>
- [17] Marraccini M, Hubal R, Middleton T, Delgaty L. Development and testing of a virtual-reality intervention with youth hospitalized for suicide-related crises. *Annu Rev Cyberther Telemed.* in press.
- [18] Smith MJ, Bornheimer LA, Li J, Blajeski S, Hiltz B, Fischer DJ, Check K, Ruffolo M. Computerized clinical training simulations with virtual clients abusing alcohol: Initial feasibility, acceptability, and effectiveness. *Clin Soc Work J.* 2021;49(2):184-196. doi: <https://doi.org/10.1007/s10615-020-00779-4>
- [19] Kizakevich PN, McCartney ML, Nissman DB, Starko K, Smith NT. Virtual medical trainer - Patient assessment and trauma care simulator. *Stud Health Technol Inform.* 1998;50:309-15. doi: <https://doi.org/10.3233/978-1-60750-894-6-309>

- [20] Brown WJ, Tortorella RAW. Hybrid medical simulation – A systematic literature review. *Smart Learn Environ.* 2020;7:16. doi: <https://doi.org/10.1186/s40561-020-00127-6>
- [21] Wang S, Lilienfeld SO, Rochat P. The uncanny valley: Existence and explanations. *Rev Gen Psychol.* 2015;19(4):393-407. doi: <https://doi.org/10.1037/gpr0000056>

The Cyberpsychology of Deception: A Mini Review of the Psychological Factors Influencing Scam Compliance

David LACEY^{a,1}, Andrew CAMPBELL^b, Sigi GOODE^c and Brad RIDOUT^b

^a*The University of Sydney, University of Queensland and IDCARE*

^b*Cyberpsychology Research Group, The University of Sydney*

^c*Research School of Management, Australian National University*

ORCID ID: David Lacey <https://orcid.org/0000-0001-9328-1400>

Abstract. Despite extensive research in addressing technical aspects of cybercrime, there is a significant gap in understanding how psychological factors impact scam compliance in differing contexts. This review evaluates studies from diverse disciplines, including psychology, criminology, and behavioral science, to identify key personality, cognitive, emotional and social factors that influence victimization. Covering a broad spectrum of fraudulent activities—including investment fraud, relationship scams, mass marketing fraud and phishing—the review aims to provide an overview of recent empirical research focusing specifically on data collected from actual scam victims. The review applied the PRISMA-P methodology to systematically search and screen literature from multiple databases to identify 18 empirical studies. Findings revealed that personality traits such as impulsivity and trust, cognitive factors such as authority bias, and other emotional and social risk factors are recurrent themes found to influence scam compliance. However, there is considerable variability in research methodologies, scam contexts, and reporting of results. This variability underscores the need for more detailed, context-specific investigations into different scam types as the psychological factors that influence scam compliance differs by scam type and context. The review concludes with recommendations for future research, emphasizing the importance of examining specific scam contexts and improving study designs to better understand scam compliance.

Keywords. Scam compliance, deception, persuasion, personality, cognition, cyberpsychology.

1. Introduction

Cybercrimes impact millions of people across the world each year. According to Cybersecurity Ventures, the global annual cost of cybercrime is predicted to reach \$9.5 trillion USD in 2024 [1]. Scams and crimes of deception are a subset of cybercrimes where the individual concerned has been deceived, and having been deceived, is compliant to the demands of the scammer. There has been considerable recent research on cybercrimes, scams and related threats across many disciplines including computer science, psychology, criminology, organizational science, and accounting. Many papers have oriented their work on the technical and cybersecurity aspects of cybercrime [2], but more recently researchers have ventured into considering the role of the victim in these crimes to shed some light on what may influence an individual's susceptibility to believing and acting on scam stimuli.

Scams and deception crimes are attacks on a person's cognition. Cognition is the mental process of acquiring knowledge and understanding through thought, experience and senses which has evolved over many thousands of years [3]. It is both a conscious and sub-conscious process by which knowledge is accumulated and engaged to allow us to perceive, recognize, conceive and reason [4]. Cognition is influenced by personality

¹ Corresponding Author: David Lacey, dlace6839@sydney.edu.au

traits (stable cognitive and behavioral patterns consistent in various contexts), emotional factors (which vary from moment to moment) and social factors (which vary across the lifespan). It includes our ability to process information, whether visual or auditory, and allows us at a high level to tap into four processing components: perception, working memory, decision-making, and action [5]. Each of these components is influenced by short-term and long-term cognitive factors, such as workload, stress, personality, and vigilance. Further, each of these components and their influences is a significant area of research in its own right. For example, personality research has occupied some interest by many researchers within the fraud context [6]. More recently, researchers have found support for behavioral predispositions as being statistically significant vulnerabilities for specific scam contexts, such as high scores on measures of impulsiveness and trust propensity attributes as compliance predictors in relationship scams [7]. Understanding scam victims' personality, cognition profiles, and their emotional and social context, can help us understand what influences a person to believe deceptive communication from a scammer, leading to scam compliance [8].

There have been a small number of reviews into cybercrime and scam victimization trends, however many have focused on age as the key risk factor [9,10], or on specific scam types [11]. In 2019, Norris et al. [12] undertook a systematic review of psychologically-based literature related to online fraud vulnerabilities and found that despite a growing body of research, a limited number of studies able to identify specific psychological factors and processes associated with increased susceptibility to victimization. Most empirical studies they identified focused on the 'phishing' context, using simulations with student or employee participants, who were often pre-warned about the nature of these tests of scam compliance. Forewarnings and pre-existing awareness have been shown to alter an individual's behavior in simulated experiments [13,14], so this is an inherent weakness when investigating crimes of deception. In contrast, studies into other scam contexts such as relationship scams have focused on comparing the psychological profiles of actual scam victims to non-victims [11].

This mini review therefore aims to summarize scientific research findings across a range of disciplines and scam types that have examined the psychological influences on scam compliance and susceptibility, focusing on studies involving real victims of crimes of deception. This provides a timely and concise overview of the current evidence regarding the psychological factors that have been identified from actual victims as playing a role in cybercrime victimization and identify any gaps in the literature.

2. Methods

2.1. Search Strategy

This mini-review was performed using the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) approach [15]. The goal was to identify the scientific studies over the last decade that empirically explore psychological influences of scam susceptibility from a victim perspective. Given the cross-disciplinary nature of the subject, the databases searched covered multiple disciplines, including computer science, psychology, law, criminology, medicine and health scientific databases (PsycINFO, Scopus, Web of Science, ProQuest, Science Direct, Wiley Online Library, Sage, PubMed, ProQuest Dissertations & Theses Global, ACM Digital Library, Informatit Online, IEEE Xplore, and EBSCOhost).

Searches were conducted in December 2023 and included only peer-reviewed English-language scientific papers, conference proceedings, book chapters, books, and academic theses, published between 2013 and 2023 inclusive. This period was selected because it has witnessed the greatest escalation of reported scam impacts on the community [16]. The search syntax rules varied slightly across databases based on their functionality, but in each case studies were identified for screening if the title, abstract or key words were found to include the term scam* (including scams, scamming, scammer) or social engineer* (including social engineering and social engineers), and either behav* (including behaviour, behaviours, and behavioural along with US spelling equivalents), persua* (including persuasion, persuasive or persuade), or psycholog* (including psychology and psychological).

2.2. Screening Process

Figure 1 shows the initial search syntax across the thirteen databases yielded 4,690 papers. Google Scholar was also searched separately, resulting in a further 14 papers for screening. After removing duplicates, 2,808 were screened at the title and abstract level reducing the number of eligible papers to 2,116. A further detailed examination at full-text level to apply the requirement to only include empirical research involving actual victims resulted in a total of 18 research papers included in the review.

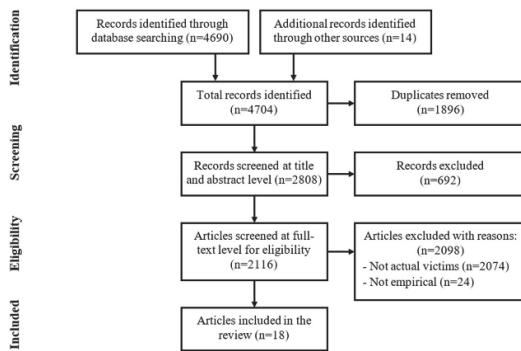


Figure 1. PRISMA flow diagram.

3. Results

3.1. Study Characteristics

One-third of the included papers were from the United States (n=6) and one-third were from the United Kingdom (n=6), however one each of these included participants from other countries. The remaining six papers were from Canada, Australia, China, Japan, Malaysia, and the Netherlands. There was a relatively steady volume of studies published over the time period (between one and three each year), despite the rapid acceleration of victimisation over that period. The exception was 2021 when no studies were published (possibly due to COVID-19 interruptions), but this was made up for in 2022 with four studies published in this year. Scientific studies published in psychology (39%) and criminology (33%) disciplines accounted for most of the eligible publications, followed by medicine and health (17%) and consumer research (11%). There were no eligible papers from computer science, despite the inclusion of computer science databases in the search strategy.

3.2. Study Participants and Design

Most studies (n=13) used quantitative surveys to compare scam victims with nonvictims. The remaining studies used qualitative methods: victim interviews (n=5) and content analysis of victim narratives posted online (n=1). Participant demographics varied widely across the studies. Several studies focused on older adults (n=6), while studies with larger samples sizes drew upon random samples that aimed to be representative of the broader population. Other populations of interest with young people aged 18-25, eBay and dating site users, and individuals with acquired brain injury. Some studies did not collect or disclose details on their participants, other than stating that they were victims of cybercrime.

3.3. Study Findings

Table 1 shows the characteristics of the included studies and the psychological factors that were found to significantly influence scam compliance. Two-thirds of the studies identified personality factors (n=12), with impulsivity and trust the most common. Almost as many studies identified emotional factors (n=11), including emotional

instability, depression or financial desperation. Cognitive factors (e.g. cognitive ability, authority bias) and social factors (e.g. loneliness) were identified in 6 studies each. A variety of scam types were studied, however half of studies either combined results for multiple different scam types (n=5) or did not disclose the scam type (n=4). Of the studies that focused on a single scam type, relationship scams (n=4) were the most common.

Table 1. Summary of study findings regarding psychological factors influencing scam compliance.

Authors and location	Scam Type	Design	Participants	Personality factors	Cognitive factors	Emotional factors	Social factors
Whitty 2018 (UK) [17]	Relationship	Survey of victims (n=200) vs non-victims (n=9,466)	Random sample	Impulsivity, addictive disposition, lower kindness, trustworthiness			
Whitty 2019 (UK) [8]	Various	Survey of victims (n=1,057) vs non-victims (n=10,723)	Random sample	Impulsivity (urgency, sensation seeking), addictive disposition			
Whitty, 2020 (UK) [18]	Various (reported separately)	Survey of victims (n=1,057) vs non-victims (n=10,723)	Random sample	Impulsivity, neuroticism, internal locus of control		Emotional instability	
Modic 2022 (UK) [19]	Online auction fraud	Study 1: Survey of victims (n=327) vs non-victims (n=2,374) Study 2: Survey of victims n=81)	eBay users	Impulsivity, modesty, creativity, gentleness		Scarcity	
Wen et al. 2022 (China) [20]	Various	Interviews with scam victims (n=17)	18–25 years	Impulsivity, trust	Poor risk perception	Emotional instability, motivated by high value monetary gain	
Fischer et al. 2013 (US) [21]	Not disclosed	Survey of victims (n=26) vs nonvictims (n=77)	Not disclosed	Trust		Motivated by high value monetary gain	
Ueno et al. 2022 (Japan) [22]	Various	Survey of victims (n=56) vs non-victims (n=99)	60+ years	Over-confidence			
Olivier et al. 2015 (UK) [23]	Mass marketing fraud	Interviews with victims (n=3)	60+ years	Trust		Emotional vulnerability	Loneliness
Gould et al. 2023 (Australia) [24]	Relationship	Interviews with scam victims (n=7) and close others (n=6)	Individuals with acquired brain injury	Trust	Lower cognitive ability	Social isolation, loneliness	
Wilson et al. 2023 (Malaysia) [25]	Various	Interviews with victims and non-victims	Not disclosed	Trust, greed		Financially desperation	
DeLiema et al. 2020 (US) [26]	Investment	Survey of victims (n=214) vs non-victims (n=813)	Investors (M=70 years)	Risk-taking, greed/materialism		Motivated by high value monetary gain	
Judges et al. 2017 (Canada) [27]	Not disclosed	Survey of victims (n=51) vs non-victims (n=100)	60-90 years without cognitive impairment	Lower conscientiousness, lower honesty-humility	Lower cognitive ability		
DeLiema 2015 (US) [28]	Not disclosed	Survey of victims (n=460) vs non-victims (n=11,524)	50+ years, not living in nursing home		Higher cognitive ability	Stressful life event	Widowed (not loneliness)
Wang et al. 2022 (US) [29]	Relationship	Content analysis of online narratives by victims	Not disclosed		Authority bias, reciprocity bias	Emotional instability	Social isolation
DeLiema et al. 2023 (US) [30]	Various	Survey of victims (n=307) vs non-victims (n=1,040)	Scam reporters		Authority bias	Financially desperation	Loneliness

<i>Lichtenberg et al.</i> 2013 (US) [31]	Not disclosed	Survey of victims (n=220) vs non-victims (n=4,240)	50+ years	Depression	Loneliness
<i>Buchanan et al.</i> 2014 (UK) [32]	Relationship	Study 1: Survey of victims (n=137) vs non-victims (n=716) Study 2: Survey of scam victims (n=325) vs non-victims (n=42)	Study 1: Dating site users Study 2: Relationship scam Support website	Idealization of romantic partners	
<i>Jansen et al</i> 2016 (Netherlands) [33]	Phishing/malware	Interviews with victims (n=30)	Reported to police	Nil (everyone is susceptible)	

4. Discussion

4.1. Personality Factors

Impulsivity was the most commonly identified personality risk factor for scam victimization, and was a significant predictor in four of the five large scale quantitative studies [8,17,18,19]. The exception was DeLiema (2015) [28] focused on people aged over 50 years, which is not surprising given that impulsivity is found to decrease with age [34]. Among the qualitative interviews with victims, being high in trust was the most identified factor, and was commonly self-reported by victims to explain why they believed the scammer's deception.

While the lack of delineation between results for different scam types in many studies made it difficult to determine trends in specific contexts, most of the other personality factors identified appear to be related to increased risk-taking (e.g., addictive disposition, over-confidence, high internal locus of control, greed, and lower consciousness), with some of these found to be directly related to scams that exploit the victim's desire for financial gain such as investment or mass-marketing scams.

4.2. Cognitive Factors

Only one-third of included studies identified cognitive factors that influence scam compliance. Authority bias, where individuals comply with a scammer's deception due to how 'official' a demand for compliance seems, was a significant predictor of victimization in DeLiema's 2023 study [30], especially in threat-based scams. Authority bias was also a factor identified in Wang et al.'s study of narratives by relationship scam victims [29], in that scammers would produce 'official' looking documents to legitimize their financial requests. Reciprocity bias was also identified, with scammers creating scenarios where victims would feel a strong inclination to return favours.

There were contradictory results regarding the influence of overall cognitive ability. Lower cognitive ability was identified as a predictor of scam compliance in Judges et al.'s study [27] of individuals aged 60-90 years (but without cognitive impairment), and in Gould et al.'s [24] interviews with individuals with acquired brain injury. However, DeLiema's large scale study [28] of individuals aged over 50 (but not living in nursing homes) found that scam victimization was associated with higher cognitive ability. The author suggested this may be explained by the victims in their study being younger and more educated than the non-victims, or that victims who are older and/or with lower cognitive ability may be less likely to report their victimization. The only study that focused on young people found that scam victims identified poor risk perception as a cognitive factor in their scam compliance [20].

4.3. Emotional Factors

Emotional instability emerged as a recurrent theme across various scam types, which some authors suggested made individuals more reactive to the visceral cues used by

scammers. Other emotional factors identified were driven by external pressures, such as financial desperation and stressful life events, which can cloud judgment and increase scam compliance. Depression and emotional vulnerability were found to be significant factors influencing scam victimization in older individuals [23,31].

There were also emotional factors specific to particular scam contexts, such as the desire of scarcity in falling for online auction scams [19], the idealization of romantic partners in relationship scams [32] and feeling motivated by lure of a high-value monetary gain in mass marketing and investment scams [20,21,26]. Overall, emotional factors appear to play a critical role in diminishing an individual's capacity to critically evaluate scam-related information.

4.4. Social Factors

All social factors identified in the included studies related to aspects of social isolation. DeLiema [28] found that compared to married respondents, those who are widowed were more likely to be victimised, despite loneliness not being a statistically significant factor. However, loneliness was frequently identified as a contributing factor in other studies, particularly among older victims [23,31] and individuals with acquired brain injuries [24]. In addition to lacking social support networks that could otherwise help them avoid scams, interviews revealed that for some older people, the ritual of interacting with and dispatching money to scammers can be exciting for the victim and provide them with rare social interaction [23].

4.5. Study Quality

Methodologically, most quantitative studies statistically compared victims to nonvictims, which allowed for researchers to elucidate scam compliance sensitivities. However, studies inconsistently defined the period and number of prior scam events, and the time elapsed between scam engagement and research participation. The latter is particularly critical for interview studies, as longer gaps can lead to memory distortion or rationalization by victims, thereby affecting the reliability of the data collected. Studies with shorter intervals between scam engagement and research participation are likely to capture more accurate reasons and aspects of scam compliance.

Studies also inconsistently characterized the specific scam context. Many studies either combined results from different scam types or did not disclose the specific scam type. This lack of specificity is problematic, as scam compliance can vary greatly among scam types. The Australian Competition & Consumer Commission (ACCC) [16] notes that compliance with phishing scams can occur within minutes, whereas compliance with relationship and investment scams often unfolds over weeks, months, or even years [35]. This distinction is crucial, yet it was often overlooked, leading to a generalized understanding that may not accurately reflect the nuanced reality of different scam types.

Finally, there was a stark difference between the contexts examined in the studies and commonly reported scam scenarios. For instance, nearly 90% of investment fraud cases reported to the ACCC [16] involve cryptocurrency, with victims typically starting with a small investment that escalates as scammers show fake returns over time. This scenario, however, was not captured in the studies reviewed, indicating a knowledge gap that needs to be addressed through more context-specific research.

5. Conclusion

This mini review provides a concise overview of the current empirical evidence of the psychological factors contributing to scam compliance, revealing insights from studies of actual victims. The recurring identification of personality traits such as impulsivity, and cognitive biases like authority bias, illustrates how deeply ingrained psychological patterns can predispose individuals to fall victim to scams. The emotional and social factors identified—ranging from emotional instability and financial desperation to social isolation—further emphasize the vulnerability of certain populations, particularly older adults and those experiencing significant life stressors.

Three main implications for future research emerge from this review. First, future research should focus on examining specific scam types, rather than aggregating results across diverse scams. Understanding the unique psychological, cognitive, emotional, and social factors at play in different types of scams will provide more actionable insights for prevention and intervention strategies tailored to not only different scams, but different psychological profiles. Given the varying influence of factors like impulsivity, trust, and cognitive ability across different age groups and scam contexts, future studies should consider age-specific vulnerabilities and how these interact with different scam types.

Second, there is a need for more consistent and rigorous methodological approaches, particularly in terms of defining the time between scam engagement and participation in research. Reducing this gap will help ensure that data collected reflects the immediate reasons for scam compliance, rather than rationalizations formed over time.

Finally, researchers should strive to align their studies more closely with the evolving landscape of scam tactics, such as the growing prevalence of cryptocurrency scams. By keeping pace with these developments, research can remain relevant and provide timely insights that can inform policy and protective measures.

In conclusion, while the existing body of research offers valuable insights into the psychological factors influencing scam compliance, addressing the gaps identified in this review and refining study designs will be crucial for advancing a more nuanced understanding of the diverse factors contributing to scam victimization in different contexts, and developing more effective interventions.

References

- [1] Morgan S. Cybercrime to cost the world \$9.5 trillion USD annually in 2024 [Internet]. Northport (NY): Cybersecurity Ventures; 2023 Oct 25 [Cited 2024 Aug 19]. Available from: <https://cybersecurityventures.com/cybercrime-to-cost-the-world-9-trillion-annually-in-2024>.
- [2] Wu L, Peng Q, Lemke M. Research trends in cybercrime and cybersecurity: A review based on web of science core collection database. *Int J Cybersecurity Intell Cybercrime*. 2023;6(1):5-28.
- [3] Tomasello M. *The cultural origins of human cognition*. Cambridge (MA): Harvard University Press; 2009.
- [4] Burton L, Westen D, Kowalski R. *Psychology: Australian and New Zealand Edition*. Milton (AU): Wiley; 2015.
- [5] Modic D, Anderson R, Palomäki J. We will make you like our research: The development of a susceptibility-to-persuasion scale. *PloS One*. 2018 Mar 15;13(3):e0194119.
- [6] Digman JM. Higher-order factors of the Big Five. *J Pers Soc Psychol*. 1997 Dec;73(6):1246.
- [7] Whitty MT. Predicting susceptibility to cyber-fraud victimhood. *J Financ Crime*. 2019 Jan;26(1):277-92.
- [8] Modic D. Willing to be scammed: How self-control impacts internet scam compliance [dissertation]. Exeter (UK): University of Exeter; 2012.
- [9] Burnes D, Henderson Jr CR, Sheppard C, Zhao R, Pillemer K, Lachs MS. Prevalence of financial fraud and scams among older adults in the United States: A systematic review and meta-analysis. *Am J Public Health*. 2017 Aug;107(8):e13-21.
- [10] Shang Y, Wu Z, Du X, Jiang Y, Ma B, Chi M. The psychology of the internet fraud victimization of older adults: A systematic review. *Front Psychol*. 2022 Sep 5;13:912242.
- [11] Coluccia A, Pozza A, Ferretti F, Carabellese F, Masti A, Gualtieri G. Online romance scams: relational dynamics and psychological characteristics of the victims and scammers. A scoping review. *Clin Pract Epidemiol Ment Health*. 2020;16:24.
- [12] Norris G, Brookes A, Dowell D. The psychology of internet fraud victimisation: A systematic review. *J Police Crim Psychol*. 2019 Sep 15;34:231-45.
- [13] Epley N, Gilovich T. When effortful thinking influences judgmental anchoring: differential effects of forewarning and incentives on self-generated and externally provided anchors. *J Behav Decis Mak*. 2005 Jul;18(3):199-212.
- [14] Wood W, Quinn JM. Forewarned and forearmed? Two meta-analysis syntheses of forewarnings of influence appeals. *Psychol Bull*. 2003 Jan;129(1):119.
- [15] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*. 2021 Mar 29;372.
- [16] Australian Competition and Consumer Commission (ACCC). Targeting scams: report of the ACCC on scams activity 2022. [Internet]. Canberra (AU): ACCC; 2022 [Cited 2024 Aug 19]. Available from: <https://www.accc.gov.au/about-us/publications/serial-publications/targeting-scams-reports-on-scamsactivity/targeting-scams-report-of-the-accc-on-scams-activity-2022>.
- [17] Whitty MT. Do you love me? Psychological characteristics of romance scam victims. *Cyberpsychol Behav Soc Netw*. 2018 Feb 1;21(2):105-9.
- [18] Whitty MT. Is there a scam for everyone? Psychologically profiling cyberscam victims. *Eur J Crim Policy Res*. 2020 Sep;26(3):399-409.
- [19] Modic D. Do not distract me while I am winning this auction: The psychology of auction fraud. In: Hanoch Y, Wood S, editors. *A fresh look at fraud*. London (UK): Routledge; 2022. p. 144-166.

- [20] Wen X, Xu L, Wang J, Gao Y, Shi J, Zhao K, Tao F, Qian X. Mental states: A key point in scam compliance and warning compliance in real life. *Int J Environ Res Public Health*. 2022 Jul 7;19(14):8294.
- [21] Fischer P, Lea SE, Evans KM. Why do individuals respond to fraudulent scam communications and lose money? The psychological determinants of scam compliance. *J Appl Soc Psychol*. 2013 Oct;43(10):2060-72.
- [22] Ueno D, Arakawa M, Fujii Y, Amano S, Kato Y, Matsuoka T, Narumoto J. Psychosocial characteristics of victims of special fraud among Japanese older adults: A cross-sectional study using scam vulnerability scale. *Front Psych*. 2022 Jul 25;13:960442.
- [23] Olivier S, Burls T, Fenge LA, Brown K. "Winning and losing": Vulnerability to mass marketing fraud. *J Adult Protect*. 2015 Dec 14;17(6):360-70.
- [24] Gould KR, Carminati JY, Ponsford JL. "They just say how stupid I was for being conned". Cyberscams and acquired brain injury: A qualitative exploration of the lived experience of survivors and close others. *Neuropsychol Rehabil*. 2023 Feb 7;33(2):325-45.
- [25] Wilson S, Hassan NA, Khor KK, Sinnappan S, Abu Bakar AR, Tan SA. A holistic qualitative exploration on the perception of scams, scam techniques and effectiveness of anti-scam campaigns in Malaysia. *J Financ Crim*. 2023 Sep 27.
- [26] DeLiema M, Shadel D, Pak K. Profiling victims of investment fraud: Mindsets and risky behaviors. *J Consum Res*. 2020 Feb 1;46(5):904-14.
- [27] Judges RA, Gallant SN, Yang L, Lee K. The role of cognition, personality, and trust in fraud victimization in older adults. *Front Psych*. 2017 Apr 13;8:588.
- [28] DeLiema, M. Using mixed methods to identify the characteristics of older fraud victims [dissertation]. Irvine (CA): University of Southern California; 2015.
- [29] Wang F, Topalli V. Understanding romance scammers through the lens of their victims: qualitative modeling of risk and protective factors in the online context. *Am J Crim Justice*. 2024 Feb;49(1):145-81.
- [30] DeLiema M, Li Y, Mottola G. Correlates of responding to and becoming victimized by fraud: Examining risk factors by scam type. *Int J Consum Stud*. 2023 May;47(3):1042-59.
- [31] Lichtenberg PA, Stickney L, Paulson D. Is psychological vulnerability related to the experience of fraud in older adults?. *Clin Gerontol*. 2013 Mar 1;36(2):132-46.
- [32] Buchanan T, Whitty MT. The online dating romance scam: causes and consequences of victimhood. *Psychol Crime Law*. 2014 Mar 16;20(3):261-83.
- [33] Jansen J, Leukfeldt R. Phishing and malware attacks on online banking customers in the Netherlands: A qualitative analysis of factors leading to victimization. *Int J Cyber Criminol*. 2016;10(1):79.
- [34] Chamorro J, Bernardi S, Potenza MN, Grant JE, Marsh R, Wang S, Blanco C. Impulsivity in the general population: a national study. *J Psychiatr Res*. 2012 Aug 1;46(8):994-1001.
- [35] Lacey D, Goode S, Pawada J, Gibson D. The application of scam compliance models to investment fraud offending. *J Criminol Res Policy Pract*. 2020 Apr 22;6(1):65-81.

Singularity: A Non-Dangerous Case

Alexander VOISKOUNSKY^{a,1}

^a*Psychology Department, Moscow State University after M.V.Lomonosov, Mokhovaya
st. 8/5, 103009 Moscow, Russia*

ORCiD ID: Alexander Voiskounsky <https://orcid.org/0000-0002-5213-1366>

Abstract. The concept of technological singularity, often associated with artificial superintelligence surpassing human capabilities, has sparked significant debate. This paper explores the unique position of high-level chess players, particularly grandmasters, as a community potentially exemplifying lifestyles adapted to singularity-related environments. Through a five-year analysis (2019–2023) of 276 game commentaries from leading chess platforms, the study identifies how chess players integrate advanced AI systems, such as chess engines, into their daily routines. The findings reveal that 84% of the analyzed commentaries referenced AI tools positively, indicating widespread adaptation and a perception of these tools as cognitive partners rather than threats. Drawing on Vygotsky's concept of the zone of proximal development, the study argues that interactions with AI can enhance cognitive development and professional competence. The paper concludes by challenging the dystopian view of singularity as a disaster, suggesting instead that it may foster human transformation and growth in specific domains.

Keywords: Technological singularity, artificial intelligence, chess players, cognitive development, zone of proximal development

1. Introduction

In the era of digital technologies cyberpsychology neighbors with various problem areas and fields referring to human beings' connections with electronic technologies and peculiar mediations, such as human-robots, human-neural networks or human-chatbots connections. Connections may be collaborative, or opposing, or any intention in between. A widely discussed theme dealing with the perspectives of the artificial intelligence (AI) is an actual topic to exchange views. Scholars, practitioners, IT and AI experts as well as ordinary people keep passionately discussing the limits of digital technologies' progress: the progress may supposedly result in overrunning human intelligence. That means humans will supposedly lose control over the most valuable decisions made by various AI subsystems – decisions dealing with our personal and social being, business perspectives, financial investments, adoption of new technologies, ecological measures, selection of specialties in demand to learn and develop, etc. No wonder, many polemists actively dislike such a perspective.

While this sort of discussions is not quite new (historical precedents are well-known: for example, anti- machinery Luddites movements in Britain in early 1810s, or disputes pro & contra the newborn cybernetics in the Soviet Union in 1950s), during the last decades certain consequences of the AI development and usage refer mainly on the idea of singularity (often called technological singularity). By singularity Vernor Vinge [1], Ray Kurzweil [2] and their followers mean the way that rapidly developing digital technologies fit more and more close to creation a so called superintelligence – the one that surpasses human intelligence and becomes capable of making super-complex decisions that would never (or maybe only rarely and by chance) come to someone's mind due to all the unexpectedness and/or huge time-consuming period which lies behind such a decision.

Enthusiasts estimate the appearance of artificial superintelligence (ASI), i.e. supercomputer systems connected via supernetworks, to happen around 2040 or 2045. Several ASI projects such as AlphaZero or ChatGPT, as well as some others, have declared themselves in the last few years as really competent instruments, useful both for experts and for ordinary public. While being competent and useful, many projects

¹ Corresponding Author: Alexander Voiskounsky, vae-msu@mail.ru

meet nevertheless protests from prominent specialists, politicians, educators, and general public.

Meetings of natural intellects with AI and ASI have not once been visualized (more correct would be to say: imagined) in theory. But can we carry on serious investigations in the field right now, two decades before 2045s? With this in mind, we need to find first of all the situations close to this or that sort of singularity.

Many people know little and feel neutral regarding the idea of singularity and the particular AI/ASI projects; the others express either delight or fear, i.e. they are positive or negative. The majority have no practice of dealing with AI/ASI, or else they believe (sometimes mistakenly) they lack such experience. Indeed, there are rather few groups being capable to organize daily lifestyles within a situation of technological singularity. To our view, such a group consists of, for example, qualified chessplayers. In the next section we will try to substantiate this responsible point of view.

2. Community of Chessplayers

The aim of the paper is to draw the specialists' attention to the lifestyles of a rather rarely analyzed international community of high-quality chessplayers: grandmasters (GMs) and masters. Moreover, the aim is to start discussing their lifestyles from the standpoint of the perceived closeness to the singularity point. Further we are going to substantiate the position according to their dealing with specialized software and advanced AI systems on a daily basis. That means, certain behavioral patterns inherent between chessplayers are worth to be identified and discussed.

The community in question is multi-layered [3], partly due to the Prof. Elo rating system accepted by the International Chess Federation (FIDE), the governing body of international chess competitions. The number of world-wide chess fans and occasional players is believed to exceed hundreds of thousands. The number of titled chess players: Grandmasters (GM), International Masters (IM), and FIDE masters (FM), are tens of thousands. The top layers are occupied by approximately two thousand grandmasters (as well as rarely – masters), the majority of them – media persons having high reputation. Two-thirds of them (approximately 1000+) players are believed to be active, that is they systematically fight in chess battles and verify their player strength. Besides, their main income is earned by playing chess.

The high-layer community is heavily biased and includes much more males than females; the latter (four or five dozens) too belong to the highest layers. Besides grandmasters, it includes many masters (international or local) and experts who systematically or from time to time participate in chess competitions, help the strongest players on a friendly basis or for salary, teach beginners, comment for chess fans the ongoing competitions and particular games, often coach strong chessplayers as their seconds. Less qualified players organize and judge chess competitions, collect chess-related information and administer chess databases, update chess engines (such as Stockfish or Leela Chess Zero, etc.). It should be noted that grandmasters may as well take upon themselves all these functions.

3. Grandmasters in the zone of proximal development?

The aim of this pioneer paper is to draw the scholars' attention to chessplayers as possible «victims» of singularity, and also to show that they hardly consider themselves being «victims». Moreover, to meet a definitely more competent and knowledgeable agent who takes a responsibility to support your psychic development is in a way a serious bonus and advantage. In Lev Vygotsky' cultural historic theory of psychic development the support of an older partner is called ZPD – a zone of proximal development [4].

Although this idea refers to the child development, in the era of life-long learning the ZPD can be easily transferred to grown-ups too. Thus, a chess engine may be taken as an elder agent that helps chess players of any age to advance his/her competence in chess. Moreover, Vygotsky's followers insist that a ZPD related experience and getting new competencies is an important way to enhance psychic development [5].

4. Results and Discussion

An analytical method has been used at a starting point: it includes reviewing the comments issued by observation authors and checking whether commentators mentioned computer supporting instruments: either specific chess engines (such as Stockfish) or simply named supporters as a computer or a software program. In fact, observers and players often refer to supporters – particular or general. This was selected to be the most useful material to carry analysis.

The games of top-level players, especially rated within the first hundred, are widely published and thoroughly discussed with detailed comments. The analysis consisted of the following. During 5 years (2019 – 2023) the author followed three popular chess media (namely: chess.com, fide.com and chesspro.ru) and selected comments to all the games with detailed explanations and interpretations (oriented to chess fans); comments have always been prepared by grandmasters. The further analysis included explicit fixations dealing with references about recommendations of a chess engine. Thus in case the chessplayers' fixations were numerous and positive, we might state that the members of the community of top-level chessplayers may be considered among the first to adapt within the groups of a singularity related lifestyle. Otherwise, we would state that the authors of comments have not yet adapted to a new lifestyle with a singularity burden.

Results show that the chessplayers adapted to the changes in their professional life. Indeed, 276 comments have been analyzed and only 43 missed any references to a chess engine. That means, the majority of observers (84%) seem to adapt to the situation when a chess engine or a computer are highly knowledgeable in the moves on a chessboard. They do not seem to consider themselves burdens of singularity times.

Moreover, previously some critical voices were heard saying that although the machine wins, it demonstrates a rather boring game, neglecting "sharp" continuations, since the results of such moves are too difficult to calculate even for an unbelievably fast computer. Nowadays such a critics turns out to be old-fashioned, if we listen to the words of Garry Kasparov, the greatest expert in the field: «I admit that I was pleased to see that AlphaZero had a dynamic, open style like my own. The conventional wisdom was that machines would approach perfection with endless dry maneuvering, usually leading to drawn games. But in my observation, AlphaZero prioritizes piece activity over material, preferring positions that to my eye looked risky and aggressive» [6]. This is a high opinion indeed.

The stated hypothesis looks to be correct: chessplayers have indeed adapted their lifestyle to exist within an idea of singularity. What looks like dependence, is not: chessplayers use possible bonuses of singularity to transform themselves mentally. And the idea of singularity as a sort of disaster needs to be revised and reconstructed.

References

- [1] Vinge V. Signs of the singularity. IEEE Spectrum. 2008;45(6):76–82. doi:<https://doi.org/10.1109/mspec.2008.4531467>.
- [2] Kurzweil R. The singularity is near. New York: Viking Books; 2005.
- [3] Desjarlais RR. Counterplay: an anthropologist at the chessboard. Berkeley–Los Angeles–London: University of California Press; 2011.
- [4] The Vygotsky anthology: a selection from his key writings. Barrs M, Richmond J, editors. London: Routledge; 2024.
- [5] Cole M. Cultural psychology: a once and future discipline. Cambridge, MA: Belknap Press; 1998.
- [6] Kasparov G. Chess, a drosophila of reasoning. Science. 2018;362(6419).

Outlining Theoretical Underpinnings of Game Transfer Phenomena

Angelica B. ORTIZ DE GORTARI^{a,1}, Andrzej CUDO^b, Julio C. LLAMAS-ALONSO^c
and Arturo LERNER^d

^a*Centre for the Science of Learning & Technology (SLATE), University of Bergen,
Norway*

^b*Department of Experimental Psychology, The John Paul II Catholic
University of Lublin, Poland*

^c*Faculty of Psychology, Universidad La Salle Bajío, Campus Campestre, México*

^d*Private Practice; former Chairman of the Israeli Society of Addiction Medicine;
Former Head of the Department of Dual Disorders (Mental Illnesses and
Addictions), Lev-Hasharon Mental Health Medical Center, Pardessya, Israel*

ORCiD ID: Angelica B. Ortiz de Gortari <https://orcid.org/0000-00029172-9264>

Abstract. Research has investigated the impact of video game play on cognitive processes and perceptions during and after gaming. These studies have primarily focused on cognitive functions such as perception, attention, working memory, and skill transfer. However, little attention has been paid to the short-term effects of gaming. Research on Game Transfer Phenomena (GTP) has emerged as a comprehensive, multi-modal approach to examining the temporary changes in sensory perception, cognition, and self-agency directly linked to video game experiences. GTP manifestations involve suddenly re-experiencing images, sounds, haptics, and perceptual distortions of surroundings, objects, and the body perception after gaming. These experiences typically happen while being awake and are triggered by internal (thoughts) or external stimuli associated with the game. The prevalence of GTP ranges from 74.8% to 96% (in a 12-month period). However, GTP occurrences are mostly infrequent. Most GTP cases are reported by gamers without a history of substance use or mental disorders. After reviewing the literature on GTP (32 studies, N>8,000) and related phenomena, seven theories were formulated, partially supported by the current knowledge about GTP. The postulated theories were categorized into three groups: 1) Vulnerability factors include susceptibility to cognitive and sensory phenomena, hypersensitivity, low working memory capacity, and tendency to experience mind-wandering. 2) Mechanisms underlying GTP involving priming mechanisms, associative learning, selective attention to game-related cues, predicting errors, disinhibition of visual processors, neurotransmission dysregulations, sensory sensitivity, and neural adaptations. 3) Risk involving GTP due to their potential disruptive nature and its contribution to the symptomatology of gaming disorder.

Keywords. Game transfer phenomena, video games effects, involuntary cognitions

1. Introduction

GTP involve the interplay of physiological, perceptual and cognitive mechanisms. Recent efforts have focused on identifying the underlying processes and potential outcomes by distinguishing between endogenous and exogenous manifestations as well as environmental stimuli acting like triggers [1]. The GTP framework [2] has been supported by cross-cultural validation studies [3-5]. This paper aims to outline theories that can provide a foundation for the future investigation of GTP.

2. Methods

After reviewing the GTP literature and related phenomena, seven theories were formulated as statements regarding vulnerability factors, underlying mechanisms, and the etiology of GTP. Studies on GTP included 32 studies (N>8,000). Most of the studies utilized cross-sectional designs and gathered data through surveys, interviews, diaries, online forums, and experimental tasks.

¹ Corresponding Author: Angelica B. Ortiz de Gortari, angelica@gametransferphenomena.com

3. Results

Individuals who are more susceptible to involuntary cognitive and sensory phenomena, are more likely to experience GTP: GTP are considered involuntary due to their unpredictable and uncontrollable nature [2]. GTP has been placed on the spectrum starting from common experiences up to symptoms of disorders causing distress and impairment. GTP are associated with cognitive phenomena resulting from exposition or encounters with stimuli (e.g., earworms or music imagery) [6] or the resurfacing of involuntary memories such as autobiographical or mind-popping (MP; brief and involuntary recollections of semantic knowledge in the form of words, images, and music), perceptual disturbances (e.g., MUSEQ; CAPS), unusual sleeping experiences [1, 7] and dissociations (e.g., DES) [8].

GTP are of a dissociative nature as they can disrupt the track of thoughts, perception, motor control and body representation: Normal dissociation involves instances of mind-wandering (daydreaming), task-unrelated thoughts, deep absorption in reading or watching a movie, and imaginative play [9]. Dissociation as a disorder imply the disruption in the usual integration of various aspects of consciousness and behavior [10]. GTP has been classified among dissociative phenomena (see review of [11] and has been associated with dissociations [8]. Gamers have also reported corporeal phenomena involved in dissociations (e.g., vestibular adaptations, perceptual changes in physical characteristics [1, 12]. GTP can be dissociative when individuals become deeply engaged in mental imagery and involuntary mental actions while recalling or envisioning gameplay after playing, particularly when feeling sensations of derealization and depersonalization triggered by game-related stimuli [1]. However, GTP usually occur in a non-dissociative manner, allowing episodic game content to emerge without interrupting concurrent activities [6, 13, 14]. Similarly, Hallucinogen Persistent Perception Disorder (HPPD) symptoms which share phenomenological similarities with some sensory experiences in GTP (e.g., retinal sensations) appear to be better explained by shift of focus from exogenous to endogenous phenomena than mimicking the alteration in states of consciousness typically accompanied by the substance use [15]. Although, GTP can also occur during the transition from wakefulness to sleep, resembling hypnagogic phenomena that do not necessarily involve declarative memory systems, as indicated by experiments using games to explore the link between wake states and dreams [16]. Since GTP typically manifests when awake and are elicited by gaming-related stimuli, comprehending the risk of GTP causing disruption is crucial, as distractions can lead to mishaps during engagement in activities that require full attention.

Gamers with low working memory capacity (WMC) and a tendency to mindwandering are more likely to experience GTP: WMC allows maintaining information in memory even when facing distractions [17]. Gamers with a low WMC can easily shift their attention away from their current activity and focus on internal thoughts or external stimuli related to the game. In this context, connections between GTP and mindwandering were found [18]. Research indicates that GTP like daydreaming tend to occur during periods of low cognitive demand as when performing routine tasks (e.g., walking) [14]. Gamers may activate memory representations associated with a previously played game while mind wandering as the focus of attention can be shifted to the most activated memory representation, which can lead to its awareness. The activation of specific hubs of the Brain's Default Mode Network (DMN), particularly the Posterior Cingulate Cortex [19] has been suggested to be associated with GTP [20]. Hyperactivity in the DMN has been observed in individuals experiencing unusual thoughts and hallucinations, such as those diagnosed with schizophrenia or at risk [21]. This increased activity may also be present in gamers who report high frequency of GTP, especially in those exhibiting positive schizotypal traits, as these traits have been found to predict GTP [20]. Therefore, examining the relationship between DMN activity and the extent of positive schizotypal traits could offer insights into whether frequent GTP could serve as an early indicator of vulnerability to developing psychosis, as seen in clinical cases [22]. Furthermore, GTP has been linked to cognitive lapses, including those caused by attention errors and low mindfulness [8].

GTP are the result of priming or associative learning leading to reactivity toward game-related cues: Long-term priming mechanisms observed in MP that can be triggered by previous encounters with similar stimuli (repetition priming) or with related content (associative priming) [23] may be involved in GTP [18]. GTP could also be explained by associative learning (classic conditioning) observed in experiments where auditory hallucinations have been induced by pairing visual and auditory stimuli [24]. Through repeated exposure to visual, auditory, and tactile cues in gaming, strong associations can be formed with specific actions, emotions, or states. These enduring connections in the brain could lead to automatic triggers when encountering similar cues outside the gaming environment, potentially resulting in GTP. Studies have found that gamers can recognize objects related to gaming which activates their GTP [12, 14]. The game-related cues may trigger memories of past gaming experiences and stand out due to the expectation of rewards or punishment from games [25]. GTP usually occur shortly after playing but also days, months, or even years later when engaging in an activity related to a particular game [26]. This suggest that repetitive exposure to words, images, and concepts associated with the game may not immediately disappear, but remain active in the mind for some time after playing [18]. Some gamers with poor impulse control seem to struggle to resist impulsive actions toward game-related stimuli and automatically approach objects or raise their arms [2]. Motor impulsivity has been shown associations with GTP [14], and negative correlations have been found in goal maintenance, inhibition, and adjournment [27]. However, no correlations have been found between GTP and the Go/No-Go task [28].

Sensory GTP experiences are explained by prediction errors and neural adaptations involving hypersensitivity: The predictive processing framework supports the idea of the role of top-down processes in hallucinations [29]. The connection between cognitive processes and sensory perception is evident in the context of GTP. The same game has led to GTP for many gamers, although with different manifestations. Some expect to see maps when searching for an address, while others see them in the corner of their eye. Some are taken aback by the non-square shape of trees, unlike the cubic ones in Minecraft, while others perceive trees as distorted.

Engaging in repetitive tasks or prolonged exposure to sensory stimuli can lead to neural adaptations [2] facilitated by fatigue [28]. Seeing game images recurrently resemble afterimages in the retina, but due to their prolonged duration (hours, days) and color preservation, they appear to be positive afterimages indicating involvement of the brain [12]. Some games contain epileptogenic precipitators such as flashing lights and specific figure patterns [30]. Sensory hypersensitivity have correlated with GTP (e.g., sensitivity to bright light, pain, textures) [7]. Symptoms of HPPD, which bear similarities with visual GTP, have been linked to photosensitivity and perceptual disturbances, including sensitivity to tinnitus, migraine with aura, vertigo synesthesia [31].

Visual GTP experiences are a result of disinhibition of visual processors and neurological: Analysis of visual experiences via qualitative studies and clinical cases [32, 33] suggests that GTP exhibits clinical features and similarities to Flashbacks (FB or Type I) and Hallucinogen Persisting Perception Disorder (HPPD or type II) derived from the use of psychedelics and consciousness or mind expanding substances. This includes the re-emergence of the perceptual disturbances experienced during substance use, in the absence of the substance (DSM-5-TR) (For review see [10]). Likewise, HPPD patients, those reporting GTP, while not exclusively, re-experienced sensory perceptions in dark conditions, such as when attempting to fall asleep, moving from well-lit environments, or in conditions where ambiguous stimuli allow re-interpretations [12, 33]. The perceived location of these images may be internal, within the mind's eye, or external, overlaying objects associated with the game. The disinhibition hypothesis appears to explain some of the HPPD symptoms (e.g., afterimages, palinopsia, and halos) which may be caused by a lack of inhibitory neurons to control the activity of pyramidal neurons after a visual stimulus [31]. This disinhibition may be linked to dysfunction of cortical serotonergic inhibitory interneurons with GABA-nergic outputs processes that are implicated with sensory filtering mechanisms of unnecessary stimuli in specific and certain brain areas [34]. A faulty sensory gating process could be linked to the

development of visual GTP experiences, enabling the ongoing central processing of visual perceptions after the image has disappeared from the visual field [35].

GTP contributes to symptoms of GD, and can be a sign of GD symptomatology and dysregulation of neurotransmission: GD involves excessive, uncontrollable and dysfunctional gaming, while GTP implies temporal changes in perceptions, cognition, and behaviors [36]. Both GD and GTP have been associated with a range of conditions, including anxiety and depressive disorders, impulsivity, ADHD, OCD, dissociative disorders, and substance use [42, 46-48]. It is noteworthy that while the co-occurrence of GD and mental disorders is common, most gamers with GTP do not have a mental disorder [20]. Nevertheless, having a disorder does increase susceptibility to GTP [37]. An abnormal activity in the release of dopamine has been connected to gaming and GD, hallucinations in schizophrenia and perceptual disturbances in HPPD [34, 38] and may play a role in GTP. Studies have consistently shown that GD can predict GTP [28, 39]. Associations between playing time and GTP has been established, but the findings have not always been substantiated [40, 41]. The individual characteristics appear to play a more significant role in the susceptibility to GTP than excessive playing [20].

GTP could also represent GD symptoms [32]. Specifically, (i) silence symptoms, a peripheral criteria, involves preoccupation for gaming when it dominates thoughts, emotions and behaviors [42]. (ii) The Elaborated Intrusion Theory (EIT) posits that cravings manifest as spontaneous intrusions (cognitive or sensory) that can become desires when elaborated [43]. Addictions are often triggered and sustained by cravings. In patients with both GD and GTP, it has been observed that re-experiencing game content or attempting to induce sensory experiences contributes to the maintenance and relapse of GD symptoms [22, 44]. The cycle of dysfunctional gaming has been posited to be worsened by susceptibility to GTP, including euphoric recall of gaming, cravings, and impulses drive to play, despite negative consequences [33]. (iii) Withdrawal symptoms in GD comprise anxiety, moodiness, sadness, and irritability. There is little evidence about the physiological symptoms of GD compared to other addictions [45]. Withdrawal symptoms in addictions may include perceptual disturbances such as hallucinations, with intact reality testing [46]. Dystonias and dyskinesias (e.g., involuntary muscle movements, twisting and repetitive movements) can be withdrawal symptoms caused by dopamine receptor-blocking agents (neuroleptics). Involuntary movements of fingers have been also observed in repetitive tasks such as playing piano [47]. In GTP involuntary and repetitive finger movements have been observed when trying to sleep and are sometimes triggered by an identifiable game cue [12]. Memory and attention changes, irrational thinking and expectancies may be cognitive features of withdrawal [45]. GTP also involves irrational thinking and false expectations associated with game outcomes in the real world [48].

4. Conclusions

This paper presents theories on GTP to prompt new research avenues. GTP encompasses a broad range of phenomena, from sensory perceptions to high cognitive levels. Consequently, the etiology and underlying mechanisms of the various forms may exhibit differences. The interpretation of studies on GTP appear to offer initial support, although further investigation is necessary to validate or refute the proposed theories.

References

- [1] Ortiz de Gortari AB, Diseth Å. Multidimensional assessment of Game Transfer Phenomena: Intrusive cognitions, perceptual distortions, hallucinations and dissociations. *Front Psychol.* 2022;13:896238.
- [2] Ortiz de Gortari AB. Game Transfer Phenomena: Origin, development and contributions to the videogame research field. In: Attrill-Smith A, Fullwood C, Kuss D, Keep M, editors. *Oxford Handbook of Cyberpsychology*. Oxford: Oxford University Press.; 2019. p. 532-56.
- [3] Ortiz de Gortari AB, Pontes HM, Griffiths MD. The Game Transfer Phenomena Scale: An Instrument for Investigating the Nonvolitional Effects of Video Game Playing. *Cyberpsychol Behav Soc Netw.* 2015;10(18):588-94.
- [4] Cudo A, Zabielska-Mendyk E, Ortiz de Gortari AB. Psychometric Assessment and Gender Invariance of the Polish Adaptation of the Game Transfer Phenomena Scale. *Adv Cogn Psychol.* 2022;18(1):48-63.

- [5] Ortiz de Gortari AB, Cudo A. Adaptation of the Game Transfer Phenomena Scale into Spanish: Sensory-Perceptual Changes, Intrusive Thoughts, and Involuntary Behaviors in Mexican Video Game Players. *Rev Int Investig Adicciones*. Pre-print. 2023.
- [6] Ortiz de Gortari AB. Sensory and cognitive intrusions with and without media content during the COVID-19 pandemic: Isolation, media use, sleep and stress factors. *Telemat Inf*. 2024;87(C):14.
- [7] Ortiz de Gortari AB, Edwards C, Palmer-Cooper E. Sensory experiences during and after playing video games and watching videos with ASMR-elicting properties. ICBA 2024; Gibraltar, UK2024.
- [8] Ortiz de Gortari AB, Panagiotidi M. The Role of ADHD and Dissociation Symptoms in Gaming Disorder and Game Transfer Phenomena. *Aust N Z J Psychiatry*. 2023;57(S1):3-198.
- [9] Kihlstrom JF, Glisky ML, Angiulo MJ. Dissociative tendencies and dissociative disorders. *J Abnorm Psychol*. 1994;103(1):117.
- [10] American Psychiatric Association. Diagnostic and statistical manual of mental disorders, fifth edition, text revision. Washington, DC: American Psychiatric Association; 2013.
- [11] Guglielmucci F, Monti M, Franzoi IG, Santoro G, Granieri A, Billieux J, et al. Dissociation in problematic gaming: A systematic review. *Current Addiction Reports Curr Addict Rep*. 2019;6(1):1-14. Epub ahead of print.
- [12] Ortiz de Gortari AB, Griffiths MD. Altered visual perception in Game Transfer Phenomena: an empirical self-report study. *Int J Hum-Comput Interact*. 2014;30(2):95-105.
- [13] Ortiz de Gortari AB, Griffiths MD. Letter to the Editor for ‘Current Addiction Reports’—Game Transfer Phenomena and Dissociation: a Reply to Guglielmucci et al. (2019). *Curr Addict Rep*. 2019;6:155-8.
- [14] Llamas-Alonso J, Kvavilashvili L, Georgiou G, Ortiz de Gortari AB. Positive schizotypy and internet gaming addiction as predictors of Game Transfer Phenomena in daily life: A diary study. *Journal of Behavioral Addictions*. 2022;11:232.
- [15] Vis PJ, Goudriaan AE, ter Meulen BC, Blom JD. On Perception and Consciousness in HPPD: A Systematic Review. *Frontiers in Neuroscience*. 2021;15:675768.
- [16] Stickgold R, Malia A, Maguire D, Roddenberry D, O'Connor M. Replaying the Game: Hypnagogic Images in Normals and Amnesics. *Science*. 2000;290(5490):350-3.
- [17] Engle RW, Kane MJ. Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychol Learn Motiv*. 2004;44:145-200.
- [18] Llamas-Alonso J. Mechanisms of the Game Transfer Phenomena: The role of cognitive, emotional, and personality variables: University of Hertfordshire; 2024.
- [19] Croft RJ, Lee A, Bertolot J, Gruzelier JH. Associations of P50 suppression and desensitization with perceptual and cognitive features of “unreality” in schizotypy. *Biol Psychiatry*. 2001;50(6):441-6.
- [20] Llamas-Alonso J, Kvavilashvilia L, Gergiou G, Ortiz de Gortari AB. Contextualising Game Transfer Phenomena into Involuntary Cognitions: The effects of Gaming Disorder, Cognitive and Psychopathological Predictors (under review).
- [21] Whitfield-Gabrieli S, Ford JM. Default mode network activity and connectivity in psychopathology. *Annual review of clinical psychology*. 2012;8:49-76.
- [22] Le HK, Ortiz de Gortari AB, Callan A, Poynton D, Vecchio D, Chen W. Game transfer phenomena in a clinical case with psychosis and gaming disorder. *Psychiatry Int*. 2023;4(3):286-96.
- [23] Kvavilashvili L, Mandler G. Out of one's mind: A study of involuntary semantic memories. *Cognitive Psychology*. 2004;48(1):47-94.
- [24] Powers A, Mathys C, Corlett P. Pavlovian conditioning-induced hallucinations result from over weighting of perceptual priors. *Science*. 2017;357(6351):596-600.
- [25] Han DH, Lyoo IK, Renshaw PF. Differential regional gray matter volumes in patients with on-line game addiction and professional gamers. *Journal of Psychiatric Research*. 2012;46(4):507-15.
- [26] Ortiz de Gortari AB, Griffiths MD. Prevalence and characteristics of Game Transfer Phenomena: A descriptive survey study. *Int J Hum-Comput Interact*. 2016;32(6):470-80.
- [27] Cudo A, Zabielska-Mendyk E. Relationship between self-control dimensions, emotional regulation and problematic gaming: The mediating role of the Game Transfer Phenomena. *Journal of Behavioral Addictions*. 2022;11:233-34.
- [28] Ortiz de Gortari AB, Panagiotidi M. The interplay between executive function deficits, psychopathological traits and dysfunctional gaming habits in the context of Game Transfer Phenomena. *Comput Hum Behav*. 2023;138:107469.
- [29] Alderson-Day B, Moffatt J, Lima CF, Krishnan S, Fernyhough C, Scott SK, et al. Susceptibility to auditory hallucinations is associated with spontaneous but not directed modulation of top-down expectations for speech. *Neuroscience of consciousness*. 2022;2022(1):niac002.
- [30] Chuang Y-C. Massively Multiplayer Online role playing game induced seizures: A neglected health problem in internet addiction. *Cyberpsychology & Behavior*. 2006;9(4):451-6.
- [31] McConnell A, He W, McConnell H, Sowman PF. Associations between Hallucinogen Persisting Perception Disorder (HPPD) and non-visual perceptual disturbances. 2023.
- [32] Ortiz de Gortari AB, Laroi F, Lerner A. Can Game Transfer Phenomena contribute to the understanding of the psychophysiological mechanisms of gaming disorder? A comparison of visual intrusions induced by videogames and mental disorders. *J Behav Addict*. 2017;6(S1):9-11.
- [33] Basche A, Ortiz de Gortari AB. The Impact of Game Transfer Phenomena on Cravings in IGD-ADHD Adolescents. *Journal of Behavioral Addiction*. 2023;11(1):24.
- [34] Jardri R, Hugdahl K, Hughes M, Brunelin J, Waters F, Alderson-Day B, et al. Are Hallucinations Due to an Imbalance Between Excitatory and Inhibitory Influences on the Brain? *Schizophr Bull*. 2016.
- [35] Martinotti G, Santacroce R, Pettoruso M, Montemiro C, Spano MC, Lorusso M, et al. Hallucinogen Persisting Perception Disorder: Etiology, Clinical Features, and Therapeutic Perspectives. *Brain Sciences*. 2018;8(3):47.
- [36] Ortiz de Gortari AB, Gackenbach J. Game Transfer Phenomena and Problematic Interactive Media Use: Dispositional and Media Habit Factors. *Frontiers in Psychology*. 2021;12(1144):585547.
- [37] Ortiz de Gortari AB, Oldfield B, Griffiths MD. An empirical examination of factors associated with Game Transfer Phenomena severity. *Computers in Human Behavior*. 2016;64:274-84.

- [38] Hou H, Jia S, Hu S, Fan R, Sun W, Sun T, et al. Reduced striatal dopamine transporters in people with internet addiction disorder. *BioMed Research International*. 2012;2012.
- [39] Llamas-Alonso J, Georgiou G, Kvavilashvili L, Ortiz de Gortari AB. The role of spontaneous cognition, positive schizotypal traits and internet gaming addiction in the frequency of Game Transfer Phenomena. In: *BPS Cyberpsychology Virtual Conference*; 2021 Jul 6-7.
- [40] Ortiz de Gortari AB, Griffiths MD. Severity levels of Game Transfer Phenomena: Associated factors and impact. *Journal of Behavioral Addictions*. 2016;5(S1):16-7.
- [41] Ortiz de Gortari AB, Pontes H, Griffiths MD. The Game Transfer Phenomena Scale: An instrument for investigating the non-volitional effects of video game playing. *Cyberpsychology, Behavior, and Social Networking*. 2015;18(10):588-94.
- [42] Kuss DJ, Griffiths MD. Internet gaming addiction: A systematic review of empirical research. *Int J Ment Health Addiction*. 2012;10(2):278-96.
- [43] Andrade J, May J, Kavanagh D. Sensory imagery in craving: From cognitive psychology to new treatments for addiction. *Journal of Experimental Psychopathology*. 2012;3(2):127 - 45.
- [44] Ortiz de Gortari AB, Basche A. Pain and gain of auditory intrusions with video game content: Game transfer phenomena in clinical cases. *European Psychiatry*. 2021;64(S1):S642-S642.
- [45] Kaptosis D, King DL, Delfabbro PH, Gradišar M. Withdrawal symptoms in internet gaming disorder: A systematic review. *Clinical Psychology Review*. 2016;43:58-66.
- [46] Saitz R. Introduction to alcohol withdrawal. *Alcohol Research and Health*. 1998;22(1):5.
- [47] Haueisen J, Knösche TR. Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience*. 2001;13(6):786-92.
- [48] Ortiz de Gortari AB, Griffiths MD. Automatic Mental Processes, Automatic Actions and Behaviours in Game Transfer Phenomena: An Empirical Self-Report Study Using Online Forum Data. *Int J Ment Health Addiction*. 2014;12(4):432-52.

Visceral Extension and the Beast-Machine: Toward a Neuropsychological Criterion for the Extended Self

Joseph NEISSE^{a,1}

^a*Department of Philosophy, Grinnell College, Grinnell, IA, USA*

ORCID ID: Joseph Neisser <https://orcid.org/0000-0002-2909-5029>

Abstract. When does a cognitive system distributed across brain, body, and world constitute an extended self or person? If cognition extends “beyond skin and skull,” then does the self also extend? Or is cognitive labor simply outsourced to an external device that is not part of the self? More generally, what does reflection on technologies of cognitive enhancement suggest about the nature of the self and its relation to the brain and body? Drawing on Anil Seth’s concept of the self as a “beast-machine,” I argue that although (some form of) the Parity Principle may plausibly provide a criterion for cognitive extension, this is not sufficient to establish an extended self, and that the self is not co-extensive with cognition.

Keywords. Extended cognition, embodiment, self, predictive processing, consciousness

1. Introduction

When does a cognitive system distributed across brain, body, and world constitute an extended self or person? If cognition extends “beyond skin and skull,”² then does the self also extend? Or is cognitive labor simply outsourced to an external device that is not part of the self? More generally, what does reflection on technologies of cognitive enhancement suggest about the nature of the self and its relation to the brain and body? Drawing on Anil Seth’s [1] concept of the self as a “beast-machine,” I argue that while the received form of the Parity Principle plausibly provides a criterion for cognitive extension, it is not sufficient to establish an extended self. That is, the self is not co-extensive with cognition. Nor is it co-extensive with consciousness. The embodied self runs deeper than consciousness; the lived body is not simply the body as it appears in consciousness. I suggest a conception of extended selves rooted in neural mechanisms for proprioception and interoception.

5. Theoretical Framework

First, I’ll quickly rehearse the background for extended cognition which, doubtless, is familiar to many readers. Begin with the *Parity Principle*:

If, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process. ([2], p.8)

Next, add the so-called *Glue & Trust Criteria*:³

1. That the external resource be reliably available and typically invoked.
2. That any information thus retrieved be more-or-less automatically endorsed. It should not usually be subject to critical scrutiny (unlike the opinions of other people, for example). It should be deemed about as trustworthy as something retrieved clearly from biological memory.
3. That information contained in the resource should be easily accessible as and

¹ Corresponding Author: Joseph Neisser, neisserj@grinnell.edu

² The phrase “beyond skin and skull” is due to [2].

³ The phrase “glue & trust” is from [3]. See also [4].

when required. ([5], pp.79).

By appeal to the Parity Principle, duly supplemented by the glue & trust criteria, it is possible to construct arguments showing that notebooks, video game controllers, smart phones, and gps devices can become parts of an integrated cognitive system that extends beyond skin and skull. These extended systems perform labor for which we would give cognitive credit, had a person done it in their heads. In that sense, thought itself is extended and enhanced.

In a later discussion of the metaphysical implications of the Parity Principle, David Chalmers [6] distinguishes between two kinds of extended cognition, circuit extension and sensorimotor extension (see also [7]). I next introduce these and then suggest a third possibility which crosscuts Chalmers' distinction: *visceral extension*.

Circuit extension is exemplified by the case of Diva, a human who suffers minor brain damage and loses some specific cognitive ability, say, arithmetic [8]. Diva has an external silicon circuit connected wirelessly to her brain, restoring the original functions without loss. By parity reasoning, then, Diva's cognition is now extended in a new way, supervening on the chip-brain circuitry rather than just the brain. Even the critics of extended cognition find this conclusion undeniable (e.g., [9]). But *sensorimotor extension* is more interesting and more difficult to accept. In sensorimotor extension, the glue & trust criteria are met via person-level perception and action rather than by subpersonal circuitry. Amnestic Otto accesses information in his trusty notebook by *reading* it, and the savvy Tetris player finds the correct orientation of an image by *turning* it on the screen with a dependable clicker [2]. Parity reasoning (again supplemented by Glue & Trust) shows that reading the notebook is a case of remembering, and turning the images is a case of mental rotation.

Thus, sensorimotor extension demonstrates that one's cognitive processes can be outside of *oneself* in surprising ways. Chalmers eventually settles on a revision of the extended mind thesis: A subject's cognitive processes and mental states can be partly constituted by entities that are external to the subject, in virtue of the subject's sensorimotor interaction with these entities ([6], p.7). The upshot is that cognitive processes can be external, not just to the skin and skull, but to *the subject whose cognitions they are*. Chalmers explains that the reason that sensorimotor extension is controversial, while circuit extension is not, is that it violates a fundamental assumption of cognitive psychology, namely, that cognition takes place *between* perception and action. If the thesis of sensorimotor extension is correct, then a core idea of cognitive psychology may turn out to be false. In contrast, the extended circuit thesis is just another piece of banal functionalism. As Chalmers puts it, the thesis of sensorimotor extension "...does not just overthrow the hegemony of skin and skull as boundaries for cognition... It also overthrows the hegemony of perception and action." ([6], p. 9)

As it stands, Chalmers' distinction between circuit extension and sensorimotor extension suggests a simple application to the question about extended selves. Circuit extensions might be self-extensions because they are functionally "between" the inputs and outputs –perception and action – that constitute the bounds of an experiencing subject. Sensorimotor extensions, meanwhile, would not be part of the extended self since they would not be between perception and action. Because one must interact with them on a personal level, they are not part of one's person. But this is too quick. Chalmers speaks of "the subject," not the self or person. Often, writers do speak of "the subject" in a way which suggests this equivalence with the self [10]. But subjectivity corresponds to the first-person *perspective*, which can be enacted in conscious creatures that lack a self [11]. Subjectivity in this sense is partly constitutive of consciousness, and it is to the possibility of *extended consciousness* that Chalmers next turns his attention, arguing that the parity cases are not cases of extended consciousness. Roughly, he holds that perception and action still constitute the bounds of consciousness, even when cognition transcends these bounds via sensorimotor extension. In this way, Chalmers' distinction between sensorimotor extension and circuit extension recapitulates the distinction between consciousness and cognition.

But the self crosscuts the distinction between consciousness and cognition. Pace any doctrinaire Husserlian or Sartrean phenomenology, the self cannot be straightforwardly identified with conscious subjectivity. If there are non-conscious embodied or neural dimensions of the self, then the bounds that constrain consciousness may not constrain the self. Like cognition, the self may extend beyond and beneath consciousness. But

unlike cognition, the self is no mere matter of representation or information processing, because information processing by itself is also insufficient to constitute a self. Whatever an extended self is, it is not just an extension of one's thought but of one's *embodiment*. A self is not just a "thinking thing" (though it is that). It is also a particular kind of being in the world, a way of being that distinguishes itself from its world and carves itself out from the world. So, an extended self is an extended, thinking body that is somehow *lived*. Just as the self is not merely identical with whatever one thinks about or values, the self is also not just whatever physical substrate realizes one's cognition. To extend the self is to become entwined or entangled with an assemblage in a distinctive way.

If the self is embodied and lived in this way, then the Parity Principle cannot be sufficient, even when supplemented by Glue & Trust, to ground the attribution of an extended self (though it may be necessary). A new category is required, which I call *visceral extension*. To evoke this idea, I turn to Anil Seth's image of the self as a *beast-machine* [1]¹. The phrase "beast-machine" is meant, in part, to evoke our basic sense of *being an embodied organism* ([1], pp.250). Roughly, Seth's thought is that embodied selfhood may be grounded in a particular kind of predictive processing, distinct from that which underlies perception of the external world.

Proponents of predictive processing hold that the brain is fundamentally a Bayesian machine. Paradigmatically applied to explain perception, the theory holds that perception consists in the brain's expectations about the next sensory inputs - an expectation which functions as an inference about the external causes of stimuli. Seth [1] extends the framework to speculate that feelings of selfhood are the result of a subtly different, non-perceptual kind of predictive processing, which he calls *control-oriented active inference*. While perception of the external world involves epistemic inference about the causes of stimulation, feelings of the internal milieu may arise from control-oriented inferences which regulate homeostasis ([1], pp. 242–243).

Seth's approach could be helpful for formulating an appropriately embodied and genuinely neuropsychological criterion for the extended self. Such a criterion would ask whether the extending technology – suitably glued & trusted – is coupled with the neural mechanisms that maintain the embodied self – i.e., the beast-machine. The brain may interface with technology in multiple ways, including either via epistemic or control-oriented mechanisms. In the latter case, it may be plausible to suppose that the extension has become part of the self. In this case, the relation to the device would be properly *interoceptive*. It would then be a further question whether this is the only possible basis for an extended self. If so, the criterion may be too restrictive – the technology might then have to be governed by specific and dedicated neural circuits that are not particularly plastic in their range of functions (e.g., brainstem regulation of heart rate), and this kind of brain-technology interface may be limited to certain very specific medical contexts (compare the case of Diva). But if the control-oriented predictive processing for the extended self can include *proprioceptive* mechanisms – mechanisms that perceive, predict, and control the position of the body – then more technological extensions might genuinely become part of the lived body, part of the beast-machine. In effect, proprioception could be a sub-type of sensorimotor extension in which the embodied self *does* overflow the boundaries perception and action set by Chalmers. This idea is not without problems. Seth's framework is designed around interoception. Proprioception would be excluded on grounds that it is perceptual rather than homeostatic. So, if visceral extension is our sole criterion, then technologically extended selves will be rare indeed. This may in fact be the best way forward – it may be best to deny that popular cases of extended cognition are also cases of the extended self. I am sympathetic to this thought. But it may also be too restrictive. It remains highly intuitive that proprioception is, properly speaking, *self*-perception. If so, and if proprioception is not illusory, then there are grounds for including the neural mechanisms of proprioception in the beast-machinery. Certainly, this does not settle the matter. But it points toward a slightly more permissive version of Seth's framework.

¹ Seth draws inspiration for the beast machine from materialist philosopher La Mettrie [12].

6. Discussion

I conclude that technology becomes a proper part of an embodied self when it is anchored in the control-oriented predictive processing mechanisms in the mid-line neural systems implicated in self-regulation, self-control, and self-perception. Thus, the way to *visceral* extension of the embodied self leads through the default mode network.

If the above is on the right track, there are a range of implications for emergent questions about personal integrity, autonomy, and privacy in the age of enhanced reality. For cases of extended cognition *without* an extended self, it may be that existing policy and case law might be applied quite directly. Extended cognition in the absence of self-extension might be comparable to other kinds of personal data – perhaps at one end of a spectrum of such data. Genuinely extended selves, however, would have a different status, and infringements by others might be best understood as violent assault.

Interesting speculative questions also ensue. The nature and even the existence of the self is not well understood. If we can identify the conditions under which extended selves really do and do not arise, how does this affect the wider debates about the self? How might we better conceptualize a “post-Cartesian” kind of being in the world, and contribute to the widely announced project of rethinking the humanistic tradition? When the prototypical human subject is just one possible formation of the self, the effect is to “decenter the human” and to suggest a continuum between human, animal, and technological forms of cognition ([12], p.182). An account of the extended self can play a role in transforming, while retaining contact with, traditional concepts like agency, responsibility, and integrity. In the world of technological entanglement these categories will not be simply jettisoned, They will be retrofitted.

Notoriously, the late Daniel Dennett asked, “Where am I?” [14]. He devised a series of cases involving a telepresence robot of a kind that was strictly science fiction at the time.¹ His characteristically rhetorical discussion pointed to a deflationary, possibly eliminativist conception in which the self is either identified with the brain or turns out to be a narrative illusion (see also [15]). But perhaps there really are embodied selves that can extend across the brain, animal body, and technology?

References

- [1] Seth AK. Being a beast machine: the origins of selfhood in control-oriented interoceptive inference. In: Columbo M, Irvine E, Stapleton M, editors. *Andy Clark and His Critics*. Oxford: Oxford University Press; 2019. p. 238-53. <https://doi.org/10.1093/oso/9780190662813.001.0001>.
- [2] Clark A, Chalmers D. The extended mind. *Analysis*. 1998;58(1):7-19.
- [3] Clark A. Coupling, constitution, and the cognitive kind: a reply to Adams and Aizawa. In: Menary R, editor. *The Extended Mind*. Cambridge, MA: MIT Press; 2010. p. 81-100.
- [4] Gallagher S. *Embodied and enactive approaches to cognition*. Cambridge: Cambridge University Press; 2023.
- [5] Clark A. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press; 2008.
- [6] Chalmers D. Extended cognition and extended consciousness. In: Columbo M, Irvine E, Stapleton M, editors. *Andy Clark and His Critics*. Oxford: Oxford University Press; 2019. p. 9-20. <https://doi.org/10.1093/oso/9780190662813.001.0001>.
- [7] Parsons T, Neisser J. Intersecting frontiers of extended reality and neuropsychology. In: Kissel A, Ramirez E, editors. *Exploring extended realities: metaphysical, psychological, and ethical challenges*. London: Routledge; 2024.
- [8] Clark A. Letter to the editor. *London Rev Books*. 2009;31(6):March 26. Available from: <https://www.lrb.co.uk/the-paper/v31/n06/letters>.
- [9] Adams F, Aizawa K. The bounds of cognition. *Philos Psychol*. 2001;14(1):43-64.
- [10] Neisser J. Subjectivity and the limits and narrative. *J Conscious Stud*. 2008;15(2):51-66.
- [11] Neisser J. *The science of subjectivity*. New York: Palgrave Macmillan; 2015.
- [12] La Mettrie JO. *L'homme Machine*. Leiden: Luzac; 1748.
- [13] Hayles NK. *Unthought: the power of the cognitive nonconscious*. Chicago: University of Chicago Press; 2017.
- [14] Dennett DC. Where am I? In: *Brainstorms: Philosophical Essays on Mind and Psychology*. Montgomery, VT: Bradford Books; 1978. p. 310-23. Available from: <https://mitpress.mit.edu/9780262540377/brainstorms/>.
- [15] Dennett DC. The self as a center of narrative gravity. In: Kessel F, Cole P, Johnson D, editors. *Self and consciousness: multiple perspectives*. Hillsdale, NJ: Erlbaum; 1992.

¹ See [7] for discussion of the extended self and a contemporary version of the telepresence robot – one that is no longer science fiction.

Onlife Subjectivities Construction in two Radicalized Ecosystems. A Practical Application of a Theoretical Perspective

Ilaria VERGINE^{a,1}, Carlo GALIMBERTI^a, Barbara LUCINI^b
and Marco LOMBARDI^b

^a*PsiCom – Research Center in Communication Psychology, Department of Psychology, Università Cattolica del Sacro Cuore*

^b*ITSTIME Italian Team for Security, Terroristic Issues & Managing Emergencies, Department of Sociology, Università Cattolica del Sacro Cuore di Milano*
ORCID ID: Ilaria Vergine <https://orcid.org/0000-0002-0921>

Abstract. Radicalized ecosystems emerged in the Horizon CounteR project as realized in a wide variety of real and digital environments, and these ecosystems can be analyzed through a theoretical perspective made of four thematic areas: communication, identity, culture, and electronic technologies. The precipitate of these four thematic areas could be the radicalized onlife subjectivity: the radicalized subject's identity that becomes intelligible and observable during a communicative interaction on the border of real and digital environments. This paper compares two radicalized ecosystems where electronic technologies concur to realize radicalized activities: Mohammed Zahir Khan and Maria Giulia Sergio. The data on these ecosystems were collected through a five-stage search: 114 sources of evidence were selected for Mohammed Zahir Khan and 31 were appropriate for Maria Giulia Sergio (a Horizon CounteR project deliverable includes the syntheses derived from retrieved sources of evidence). To build the comparison between the two ecosystems, the researchers analyzed them by using the four thematic areas mentioned above and, consequently, the configuration of their subjectivities. The findings show an example of applying the theoretical perspective developed during the Horizon CounteR project and two ways of producing radicalized subjectivities in the case of using electronic technologies for radicalized activities. The conclusions, stemming from the findings, consist of practical fallouts for intelligence.

Keywords. Onlife, subjectivity, radicalization, radicalized ecosystems

1. Introduction

Radicalization is a complex phenomenon generated by various causes/factors. Mapping factors and understanding radicalization are essential to countering this phenomenon. The threat it poses, particularly for the possibility of subsequent violent extremism, has led to funding projects and the publication of scientific literature on radicalization, as a scoping review conducted on psychological and psychosocial factors of radicalization during the Horizon CounteR project (HCP) revealed [1]. Among the causes mentioned above, as implied by the locution “psychological and psychosocial radicalization factor”, some are social, psychological, or even psychosocial. This is why the psychological and psychosocial study of radicalization phenomena is crucial.

This contribution stems from the results and conclusions a team of researchers, including psychologists, achieved during Work Package 2 of the HCP. According to the adopted psychosocial perspective, radicalization emerged as a phenomenon that could be traced in various real and digital [e.g., 2] environments, and radicalized ecosystems could be analyzed through a theoretical perspective made of four thematic areas: communication, identity, culture, and electronic technologies. The precipitate of these areas could be the onlife subjectivity. “*Onlife*” indicates that it is impossible to establish

¹ Corresponding Author: Ilaria Vergine, ilaria.vergine@unicatt.it.

whether a subject is online or offline because the subject always lives on the border of real and digital environments and is always in contact with digital technologies and interfaces. “*Subjectivity*” is the subject’s identity that becomes intelligible and observable during a communicative interaction. It depends on the conscious or unconscious decisions the subject makes in interacting with the environments, artifacts, and other subjects to show who they are. Consequently, onlife subjectivity is the subjectivity that blooms from being in an onlife situation, and radicalized onlife subjectivity is characterized by the process of radicalization. Mentioning digital environments, electronic technologies, and onlife subjectivity shows the relevance of specific branches of psychology: cyberpsychology [3] and social psychology of cyberplaces [4]. Consequently, the present paper aims to: (a) provide an application of the theoretical perspective made of the four thematic areas and onlife subjectivity to two ecosystems (Mohammed Zahir Khan [MZK] and Maria Giulia Sergio [MGS]) characterized by being Jihadist and related to ISIS propaganda; (b) compare MZK and MGS, and show the variability in producing subjectivity when electronic technologies are used to realize radicalized activities; and (c) suggest practical fallouts for intelligence.

2. Methods

2.1 Ecosystems and data sources

MZK and MGS (see **Table 1**) were included among ten ecosystems selected in collaboration with law enforcement agencies during Task 2.3 of Work Package 2 of the HCP because of their relevance and variability in characteristics (e.g., gender). During Task 2.3, MZK and MGS were synthesized by IV and CG. Each synthesis resulted from performing a systematic search of sources of evidence on five different data sources: (a) Counter Extremist Project Database search, (b) general open-source search via the internet (Google browser), (c) European Media Monitor Tool search, (d) Summon – Unicatt (tool for academic research provided by Università Cattolica del Sacro Cuore di Milano) search, and (e) Yewno – Unicatt (tool for creation of knowledge maps) search.

Table 1. Descriptive information on Mohammed Zahir Khan and Maria Giulia Sergio.

Information	Mohammed Zahir Khan	Maria Giulia Sergio
Country	Great Britain	Italy
Birth	18 July 1977	23 September 1987
Gender	Male	Female
Job	Shopkeeper	Part-time worker at a call Center/university student
Type of radicalization	Religious: Jihadist	Religious: Jihadist
Typology of radicalized subjects	ISIS Propagandist	ISIS Propagandist
Radicalized activities	Collecting, posting, sharing Radicalized material	Posting radicalized material, Attending and delivering lessons on radicalized content, converting family and others, joining ISIS, and attending firearms trainings
Charged of		Joining an international terrorist group/entity
Link to Syria	Encouraging terrorism and stirring up hatred Yes, anger against Assad's regime because he attacked women and children	Yes, she joined Islamic State
Family found involved in radicalization	No	Yes, the family of origin, family with her second husband, and her husband's family

The two ecosystems pertain to violent and Islamic typologies of radicalization; referring to Wiktorowicz's [5] classification, they fall in the realm of Sunni Islam. It denotes individuals who assert the need to revert to the original religious practices established by Mohammed and his initial three generations of followers, known as “salaf” (a term signifying “ancestors”) [6,7]. Within the domain of Sunni Islam, Salafist orientation encompasses those who advocate for a puritanical interpretation of the scriptures to restore the religious practices of Mohammed and his early followers. Specifically, the ecosystems align with the Jihadist perspective, where violence is an acceptable means to

fulfill Salafist objectives [8]. Notably, ISIS is recognized as part of this Jihadist subdivision [6].

2.2 Data collection

A five-stage search allowed the retrieval of the sources of evidence (listed in Deliverable 2.3 of the HCP, mostly newspaper articles and scientific literature mainly related to MGS; see Table 2).

Table 2. Sources of evidence retrieved during the five-stage search for each ecosystem.

Data Source	Mohammed Zahir Khan	Maria Giulia Sergio
Counter Extremist Project	1	0
Database		
Google browser	51	29
European Media Monitor Tool	0	0
Summon – Unicatt	70	3
Yewno – Unicatt	0	0
Duplicates among the five activities	8	
Total Number	114	31
Unretrieved full-texts	1	0

2.3. Analysis

First, the psychologists, IV and CG, independently read the syntheses.

Second, they independently read the definition and description of each thematic area included in the theoretical perspective developed to analyze radicalized ecosystems. They wrote down the characteristics of MZK's ecosystem for each thematic area. They did the same for MGS' ecosystem. Consequently, they were able to outline the subjectivities of each ecosystem.

Third, they compared the characteristics attributed to each ecosystem for each thematic area and the subjectivities. In case of disagreement, IV and CG discussed until they reached a common opinion. To allow the comparison, regarding MZK, the researchers considered data until he was arrested in May 2018 and MGS until she was sentenced in absentia in December 2016.

Fourth, the authors discussed what emerged from these three steps to structure the findings and the conclusions.

3. Findings

The four thematic areas and the subjectivities are compared below.

Regarding communication, MZK is related to Great Britain, and MGS is related to Italy, especially at the beginning. This means that, at a macro level, they lived in different societies and communities. However, they moved from one city/town to another (the first, from Birmingham to Sunderland; the second, from Torre del Greco to Inzago). Consequently, they moved from one community to another. Nonetheless, due to their radicalization related to ISIS, they were both connected to what was happening in the Syrian and ISIS communities. MGS was linked to additional radicalized communities. At the meso level, in both ecosystems, the family was mentioned by sources of evidence. MZK's family was not radicalized and had two members, unlike MGS' family, which was radicalized and comprised many members (her relatives and her husband's). At the micro level, MZK appeared to be isolated from other radicalized persons, while MGS was deeply involved in relationships with radicalized persons, for example, Bushra Haik. Lastly, the more MGS and her family progressed in radicalization, the more they decreased their contact with other people, especially non-Muslim ones. MZK and MGS had psychological mechanisms that characterized their communication: MZK's undoing defense mechanism [9] and MGS's contradiction mechanism. For instance, MZK moved from Birmingham to Sunderland to delete his criminal past, but he then committed radicalized crimes. MGS affirmed that men are superior, but she did not hesitate to order her father to move to Syria. Their rhetoric used words that divided the world into believers (ingroup) and unbelievers (outgroup). Indeed, they used words like "scum"

[10] and “kuffar” [11]. Only MGS’ ecosystem showed traces of radicalization in the real environments related to (e.g., the Islamic center) or not related to religion (e.g., the Pomeriggio 5 TV talk show). MGS used real objects to show her radicalization (e.g., the hijab and niqab). However, selected sources of evidence do not mention anything like this for MZK.

Regarding identity, the two ecosystems appear to be different in produced radicalized subjectivity. MZK produced a propagandist subjectivity, an online one. All his radicalized actions were online. MGS produced a propagandist subjectivity that was an onlife one. She tried to convert people in many environments, even ones on the border of digital and real, like Skype calls, which are real environments mediated by a screen according to Milgram’s reality-virtuality continuum [2]. In addition, according to the sources of evidence, MZK did not show any sign of radicalization in his other subjectivities: the shopkeeper, the son, and the father. His mother’s continued defense of him by declaring that he was not a violent person could further prove this. MGS, instead, showed radicalization in all her subjectivities, such as the daughter, the wife, and the worker. At a certain point, she changed her name, the first element through which the others recognized the existence of one subject [4], into Fatima Az-Zhara.

Regarding culture and more specifically the social context, others discriminated against MZK and MGS. The first also suffered because of racism. In MGS’ ecosystem, the primary reason she was discriminated against in different contexts was the use of the veil. Regarding the historical moment, the common point of the two ecosystems was the conflict that was taking place in Syria. They were different regarding where they lived (the first was settled in Great Britain and the second in Italy). All these aspects concur to shape the culture of the two ecosystems. In addition, nothing is known about MZK’s education. Instead, MGS was well-educated. She attended university, but she did not graduate. It makes clear that having a high level of education does not prevent one from becoming radicalized. Considering what was stated about the two ecosystems’ networks at the macro, meso, and micro levels, it appears clear that MZK and MGS were both involved in deviant networks that could allow them to introject a deviant culture (the first was involved in a criminal culture, and the second was involved in a radicalized kind). Since he was in Sunderland, MZK’s networks appear to be uninvolved in radicalization. MGS was involved in radicalized networks at the meso level and in relationships with radicalized people at the micro level. Consequently, her radicalized culture resulted from what she learned and reinforced during her interactions at the meso and macro levels, impacting her identity.

Regarding electronic technology, MZK and MGS used it in their radicalized activities. They used social media (a digital environment) and had Facebook. However, MGS allowed radicalization to become more pervasive in her life by consistently using Skype to acquire and spread radicalized knowledge. Regarding the use of digital and radicalized objects, they possessed radicalized videos. In addition, they used radicalized symbols in their Facebook activities, such as the ISIS flag.

4. Conclusion

The following conclusions can be built from the previous sections.

First, **Table 1** and the findings show that females could play a role in radicalization processes and networks, which is in disagreement with the literature on radicalization.

Second, intelligence must be updated and trained to use electronic technologies. Indeed, these technologies were relevant for the analyzed ecosystems. Consequently, the social psychology of cyberplaces can give valuable insights into this field.

Third, this paper shows a practical application, albeit brief and incomplete, of the theoretical framework developed to study radicalized ecosystems during the HCP. MGS provided an example of radicalized onlife subjectivities. For instance, being a radicalized daughter was the result of inhabiting real and digital environments (see the use of Skype), and the same was about being a propagandist on Skype, in TV shows, and in Syria. MZK showed a single radicalized subjectivity, the one of being a propagandist, and it was online. Mapping radicalized subjectivities and whether they are online or onlife allows us to elaborate on two final remarks for intelligence. First, they should work by pursuing what the concept of digital humint [12] suggests: tracing radicalized activities in all types

of environments, a priori focus on real or on digital environment (without considering the specificity of the ecosystems) gives back a reductionist analysis. In this sense, a lot of work should be done to improve terrorist risk assessment instruments that just partially address radicalization phenomenology mediated by electronic technologies. Second, according to Work Package 2 of the HCP and what emerged from sources of evidence on MZK, analyzing subjectivities allows us to understand how much radicalization is pervasive in the subject's identity. The more subjectivities a subject has in their life that are radicalized, the more difficult it is to find a non-radicalized identity from which it is possible to start a deradicalization process. According to that, MZK could have a starting point for deradicalization (being a son and a father did not seem to have traces of radicalization). It agrees with what "An Imam from Pentonville prison ... sent a letter to the judge who suggested the shopkeeper's extreme views may be changing in a more positive direction" [7]. MGS could not have a starting point: every single subjectivity was radicalized.

This paper is limited to analyzing data from sources of evidence rather than from direct observation of radicalized ecosystems. Future research should provide further applications of the theoretical perspective.

References

- [1] CounteR Project. CounteR Project: countering radicalisation for a safer world [Internet]. 2024 [cited 2024 Jun 13]. Available from: <https://counter-project.eu/>
- [2] Milgram P, Kishino F. A taxonomy of mixed reality visual displays. IEICE Transactions on Information and Systems. 1994;E77-D(12):1321–1329.
- [3] Barak A, Suler J. Reflections on the psychology and social science of cyberspace. In: Barak A, editor. Psychological aspects of cyberspace: theory, research, applications. Cambridge, EN: Cambridge University Press; 2008. p. 1–12.
- [4] Galimberti C. Segui il coniglio bianco: la costruzione della soggettività nelle interazioni medicate [Follow the white rabbit: the construction of subjectivity in mediated interactions]. In: Regalia C, Marta E, editors. Identità in relazione: le sfide odiere dell'essere adulto [Identity in relationship: today's challenges of being an adult]. Milano, IT: McGraw-Hill; 2011. p. 72–127.
- [5] Wiktorowicz Q. Anatomy of the Salafi Movement. Studies in Conflict & Terrorism. 2006 May;29(3):207–239.
- [6] Guhl J. A theoretical introduction to contemporary Salafism. London, EN: ISD: Powering solutions to extremism, hate and disinformation; n.d. 11 p.
- [7] The Newsroom. Sunderland shopkeeper jailed for terror posts. Sunderland Echo. 2018 May 2 [cited 2024 Jun 13]. Available from: <https://www.sunderlandecho.com/news/sunderland-shopkeeper-jailed-terror-posts-350919>
- [8] Kepel G. Jihad: the trail of political Islam. Harvard, MA: Harvard University Press; 2002.
- [9] Laplanche J, Pontalis JB. Encyclopédia della psicoanalisi [Psychoanalysis encyclopedia]. Laterza; 2008.
- [10] Davis M. Convicted terrorists thought to be due for release before end of March. Press Association. 2020. Available from: <https://go.exlibris.link/twCKwjnw>
- [11] Giacalone G. Islamic extremism from the Balkans emerges in Italy. Sicurezza, Terrorismo e Società/Security, Terrorism, Society. 2015;2:87–92.
- [12] Lombardi M. Le nuove prospettive della digital humint [The new perspectives on digital humint]. RA Rivista Aeronautica, Difesa Servizi. 2023;3:16–21.

The Structural Considerations of an Online Games Community

Eric J. CHOU^{a,1}

^aHonor Games, University of California Irvine Alumni

Abstract. This 32-month qualitative quarterly longitudinal study examines the structures of an online gaming clan community of the popular free to play massively multiplayer online game (MMOG) World of Tanks developed by Wargaming. During the participant observational period, from formation through several structural shifts, this study, which originally started under a Qualitative Field research course, has become a comprehensive online ethnography, documenting roles, responsibilities, organizational citizenship behaviors, as well as features of remote teams. 53 interviews were conducted with key personnel in the organization, durations varying from 15 minutes to an hour. Four distinct themes arose from the research; freedom, friendship, responsibility, and balance. This study grants insight into gaming organizations, which will better help understand the gaming culture from a foundational constructivist perspective. A framework for understanding and managing organizations was also discovered, regardless of their size, Form, function or local or global presence.

Keywords. Clans, videogames, global teams, virtual teams, organizational citizenship, online ethnography.

1. Introduction

As our societies evolve, there is no question the tech industries are right there with us, improving, simplifying and complicating our lives. We compensate these facilitating organizations for their service and expect them to continue to connect with diverse people under their brand. Similarly, we gather at bars, churches, schools and the like, for a sense of kinship, community, education, support, service and purpose. All of these traditionally serve as a means of communication, and few would question the important role they play in our lives. So, why is it that as soon as you mention "video games", as a valid form of meeting these human needs and societal functions, you tend to get drastic polarized perspectives? Even within academia, researchers are usually split between alarmists who claim videogames are dangerous violent /addictive pastimes, humanists who explore human expression/ social dynamics within games, and technologists who focus on the mechanical possibilities of such applications. Video games are stereotyped as unstructured play", and players and designers get inconsistent reactions from those in other industries and fields. The misattribution inevitably stems from not knowing or understanding the base structure of the activities nor the industry that has spawned them.

The mission of this research was to partially de-stigmatize gaming clans, illustrating that these vibrant communities can have defined organizational roles that mirror business and military structures. Furthermore, the life of a clan can grant insights into issues of governance, civics, organizational citizenship, personality, management, leadership, training, engagement, organizational development, and responsibility. These themes go beyond the social psychological paradigm traditionally pigeonholing these avenues of interest and extend into motivation and potentially work ethic. Despite the plethora of research fields mentioned, the aim of this study was to examine the organizational roles and structures that allow such rich online ecosystems to exist.

For my research I have chosen to qualitatively examine the 102nd multi gaming community, based around the popular free to play massively multiplayer online game (MMOG) World of Tanks, developed by Wargaming. The WWII shooter game involves 15 on 15 battle rounds, which players drive semi historically accurate light/ medium/ heavy tanks and artillery from that era. As you play, you obtain experience and in game

¹ Corresponding Author: Eric Chou, eric.chou@honorgames.co

currencies which are used to purchase, modify, and enhance higher tier vehicles. There are two primary victory conditions, in which your team can either eliminate all opposing forces, or capture their spawn point, which is marked by a flag. Clans do play an important aspect in the game, by organizing and facilitating teamwork to achieve these conditions. Furthermore, clan vs. clan competitions take place regularly, for the reward of gaining in game currency, which can be used to purchase and speed up game progression.

2. Theoretical Framework

There is a growing body of research that explores the duality of social gaming, prime among them the benefits of online communities. The spectrum of these stretch from online social networking tools such as facebook and forums to skype and gaming organizations, each with their own levels of social engagement and social distance. For the sake of streamlining my research, I will be using the term clan and defining it as a group of gamers who have created a social framework to assist each other to achieve game goals. The topic I've chosen to explore is the commonly overlooked structural foundations of these organizations, and the roles that must be filled to facilitate operation. The research in this particular aspect of gaming and online communities was lacking, back in summer of 2012, when I had started the study. This was likely due to the plethora of issues that can be observed and studied as a result of the structure, however, as we unveil the social aspects of "the clan" we are able to examine the core organizational structures more directly. My guiding research question asks what specific foundational roles and factors are necessary for a clan to survive and thrive?

To help counterbalance the dearth of research into clan structure, I have broadened my conceptual framework to include online communities, yet I will still aim to focus on structural factors (The following is my review of the literature during the summer field research course).

The 1st key concept is that of Organizational citizenship, "originally defined citizenship behavior as "individual behavior that is discretionary, not directly or explicitly recognized by the formal reward system, and that in the aggregate promotes the effective functioning of the organization." Organizational citizenship behaviors are summarized into seven types. These types are helping behavior, sportsmanship, organizational loyalty, compliance, individual initiative, civic virtue, and self-development (Pearce & Herbik, 2004)...community is loosely structured and self-organized without a formally prescribed structure and authority. Therefore, the membership is voluntary. However, research finds that some members in virtual communities show pro-social behaviors, as in physical organizations or groups (Wasko & Faraj, 2000)" [1]. The authors used survey methods to measure the following concepts: Social Support Orientation, Disposition to Trust, Cognitive Trust, Affective Trust, Citizenship, and Knowledge-Sharing Behavior for online communities.

Another key concept of interest is team cooperation, this was studied using a shooter game, Counterstrike, in which two opposing teams must fulfill mutually exclusive objectives to win. This study examined team cooperation and found a 75% cooperation rate [2].

The potential of clans was briefly mentioned, but was not significantly discussed, rather, the researchers focused more on examining cooperation and reactions to deviant behaviors. Although this is indeed important, the next step would be to analyze group dynamics.

Now that we have a basic framework, the following study of Ultima Online expands the scope of social interaction, being one of the 1st massive Multiplayer Online Role-Playing Games, which came out in 1997. This study granted further insight into the social rationale for gaming. "In order to stay in the game only a minimum level of social interaction is necessary, yet most players seek more. There are not merely strategic considerations behind the encounters of characters in Ultima Online but rather the search for communication and persistent social relations. Most players (88%) not only connect to Ultima Online in order to play but also to stay simultaneously in contact with the fellow players by a messaging system. Additionally, the large majority of players (84%) are members of several social units" The Researchers determined that the tenuous players

were involved in clans at an 80% rate, the Regulars at an 85% rate, and the Heavy players were involved at an 89% rate [3]. The other given 84% includes non-clan social units.

Despite the richness of all that can be learned from these clans, which are usually called guilds in these genres, the sheer scale of these guilds, are quite large, and thus increasingly complex, these usually range from anywhere from 30 people to 300, thus analyzing them may be rather a large step, especially since they involve a virtual world.

"Based on (Herrmann et al. 2004), a role has the following four characteristics: position: the social strata in a society or the hierarchy in an organization; function/tasks: roles imply special functions and tasks in the form of explicit and documented expectations and responsibility; behavior expectations: implicit and informal agreements of individual behaviors, which are expected of the roles; social interaction: roles are the result of a social interaction and negotiation between the actor and those with whom he/she interacts.... Most MMORPGs feature a mode of play known as the guild. A guild is a formally organized system, in which tasks and formal roles are normally assigned explicitly to the players. During the social interaction, however, informal roles are developed, assigned or taken over. Although communities usually have only few formal roles but a lot of dynamically developing informal roles, the stability of a guild community is based on explicit formal roles and less on informal, emerging roles. In this kind of setting, informal roles are pretty much influenced by knowledge and expectations the players already have about each other through existing formal roles. As Smith (1966) noted in observations of an Antarctic work group, informal or emergent roles will often replace or supplement more formal roles" [4]. Although the limitation of this research is that it is too general, identifying the specific formal and informal roles in a clan would be very insightful, and given the scope, cross referencing would add significantly to the research.

The final study I reviewed during the course period was a quantitative study of World of Warcraft. This study examined the rates of turn over between players in clans, as well as suggesting optimum clan size. One of the issues the researchers highlighted was that the size usually hovers around the range of 35-50, and any larger requires significantly more organization. The few qualitative conclusions drawn were regarding how important both cohort progression and how a diversity of player skill level would extend the life of a clan [4]. This study was chosen despite its quantitative nature because this was the only study of its kind, it aimed to examine clan structure and did so through statistical models. However, this is still just scratching the surface.

As the years went by, more and more research on clans has come to light, leading to a more comprehensive framework. In Poor and Skoric's study [5], they examined the potential reasons for the death of a six-year EverQuest II clan. In this study, the researchers had joined a prominent clan, seeking to understand reasons for collapse. They identified multiple factors that led to clan inactivity, prime among them as individual member conflicts, competitor games being released, inadequate leadership adaptability, and members becoming inactive due to off-line engagements. The clan was split into two competing in game identities, based on player burnout, reaching the maximum level, as well as competitive factors. Originally, they were both autonomous, and interacted on friendly terms, but as time went on, they began to siphon members from each other, causing social unrest. The researchers tracked membership through the in-game rosters, which neared 1000+ characters combined. They noted that the numbers specified both active and long time in active characters averaged six characters per member. They concluded the study with a questionnaire which they sent out after the clan had dispersed, attempting to track the strength of player friendships. Unfortunately, they only obtained 15 complete responses. Despite my concerns with their questionnaire results, the researchers do a proper job of identifying the differences between technologically mediated communities and ritual mediated communities.

They further correctly identified the need for understanding guild structure and behaviors "Although it is generally understood how guilds work—there are hundreds of successful guilds across many games—and although humans are instinctually predisposed to take part in and support communities (Dunbar, 2005), the specific issues involved with how guild members cohere to sustain a guild over time and the challenges they face are not well understood. If we can help guilds survive over time, certainly game companies will be pleased, but the lessons should apply more generally to mediated

communities and further our understanding of the general online society" [5]. Beyond the potential application to remote work teams and companies.

And yet, the research of clans is not all so straight forward. An online survey was conducted to determine the demographics and player attitudes of first-person shooter games (FPS), and they found that 80% of their 751 respondents were in clans. They further found that "Members of an amateur or (semi)professional clan scored significantly higher on social interaction than non-clan members. The amateurs expressed a stronger motivation for social interaction than the (semi)professionals. A (semi) professional clan provides its members with a functional social network for gaming. The special position of amateur clans in comparison with (semi) professionals is underlined also by the differences in scores on the excitement motive.

A key article I found takes a very intriguing perspective on the potential of clans. In interviewing and exploring China's gaming industry, they've uncovered the duality and paradox the guild can represent. "It is not uncommon for avid gamers in the West to turn a casual leisure into a serious hobby; as Kücklich (2005) has remarked, "the relationship between work and play is changing, leading, as it were, to a hybrid form of 'playbour'" [6], "In China, the spectrum of businesses—and the range of new professions—that have resulted from community-based play is simply staggering. While Western cultural reformers wage war on games for violent content and graphic representations, these concerns believe the often-egregious circumstances of game production itself: the unpaid work of guild laborers who drive innovation and absorb commercial risks to the enormous benefit of the games industry. By appropriating gamer-created content, the industry effectively "close[s] the loop' between corporation and customer [by] reinscribing the consumer into the production process (Kline et al., 2003)" [6].

Guilds have undergone rapid commercialization, becoming mediators in the relationship between the industry and gamer communities. The guild's key role in the secondary industry of online gaming—coupled with the sheer number of guilds and guild members in China, makes it an ideal site in which to explore complex negotiations between commerce and community, and capital and labor, on the micro level, and the shifting dynamics and impetus behind the rise of the secondary industry in China on the macro level" [6]. "By 2012, Chinese guilds developed distinctive characteristics in response to increasing commercialization of gamer involvement. For instance, major gaming guilds in China are often comprised of thousands—or even tens of thousands—of players, far outnumbering their counterparts in the West [7].

Clans are ripe for learning, leadership and training. "From a theoretical standpoint there is enormous potential for examining leadership practices in game worlds due to the fact that games provide simplified decision spaces in which social groups can engage with structured and unstructured tasks and problems. As a result, the real leaders in these virtual spaces are engaging in some of the same basic practices as their counterparts in analog domains, but in a context that is far more accessible for research than most other settings 'in the wild'...On the other hand, the game's designers and developers at the game company are continuously modifying the rules that govern how the game is played in an effort to maintain player interest and keep the game fun. As a result, players also need to periodically revise or relearn their roles so as to maintain a high level of performance. In effect, this forces players to engage in a process not unlike double loop learning where previously mastered skills or tasks are revisited and refined rather than simply repeated according to formula (Argyris, 1991). While teaching expertise requires engaging in this pattern of development, gamers in WoW have to engage in a similar pattern just to stay afloat with changes to the game's content. In other words, end game players in WoW are in some sense engaged in an active simulation of double loop learning" [8]. These lessons and commonalities may extend out of the virtual world in macro educational contexts.

Leadership in both schools and guilds involves fundamental practices that are required by leaders in virtually any kind of organization. Both types of leaders have to do short- and long-term planning. They are periodically faced with the integration of new technologies. As with any work with people, leadership practices in school and guilds entail ongoing conflict management for persistent or intractable issues, and intervention in escalated conflicts for issues that have boiled over. Resource allocation, visioning, maintaining morale, and work within and through organizational culture are all tasks that are also present in guilds, just as they are in schools" [8].

After all, "it is in large part this emergent quality that has led corporate researchers to look seriously at leaders in these virtual worlds (IBM, 2006; Reeves & Malone, 2007). In general, with the world of business looking at the potential of video games as tools for leadership and a generation of students who are deeply involved with games (Gee, 2003; Lenhart et al, 2008), educational researchers and practitioners can ill afford to ignore the implications of gaming technology for both learning and leading more generally (Halverson, 2005). However, even more important than all of these factors are the simple fact that end-game raiding in WoW is a learning activity. While the content players learn may not be particularly relevant for questions of instructional leadership, the organizational features that support this learning are" [8].

3. Methods

I played the role of participant observer in the organization during my 32 months, which originally started as 3 weeks. Construction and maintenance were my prime interests, and I only directly acted as "researcher", during the construction period, August 17-September 12, 2012, and each follow up. Furthermore, I decided to continue to interview on a quarterly longitudinal basis, which the dates of focus were: 12/17/2012, 3/29/2013, 11/9/2013, 3/3/2014, and 7/25/2014. Limited participant observations were done from the last aforementioned date and beginning of April 2015, which the existing structure and community were finally eroded and largely dispersed.

3.1. Setting and Study Participants

The participants were primarily North American World of Tanks players, which are the largest subset in the overall multi-gaming organization, which was at one point over 200 members strong. The demographics vary, with age, being 18-56 years, and the average being 36. As for educational background, it ranged from those who had graduated high school, those who had a GED, a majority with Associates or bachelors, as well as current students. As for occupations, there were auto mechanics, farmers, truck drivers, web designers, small business owners, cooks, warehouse supervisors, former/ current/ active military personnel, and those in several US. departments. Although the aforementioned demographics are not completely exhaustive, it paints an incredibly diverse picture of these individuals. It's important to note however, that the participants were disproportionately male, and since this was a World War II shooter game, it's no surprise to find militarily minded individuals in this community.

As for the sampling strategy, I used a combination of snowball sampling and key informants; however, more of the latter were employed, given my participant observer status, and my existing status in the organization, which was co-founder, lead recruiter, and staff officer.

Clan Structural overview: When the research started, the clan was reforming from a recent shift. The structure was formed as a democracy, which elects a commander every 4 months. During the initial research period, the following structure was voted on and ratified by the attending membership, which were 5-10 individuals. As time went on, the structure evolved to include more formal roles to keep up with expansion.

The commander appoints an executive officer (responsible for running the day-to-day operations), a training officer (responsible for scheduling and conducting training), and a staff officer responsible for all recruiting, diplomatic and documentation duties.

Each platoon Sgt. was responsible for 15 general members. The diplomat roles were dual positions, serving as liaisons to other clans. As the position evolved, it took on support for coordinating joint clan events. Field commanders led in game clan skirmishes.

The structure remained largely consistent for a year until August to December 2013, in which a new armored battalion structure was developed, and later transitioned into. This splits up the staff officer's position into 6 (S- departments) and consolidates the training roles.

S1: Personnel: Activity & documentation tracking, S2: Intelligence: Diplomacy & strategic development, S3: Training: In game training management, S4: Operations

Event scheduling, S5: Recruiting: Full cycle Recruiting, S6: Communications: Team speak & website administration.

The above structure held from its inception to the clan's decline from January to March of 2015.

In addition to the research recordings, Formation meetings and staff meetings were also kept while I was staff officer, which I held from 4/22/2012 to 4/21/13. A clan history log was also created, documenting command staff officer movements from regime to regime.

3.2. Apparatus

During non-formal observations I documented any thoughts in word documents, but all formal observations, meetings and interviews were audio recorded, (a feature included in the program.) I've obtained permission to study this through my existing position in the organization.

3.3. Data Collection

Initial Interviews: For each of the 36 1st time interviews, I briefed the interviewee off record, explained the aims of the study, and answer any questions they had. The 2nd thing done was I informed them the interview would be recorded for research and recall purposes. Using a semiformal conversational way of interviewing, I asked a set of standardized questions. 1st, I established their demographics through a few closed ended questions, and then progressed into more open-ended ones, such as describing their responsibilities, asking them to note any checks and balances of the structure, and their personal vision of the organization. When those were complete, I asked a few probing questions, which depended on the course of the interview, which were not previously scripted. Next, I asked them about their impressions of online social media, and lastly, I asked them if there were anything they would like to add. Then I wrapped it up and thanked them for their time and insights. **Follow-up Interviews:** These were conducted based on individual availability, during the target dates: (12/17/2012, 3/29/2013, 11/9/2013, 3/3/2014, and 7/25/2014.) Before each follow-up, I reviewed their previous recording, taking down notes for specific in-depth questions. If both of us were available the next round, I would repeat the process. Furthermore, as I went through the interviews, I refined the questions asked. And lastly, I went through the questions and responses and coded them for different themes.

Observations: These were done intermittently, usually informal, this means everyone has already been informed I am doing research and are just behaving naturally. These informal observations varied in duration. I recorded one informal observation, this was a training that took place with about 20-30 individuals, just acting normally during training. It serves as a benchmark for normal training activities.

Formal observations, however, are almost akin to an interview, however in this case, I am not asking the questions, an example of this is a meeting or discussion that naturally takes place to discuss structure, roles and responsibilities. As the secretary, it was my responsibility to take notes and record these formal meetings. The structure is in place, and weekly staff meetings were scheduled to share recent developments.

4. Results

Starting something is relatively easy, maintaining it over time is the hard part. Nothing is permanent, especially on the Internet, there are countless clans out there, yet the ones that last longer are the ones that understand the importance of defined roles & structure. In fact, once when I was recruiting, a person asked me what made my clan different from all the others, I immediately answered "structure", and he ended up joining, because I was the only clan recruiter that had ever had an answer for him. Furthermore, other times I had brought new recruits in, and they were amazed, since this was the most structured clan they had ever seen.

The interesting thing that I found from interviewing such a diverse community, was that when people get online, most of them want to relax and just have fun, however, fun

doesn't just happen in a vacuum, fun must be facilitated by those who are willing and able to serve. Clans can provide a sense of purpose, direction, and community, but only if truly dedicated individuals are at the helm. Another consistent finding is that "drama" will occur, period. Many of the more experienced staff that I interviewed realized that fact. Whenever you get people together, diverse or not, you will have conflict of some level, understanding that at the core, and having established procedures or individuals specifically dedicated to conflict resolution, at every level is critical. Furthermore, those at the top need to understand, at the end of the day, no matter what their personal perspectives or desires, drama management, aka conflict resolution is their main responsibility.

The next important point is that the 102nd was a democracy, most clans are dictatorships or autocracies. The game is based around small group tactics, and one of the concerning things that we had consistently was clicks, it's no surprise that these occur, but at times if left unchecked, these can be problematic, since they tended to be exclusionary. In fact, we had a commander, which was our longest running platoon Sgt., who wanted a more competitive atmosphere. Once he got elected, his concern was only for those who wanted the same thing. An email was circulated criticizing the commander rewarding the proletariat, which caused a schism. In the end, he left with his click, after his term, since the atmosphere became hostile. Although this was a rather depressing time in the clan's history, upon retrospect, many of the members that I interviewed that remember that time, felt that it was appropriate it turned out that way, since it helped define what the clan wasn't.

Throughout the history of the clan, there have been many shifts in power, most of them amicable, however, there was a shift that took place before the research began, which actually was the catalyst for why the research was started in the first place. Without going into gruesome detail, five individuals were forcibly removed from an older version of the clan, because they didn't agree with how the current regime was operating. Those that were cast out, managed to retain the name, and rebuild. The key lesson from this, is that the earlier regime before August 2012, didn't understand the key issue of service, by that, they didn't understand that the individuals that they had dispelled were actually the individuals that were responsible for keeping things running behind the scenes. Thus, that older regime shortly fell, giving birth to the opportunity of this research. I now call this understanding concept "Loadbearing".

Interviews done during that time focused on the theme of how deception might be easier in this environment, which is no surprise. However, the interesting thing is, despite popular belief, most people are actually quite genuine in the online gaming sphere. When discussing this environment with those unfamiliar with online gaming, they usually have the concern that more people are deceptive because they don't "see" them face to face, thus opening up the increased potential for deception. I'd counter that with the question, do you think people over the phone are more likely to be dishonest just because you're not in their presence? I've found consistently, people in this environment actually strive to be more communicative, since they are so limited. It's a clear testament to the fact that at the end of the day, people want to be understood. Although this study only is consistent with voice mediated technologies, I would propose such a phenomenon is seen throughout other technological facilitators. Furthermore, the common old communication adage, that a disproportionate amount of human communication is based on nonverbal cues may be overturned in such an environment.

Theme 1: Balancing teamwork vs. Individuality = freedom: This is perhaps the biggest theme, both in scope and in the importance of research. In any organization it's important to respect individual differences, while at the same time keeping in mind the goals of the group. Team cooperation is naturally one of the most important factors, however, often organizations lean too far one way or the other, and they either become too cohesive and stagnate or become too loose and are unproductive.

This concept can be best described as organizational and individual respect.

"And that is what I'm going to make sure happens, is all of us are going to treat each other with respect. Hey, I might think this guy is an ass hole, but if he is a good trainer, ok great, train the troops, because the bottom line about what it is, is about the guys in the clan, and how we can work together, so we have to get personalities together that can mesh" (Repair guy 1st XO, 2nd CO))

Theme 2: Friendship: Friendship is that connection with others that people strive for, that feeling of being understood and accepted, and a huge aspect of why clans exist particularly in the gaming sphere. In fact, a majority of the officers said both during interviews, and during informal observations, that they wouldn't be playing the game if it wasn't for the clan, that they were referring to the bonds of friendship they had formed. This potentially sounds like a case of collective ignorance, since numerous times they would complain about how frustrating the game was, yet upon asking why they still played, they said because of their friends here. This finding suggests that gaming is more about the social aspect, than the actual game itself. In other words, games are just facilitators. As a hobbyist game developer, this begs the question of whether the success of the games industry is due to the investment of technology or more the investment of social capital.

"Well, I actually want to be involved with the 102nd clan because I can make friends all around the world. One of the cool things about using this venue to do it is we can all get together, and quite frankly it's just a tool for me to get to know a bunch of guys around the world so we can trade ideas, and you know, this is similar to the elks lodge or moose lodge, where I might have certain personal problem with things, I can talk to a couple of my friends that are in Israel and me in Russia and get their idea on how they look at things, plus get their idea of how it is for them in their country and how they see us, understanding that we are a lot more similar than what we think or the media is trying to portray us, um, I'm very surprised how easy it is to make friends across the country and across the world in this respect" (Repair guy, 1st XO, 2nd CO)

The other important part about this theme is how you can get to know someone with few inherent biases, and this has the potential to lead to a clear unconditional positive regard from inside out, instead of outside in.

"No visual cues, they are deceiving, "everyone here, in their little blue dot I see on TS is their own individual, they are their words, emotions and actions to me, they're not their looks, I don't know what these guys look like, I don't care what they look like, I shouldn't care what they look like, I'm not judging the book by its cover, there is no cover here. It unbelievably frees people, but I believe it also makes deception easier, offline I likely put people into categories before I know how they think. I cannot know you by looking at you I can know you by us talking to you, that is the defining difference about using this interface By hearing what you say, how you say it, and how you respond to me, I know you far better than what you look like I think that's a paradigm shift and a social interchange, I'm not putting any preconceived ideas on how I think you are by how I perceive how you look" (Repair guy, 1st XO, 2nd CO)

This puts paralanguage in the spotlight, and potentially drastically increases its place in traditional communication. Allowing near isolation of this feature may lead to increased familiarity.

"Well, personal goals, uh, are goals for myself, um, I would not want to put my personal goals on the rest of the guys. Um, to be honest, my personal goals in this are to make sure everyone gets along, has fun, and we all enjoy being around each other, um, and have the kind of interpersonal relationship that is like best friends." (MrBlixx, 1st CO)

Theme 3: Responsibility: This theme in the current context refers to volunteer positions that a person has taken. In such an online informal environment, this really needs to be intrinsically motivated. The responsibilities that must be taken are quite important, and one of the most critical aspects of clan survival.

"So, it's a communication thing, uh, has been the problem, uh, in the past, confusion on that and the lack of focus, I'm very, I am very shocked at the lack of focus by the people who held jobs, that the understanding that you held a job for the group, don't do it for yourself, you do it for the group of people, when you're at a job, you show up for the people, and for the job." (Repair guy)

Theme 4: Organizational Balance: Every organizational system needs checks and balances if they are to survive the people that are in them. Balance in this context is easily defined as ways to insure a person, or a group of people aren't able to upset the organizational structure, no matter what position they are in. Although this is arguably the most important aspect both from a general standpoint, and from what I've found in my research, this concept is the last one because to fully grasp and implement balance, all the other themes need to be accommodated in the structure.

“There is complete check and balance. There is no I am God; this is how it is going to be.” (MrBlixx, 1st CO).

Unfortunately, this is where the clan finally reached its conclusion. The balance of the organization was not adequately maintained in the last regime. After the longest running commander stepped down after a year, in December of 2014, his XO was elected. This individual, although a strong tactical leader, failed to connect or impinged on all of the themes.

Throughout this research I’ve come to identify 5 key aspects of an online gaming community that can also be applied to any online or offline organization regardless of form function, size or capability. They are Vision, Purpose, Structure, Passion and Community. These are developed in sequential order or built in reverse. Each naturally leads into the next aspect. A change to any one of these core pillars of an organization will affect all the others.

Vision/ Mission: This encompasses the direction of the organization. The goals, aims, directives, the big picture of the end product etc. Alternative to vision is mission, this can be an altruistic goal for providing goods or services, a nonprofit aim or ideal to serve an underrepresented population or a social good.

Purpose: Relatively straight forward, once your vision is set you can figure out a function or role. This can either be on an individual basis, or an overall organizational basis. Alternatively, if you have a formalized structure, you can use it to define individual or organizational functions roles and goals leading us back to vision. (If you’re going at this model through the reverse.)

Structure: This refers to the formal organization of the organization, and what roles and responsibilities are formally documented and expected of the members. Once a purpose is laid down, a structure is needed to formalize everything. If you’re coming at it from the passion side, then what roles and responsibilities do the members of the organization naturally have excitement of doing?

Passion: This is a person’s excitement and inner drive to do something. Once we have the vision, the purpose, and the structure, a person develops a work ethic around those things. If it’s coming from the community end, they are developing a work ethic to serve their community.

Community: This encompasses the social aspects of the organization, the group dynamics, the sense of belonging and comradery, and arguably what we are all looking for in an organization.

5. Discussion

My findings supported the research of Bo Xu [1] who identified key factors which contribute to organizational citizenship. The themes of individual initiative, helping behavior, and organizational loyalty were particularly prevalent. Furthermore, the factors that define a role by Shen [7] are consistent, position, function/task, behavior expectation and social interaction.

Along with this foundation [4], findings of clan diversity were highlighted. In addition, Koch’s conclusions [2] on the importance of team Cooperation were also supported.

Despite this strong support, this study was limited in scope, since it only surveyed one clan, there are hundreds out there, spanning a plethora of games, each with their own structure and unique social dynamics. From group dynamics and communication in virtual environments, to ways to inspire and engage an online community, this research is a steppingstone in understanding communication at its most core foundations.

Leadership, training, engagement, these are all features of clans and work, as we look to the future of information-based organizations to the horizon, we must all strive to continue to engage in this dichotomous discussion. Only then will we ensure our place in this highly technological future.

References

- [1] Bo Xu AA, D.L. Knowledge sharing in virtual communities: A study. International Journal of Human-Computer Interaction. 2012;347–359.

- [2] Koch MH. Cooperation and disruptive behavior – learning from a multi-player Internet gaming community. Preprint - To appear in Proceedings of the IADIS International Conference Web-Based Communities 2004, Lisbon, Portugal, 24–26 March 2004;1–8 (2004).
- [3] Baur CK. Living a virtual life: Social dynamics of online gaming. *The International Journal of Computer Game Research*. 2004;4(1).
- [4] Ducheneaut N, Yee N. The life and death of online gaming communities. *CHI 2007 Proceedings Games*. 2007;839–848.
- [5] Poor N, Skoric MM. Death of a guild, birth of a network: Online community ties within and beyond code. *Games and Culture: A Journal of Interactive Media*. 2014;9(3):182–202. doi: <https://doi.org/10.1177/1555412014537401>
- [6] Zhang L, Fung AYH. Working as playing? Consumer labor, guild and the secondary industry of online gaming in China. *New Media & Society*. 2014;16(1):38–54. doi: <https://doi.org/10.1177/1461444813477077>
- [7] Shen C. The patterns, effects and evolution of player social networks in online gaming communities. 2010.
- [8] Wolfenstein M. Leadership at play: How leadership in digital games can inform the future of instructional leadership. 2011.

SECTION III

EVALUATION STUDIES

To date, some cybertherapy applications have improved the quality of health care, and later they will probably lead to substantial cost savings.

However, cybertherapy is not simply a technology but a complex technological and relational process.

In this sense, clinicians and health care providers that want to successfully exploit cybertherapy need a significant attention to clinical issues, technology, ergonomics, human factors and organizational changes in the structure of the relevant health service.

Wiederhold & Riva, 2004

My Circle: Development and Evaluation of an Online Social Networking Platform for Clinically Guided Peer-to-Peer Support for Young People with Mental Health Concerns

Brad RIDOUT^a, Krestina L. AMON^a, Rowena FORSYTH^a and Andrew J. CAMPBELL^{a,1}

^a*The University of Sydney, Cyberpsychology Research Group, Biomedical Informatics and Digital Health Theme, School of Medical Sciences, Faculty of Medicine and Health, New South Wales, Australia.*

ORCID ID: Brad Ridout, <https://orcid.org/0000-0001-8067-0490>

Abstract. Supporting the mental wellbeing of young people is critical, as half of all lifetime mental illnesses develop before age 14 and 75% before age 25. Despite this, many young people lack access to care due to safety and accessibility barriers, increasing the risk of chronic mental health disorders. As internet-connected mobile devices have become ubiquitous, young people increasingly seek mental health support online, particularly through social media. However, this can expose them to inaccurate information and hostile comments. To address this, yourtown and The University of Sydney developed My Circle, a safe, anonymous, and clinically guided peer-to-peer support platform for young people. My Circle is moderated by clinical professionals to ensure safe, respectful conversations based on accurate information. It aims to validate experiences, reduce stigma, and educate users. It is integrated with yourtown's Kids Helpline for 24/7 support. The My Circle platform was evaluated over a 36-week mixed-method study involving 2,791 young people aged 13-25, 63% of whom identified as women/girls. The evaluation showed that My Circle users experienced improved mental wellbeing and reduced psychological distress, reduced stigma, and increased coping skills and willingness to seek help. Those who entered the service with the most severe psychological distress reported smaller changes in distress levels over time but were more engaged with the service. My Circle demonstrates the efficacy of custom-built online social networking platforms in supporting young people's mental health, offering a scalable model of care for other vulnerable groups.

Keywords. Social media, online counselling, mental health, youth

1. Introduction

From 2017 to 2023, five evaluation phases gathered extensive data and insights into the value and role of Kids Helpline's My Circle as an online social networking platform for youth mental health support. My Circle provides young Australians aged 13-25 with a safe space to anonymously connect with peers, access accurate wellbeing information, and link to mental health services [2]. In Phase 1, Kids Helpline Circles (later renamed My Circle) on the open-source Elgg platform attracted youth seeking peer connection and participation in 8-week online psychoeducation modules, but faced high disengagement, partly due to platform usability issues [2]. Phase 2 transitioned to the HumHub platform, showing strong user satisfaction, symptom reduction, and improved coping skills [3]. Phase 3's qualitative analysis highlighted positive peer support but noted technical and engagement challenges, as well as a desire from users for ongoing open discussion forums [4]. Following the participatory action research findings of Phases 1-3, the service operating model for the platform was reconceptualized and rebranded as My Circle, featuring separate sub-forums called 'Circles', where users congregate to discuss specific topics related to mental well-being (selected based on user feedback). Phase 4 trialed this reconceptualized operating model following other

¹ Corresponding Author: Andrew J. Campbell, andrew.campbell@sydney.edu.au

implemented changes including a streamlined entry process, briefer feedback surveys, and platform functionality updates [4]. User feedback was very positive, with higher engagement and significant reductions in psychological distress reported. The current study, Phase 5, involved a 36-week naturalistic evaluation of the newly established operating model for My Circle, which was relaunched following further technological and service enhancements based on user and moderator feedback from the Phase 4 trial. The Phase 5 evaluation measured user experience, satisfaction, and engagement, and changes in psychological distress over time.

2. Methods

Between February 2022 to October 2022, all My Circle users (self-selected Australian youth aged 13 – 25) were asked to complete a pop-up survey upon signing in to the My Circle social network. This survey included a mandatory Kessler Psychological Distress Scale (K10) questionnaire [5] that provided baseline data. Demographic data was obtained from information already collected when users created their My Circle profile upon signing up to the platform. Non-mandatory follow-up surveys (also pop-ups within My Circle) that included the K10, and user experience and satisfaction questions, were presented to users when they logged in to My Circle at a monthly cadence.

3. Results

3.1. Demographic

There were 2,900 respondents to the My Circle baseline survey. Removal of unusable data (e.g., incomplete surveys, and surveys returned that were ‘straight-lining’ answers), resulted in a baseline sample of N=2,791 for analysis. The mean age at baseline was 15.36 years ($SD=2.26$), with the sample skewed towards younger participants (64.1% were under 16 years, and only 3.0% were over 21 years). Most of the sample identified as women/girls (62.7%), with only 12.4% as men/boys. Gender-diverse, non-binary and custom gender responses (combined for analysis and referred to collectively from here as ‘gender-diverse’ [6]) made up 16.7% of the sample, with 8.2% either not declaring or missing input for gender identification. Most participants (72.2%) resided in major cities, 21.2% in regional areas and 1.4% in remote areas, with 5.2% not declaring their postcode. Gender-diverse individuals were more highly represented outside of major cities, in particular remote locations (26.5%), and were significantly younger on average ($M=14.8$ years, $SD=1.78$) compared to cis-gender participants ($M=15.48$ years, $SD=2.33$; $t(2551)=6.66$, $p<.001$), suggesting that My Circle may be particularly appealing to young gender-diverse individuals.

3.2. Psychological Distress at Baseline

At the relaunch of My Circle in February 2022, 70% of users reported severe psychological distress, indicating a high likelihood of severe mental disorders. The average distress levels among My Circle users at baseline were notably higher than those reported in face-to-face studies of young people seeking mental health support [7]. Gender-diverse users reported significantly higher distress ($M=36.57$, $SD=7.09$) compared to women/girls ($M=33.23$, $SD=8.04$; $p<.001$), who in turn reported higher distress levels than men/boys ($M=31.21$, $SD=8.51$; $p<.001$).

3.3. Psychological Distress Over Time

Figure 1 shows a breakdown of K10 scores according to how many follow-up surveys participants completed. Though not significant, there was a general trend in which the more follow-up surveys participants completed, the higher their K10 scores were at baseline on average. Users who initially reported higher levels of distress tended to remain engaged with My Circle and its follow-up surveys for a longer duration. Over time, there was a significant linear increase in K10 scores among users who demonstrated

maintained engagement with My Circle for at least six months by completing 5 or more follow-up surveys ($F(1,89)=4.38; p<.05; \eta^2=.047$). This increasing and ongoing distress may have contributed to their sustained engagement with the platform.

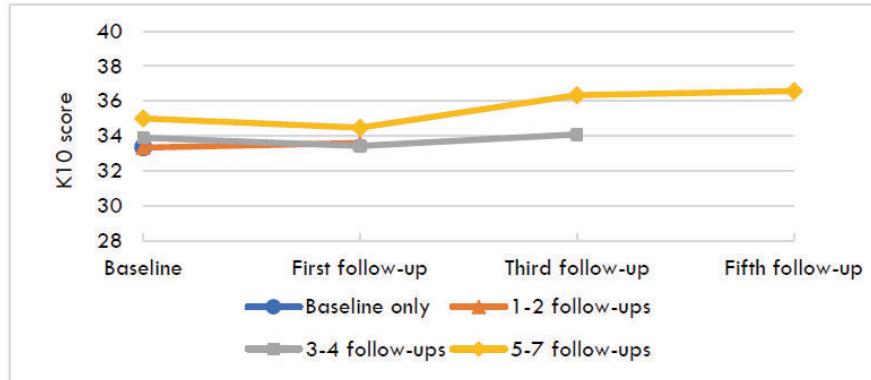


Figure 1. Mean psychological distress scores (K10) at different time points by number of follow-ups completed.

3.4. User Experience and Satisfaction

The introduction of technological and service improvements in Phase 5 led to significant increases in user experience and satisfaction scores. Despite no significant overall changes in psychological distress scores on the K10, more than half of My Circle users felt their overall mental health had improved due to the platform (56.0%). Users who remained engaged for at least three months were more likely to report mental health improvements at their first follow-up (67.3%) compared to their third follow-up (51.1%; $t(203)=3.27, p<.01$). This finding reflects the increased distress reported by this subsample over time but may also indicate changing perceptions as users became longer-term participants. Longer engagement was also associated with increased feelings of safety, acquisition of new coping skills, reduced stigma around seeking help for mental health issues, and greater awareness of mental health services (full results published in Ridout et al. [4]).

3.5. User Engagement

During the 36-week evaluation period, there were 16,849 user posts (average of 468 posts per week). There were 1,470 unique posters, representing half of all registered users. Additionally, there were 25,204 comments (700 per week) and 24,379 likes (677 per week). The average session length was 15:28 minutes, during which users viewed an average of 6.75 pages. The most active Circles in terms of posts and comments were ‘Mental health’, ‘Suicidal thoughts and self-harm’, and ‘The Vibezzz’. Posts in the ‘LGBTIQAP+’ Circle received the most engagement, with around two comments per post and 0.75 likes per comment/post. Engagement dropped during school holidays and increased with the resumption of the school term.

In addition to the ‘discussion board’ format Circles, self-help courses similar to the psychoeducation modules in Phases 1 and 2 were also available. These courses received lower engagement in terms of comments and likes in comparison to the Circles but saw high engagement with poll question activities. The most engaged self-help courses were ‘Managing Anxiety’, ‘Dealing with depression’, and ‘Family relationships and conflict’.

There were 906 private message conversations between clients and moderators, averaging 25 per week. Half of these private messages were ‘follow-up of risk’ messages initiated by moderators to check on users who were identified as potentially at risk based on their posts or comments. The second most common reason for sending private messages was to address rule breaches, accounting for 21% of these conversations.

4. Discussion

My Circle aims to improve the social and psychological wellbeing of young Australians by offering clinically supervised anonymous peer-to-peer support, access to accurate wellbeing information, and referrals for individuals in distress. Phase 5 showcases the success of these goals through participatory action research, involving collaboration with young people, stakeholders, and experts to ensure safety, accessibility, and inclusivity.

4.1. Safety

My Circle provides a safe online space where young people can share experiences without judgment. Users found it safer and more positive than other social media platforms. Clinical moderators screened new accounts, assisted distressed members, and ensured strict anonymity. Despite some requests from users for private messaging with other users, all user-to-user communication remains restricted to open conversations within ‘Circles’ for better monitoring. Further safety improvements implemented following launch included a ‘post quarantine’ feature to flag posts containing high-risk keywords for moderator review before they are published and tracking of potentially at-risk users and duplicate accounts from banned members.

4.2. Accessibility

My Circle addresses major barriers to mental health support, such as confidentiality, stigma, cost, and accessibility. It is a free service, available 24/7, and offers a clinician-supervised alternative to public social networking sites. Evaluation results show that My Circle improved mental health, reduced self-stigma, and increased awareness and help-seeking behavior among young people [4]. It engaged those who might otherwise not seek help, providing early intervention and pathways to additional services.

4.3. Inclusivity

The development and evaluation of My Circle adopted a person-centered approach, involving young people in decision-making. It supported those with impairments through features like voice-to-text and screen reader compatibility. It appealed to gender-diverse individuals and those in remote areas, with the ‘LGBTIQAP+’ Circle among the most popular with users. The platform continually adapted to meet the diverse needs of its users, including those with unique conditions like dissociative identity disorder.

5. Conclusion

My Circle, a community of over 10,000 young Australians, was developed through evidence-based refinement with active input from young people, counsellors, and stakeholders. Led by the University of Sydney’s Cyberpsychology Research Group, this participatory approach enabled continuous improvement. Prioritizing safety, accessibility, and inclusivity, My Circle engaged young people who might not otherwise seek support, offering early intervention and pathways to further services. Clinical moderators supervise interactions, providing validation, crisis support, and enhancing emotion regulation and social skills. My Circle effectively overcomes barriers to mental health support, normalizes help-seeking, and reduces psychological distress, offering a scalable, cost-effective care model.

References

- [1] Kids Helpline. My Circle [Internet]. Brisbane (AU): Kids Helpline; 2024 [cited 2024 July 3]. Available from: <https://kidshelpline.com.au/my-circle>
- [2] Campbell A, Ridout B, Amon K, Navarro P, Collyer B, Dalgleish J. A customized social network platform (Kids Helpline Circles) for delivering group counseling to young people experiencing family discord that impacts their well-being: exploratory study. Journal of Medical Internet Research. 2019;21(12):e16176. doi: <https://doi.org/10.2196/16176>

- [3] Amon K*, Ridout B*, Forsyth R, Campbell A. Online group counseling for young people through a customized social networking platform: Phase 2 of Kids Helpline Circles. *Cyberpsychol Behav Soc Netw.* 2022;25(9):580–588. doi: <https://doi.org/10.1089/cyber.2022.0044>
- [4] Ridout B, Amon K, Forsyth R, Campbell A. Development and evaluation of Kids Helpline's My Circle. Sydney (NSW): The University of Sydney; 2023.
- [5] Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, Walters EE, Zaslavsky AM. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med.* 2002;32(6):959–976. doi: <https://doi.org/10.1017/S0033291702006074>
- [6] Cameron JJ, Stinson DA. Gender (mis)measurement: guidelines for respecting gender diversity in psychological research. *Soc Personal Psychol Compass.* 2019;13(11):e12506. doi: <https://doi.org/10.1111/spc3.12506>
- [7] Rickwood DJ, Mazzer KR, Telford NR, Parker AG, Tanti CJ, McGorry PD. Changes in psychological distress and psychosocial functioning in young people visiting headspace centres for mental health problems. *Med J Aust.* 2015;202(10):537–542. doi: <https://doi.org/10.5694/mja14.01696>

Application of Abstract Generative Imagery for Psychological Diagnosis: A Pilot Study

Andrey V. VLASOV^{a,b, 1}, Natalia E. SERGEEVA^c,

Vasily V. SERIKOV^d and George A. KARKASHADZE^c

^a*HSE University, International Lab for Experimental and Behavioural Economics*

^b*DDT*

^c*Research Institute of Pediatrics and Children's Health, Scientific Center No.*

2, Russian Scientific Center for Surgery named after. acad. B.V. Petrovsky

^d*Russian Medical Academy Continuing Professional Education, Department of*

Occupational Pathology and Occupational Medicine, Moscow

ORCiD ID: Andrey V. Vlasov <https://orcid.org/0000-0001-9227-1892>

Abstract. *Purpose:* of this pilot study was to determine the practical use of AI-generated images for psychological diagnosis in an adult group. *Methods.* A set of abstract images generated by a text-to-image model was developed and applied. Prompting (creating a text input) included "polar words" that created a visual dichotomy, positive or non-positive coloration of the images presented: e.g., "calmness – excitement". The following psychological techniques were used in the experiment to test correlations: Satisfaction with Life Scale, Five Well-Being Index, Questionnaire for the determination of accentuated personalities (Schmieschek), Hospital Anxiety and Depression Scale. *Results.* The main findings are: (i) Correlations were found between preference for positive visual content in the high life satisfaction group with the "pacification and rabies" dichotomy and the "delight and despondency" dichotomy; and SWLS score with the "delight" prompt, as well as WHO5 volume ($p<.05$ and $p<.05$, respectively). (ii) When assessing the relationship between individual image dichotomies of concepts and psychological characteristics, it was revealed that 5 out of 10 personality accentuation scales were significantly correlated with the coloring of preferred generative content. Correlations were found for the visual dichotomy of "calmness and excitement" and the following personality accentuations (Schmieschek): demonstrative, hyperthymic, and dysthymic types. *Conclusion.* The study has demonstrated the potential use of AI-generated images in diagnosing (dis)satisfaction with life, distress, increased anxiety, as well as excitable personality traits. The identified connections suggest that this stimulus is effective for diagnosing states and psychological traits characterized by heightened sensitivity to external visual content.

Keywords. Generative images, delight, psychological diagnosis, satisfaction with life scale, accentuated personalities, anxiety

1. Introduction

In recent years, the development of artificial intelligence (AI) technologies has significantly impacted numerous aspects of our lives, including scientific research and practice across various fields. New frontiers are emerging in psychology as well: AI is being actively integrated into psychological counseling practice (development of bots for remote psychotherapy) and into scientific psychological research. One of the promising avenues for assessing diverse psychological constructs is the use of generative images. Projective techniques are employed in psychodiagnostics, wherein the presented stimulus material exhibits a reduced level of structuredness, thereby facilitating the overcoming of defensive attitudes and conscious control exerted by the subject.

Frequently, the test material utilized in projective techniques is visual. A notable example of the utilization of visual projective techniques is the Rorschach inkblot test, which was developed by the psychiatrist Rorschach. This test involves presenting subjects with a series of inkblots of varying colors and saturation levels, and then

¹ Corresponding Author: Andrey V. Vlasov, avvlasov@hse.ru; study@demonroyal.com.

analyzing their responses to identify specific personality traits. Despite its historical significance, the relevance of the Rorschach inkblot test has diminished in modern clinical practice due to changes in sociocultural contexts and conditions, as well as improvements in quality of life.

In contemporary clinical settings, various forms of visual projective techniques remain actively employed. However, the relevance of the material utilized in these techniques is contingent upon the aforementioned sociocultural and contextual factors. The application of AI technologies for the development of psycho-diagnostic tools enables the expedited modernization and adaptation of the material, thereby expanding the scope of its applicability and facilitating its integration into practice for specialists. Consequently, the objective of this pilot study was to examine the potential for the practical utilization of generative images in psychological diagnostics.

2. Methods

The study employed generative art images as stimuli for psychological assessment. The images were employed to obtain feedback from the user and to ascertain his or her mental state (wellbeing) for diagnostic purposes. Concurrently, we sought to mitigate the influence of the emotional component on the choice of any color by relying on the context (hint – prompt) of the generated image and using a set of images in a single-color scheme. Polar keywords were employed as the context for the "input" cue.

The set included two-dimensional images (GALA [2,3]), which were combined into ontological pairs — categories of words with opposite meanings. The pairs were as follows: 1) anger and kindness; 2) wrath and forgiveness; 3) rabies and pacification; 4) aggression and peacefulness; 5) excitement and calmness; 6) fear and courage; 7) despondency and delight; 8) love and hate; 9) life and death.

A three-part plan was developed to evaluate the potential of using generative imagery in psychological diagnosis. The first part of the plan was designed to identify the ability of the method to be a screening tool for determining the subjective level of satisfaction and well-being as a general characteristic of emotional state. For this purpose, the Satisfaction with Life Scale (SWLS) [4] and Wellbeing Index (WHO-5) [5, 6] were included in the research plan.



Figure 1. Anger - kindness [2; 3]. CC BY-NC-ND.

The second part of the study was designed to assess the potential of using the generated images in diagnosing characterological features of personality. To this end, the examination plan included Schmieschek's personality accentuation test [7], which is based on Leonhard's typological approach [8; 9]. The questionnaire consists of 88 statements, and the respondent was required to assess whether they correspond to their behavior. The result of the method is an assessment of the expression of each characterological type, as follows: 1) Demonstrative; 2) Pedantic; 3) Stuck; 4) Excitable; 5) Hyperthymic; 6) Dysthymic; 7) Anxious and fearful; 8) Cyclothymic; 9) Affective exalted; 10) Emotive.

The final objective of the study was to ascertain the potential for clinical utilization of the methodology. The preferences of generative images were compared with the results of the Hospital Anxiety and Depression Scale (HADS) questionnaire, which was designed to identify the level of anxiety and depressive disorders in practice by clinical

specialists. The HADS simultaneously contains questions designed to diagnose anxiety and questions designed to detect depression. A score of 11 points or higher indicates the presence of clinical anxiety and/or depression.

The data were subjected to Spearman correlation analysis using IBM® SPSS Statistics 26 (SPSS Inc.). The hypothesis that there is a correlation between the total score on the method of preferences of generative images with positive coloring and the results of psychological diagnostics was tested, as well as separately between individual preferences in a pair — a set of images with dichotomous semantic content and psychological techniques.

The pilot sample, which was analyzed, included 30 respondents (15 women), with a mean age of 30.2 years (SD 7.792).

The study was approved by the local ethical committee of the IRIOH (No. 4; May 25, 2022). The experimental procedures were conducted in accordance with local legislation and the Declaration of Helsinki.

3. Results

The surveyed sample is defined by the following indicators (**Fig. 1**).

Table 1. The overview of characteristics.

Measurable indicators, name / (scale)	Mean	SD
Choosing positively colored images	6,67	7,34
Satisfaction with life (SWLS)	23,73	6,63
Well-being index (WHO-5)	15,52	4,32
Demonstrative type	13,20	4,25
Stuck type	13,47	3,32
Pedantic type	10,27	4,45
Excitable type	10,10	4,69
Hyperthymic (hypomanic) type	14,70	5,81
Dysthymic type	9,30	5,65
Fearful type	6,50	5,05
Cyclothymic type	12,30	5,00
Exalted type	13,80	5,93
Emotive type	14,90	4,42
Anxious type (HADS)	6,20	3,22
Depressive type (HADS)	4,33	2,45

In general, all respondents can be characterized by the presence of average values of life satisfaction, normal values of subjective well-being, and accentuation of such personality traits as demonstrativeness, stuckness, hyperintimacy, cyclothymicity, exaltedness, and emotionality. Additionally, there is an absence of clinical expression of symptoms of anxiety and depression.

Subjective well-being and level of life satisfaction. Correlation analysis of the total score calculated on the basis of all generative image selection preferences (nine pairs of opposite prompts, as detailed in the Methods section) and the SWLS and WHO-5 techniques revealed no significant relationships. However, relationships were identified for individual dichotomies (word pairs), wherein the highest level of life satisfaction (SWLS) was associated with a preference for meaningfully more positive images in the dichotomy "forgiveness – wrath" ($r(30)=0.437$, $p<.05$) and in the dichotomy "delight – despondency" ($r(30)=0.433$, $p<.05$). Those with higher subjective well-being (WHO-5) scores were more likely to select the image generated by the "delight" prompt ($r(30)=0.382$, $p<.05$).

Characterological features. Correlation analysis of the preference for generated images and the expression of personal traits revealed a significant association between the selection of meaningful negative visual content and the expression of excitable personality traits ($r(30)=0.429$, $p<.05$). Additionally, the highest number of significant relationships was observed for the "calmness – excitement" and "kindness – anger" dichotomies. In the initial case, it was determined that the expression of dysthymia as a personality trait was associated with the selection of images representing the prompt "calmness" ($r(30)=0.366$, $p<.05$). This was observed particularly among respondents

exhibiting more pronounced those exhibiting demonstrative and hyperthymic traits selected images representing the "excitement" prompt as a pole ($r(30)=-0.477$, $p<.01$; $r(30)=-0.425$, $p<.05$). The same traits are also associated with the second dichotomy: dysthymia is positively correlated with the selection of the "kindness" prompt as a pole ($r(30)=0.463$, $p=.01$), while anger is chosen by respondents with traits of demonstrativeness ($r(30)=0.384$, $p<.05$) and hyperintimacy ($r(30)=0.400$, $p<.05$).

In regard to the ten personality traits, five demonstrated at least two significant correlations with the selection of a specific pole in the semantic dichotomy of the generated images. These included traits of demonstrativeness, pedantry, excitability, hyperintimacy, and dysthymicism. Two traits, cyclothymy and affectivity, exhibited a single statistically significant correlation. Conversely, three traits, stuckness, fearfulness, and emotionality, did not demonstrate any correlations.

Depressive and Anxiety Symptoms. No significant relationships were observed between depressive symptom severity and content preference, either overall or with individual dichotomies. However, a positive correlation was observed between preference for semantically positive images and lower clinical anxiety symptom severity scores ($r(30)=0.388$, $p<.05$). Conversely, higher anxiety was also associated with a greater preference for images with the meaning of "despondency" over in the "despondency – delight" dichotomy ($r(30)=0.548$, $p<.01$), as well as images with the meaning of "aggression" over "peacefulness" ($r(30)=0.386$, $p<.05$).

4. Discussion

In general, the results provide partial confirmation of the hypotheses. With regard to the objective of employing the method as a screening tool for the assessment of subjective wellbeing and life satisfaction, the selected test material did not demonstrate the anticipated efficacy. Nevertheless, connections with separate sets of images were revealed, which were generated according to the concepts semantically close to the concept of subjective wellbeing and life satisfaction ("peacefulness – aggression"; "delight – despondency"). It is noteworthy that the semantic dichotomies themselves have a high intensity of emotional coloring. It is reasonable to posit that the potential of this method will be more fully realized if we form a set with semantic values of intense emotional content, excluding the values of another domain with different emotional states. Furthermore, it is proposed that words with a high level of expression of negative emotions (e.g., despair, hopelessness, etc.) should be included in future attempts to diagnose depressive states. This assumption is based on the observation that both subjective well-being and depressive states are typified by the expression of negative emotions and a negative evaluation of one's life circumstances. Consequently, it may be valuable to explore the potential for identifying shared criteria for the development of test materials for these states. The results of the second stage indicate that sensitive dichotomies of concepts are revealed for the majority of the measured characterological traits. This is a notable outcome that suggests the potential for AI to be utilized in the generation of stimulus material for personality profile diagnosis, contingent upon further examination of semantically sensitive material pertaining to the characterological traits of those who are stuck, fearful, and emotive.

The fact that excitable, demonstrative, and hyperintense traits demonstrated sensitivity to this test material can be attributed to the fact that these traits are characterized by increased reactivity to external influences and emotional lability. This can also account for the higher diagnostic efficiency of the generated images for other phenomena and states associated with these traits, including specific types of psychological defenses and mental disorders characterized by an increased risk of development with these characterological traits.

A similar mechanism can be assumed in the case of a more pronounced level of anxiety, which is also characterized by a pronounced sensitivity to external stimuli [13]. At the same time, the general tendency to choose negatively colored imagery indicates a possibly lower significance of negative connotations than for other psychological features. This indicates that the development of test material for diagnosing anxiety is the most promising area for further research.

With the noted potential of the method, it is important to emphasize the limitations of the study. The narrowness of the sample, the potential influence of other factors, and the insufficiently high indicators of the significance level in some cases indicate the need for caution in drawing conclusions. However, given that the purpose of the study is to determine the potential and possible avenues for further research into the use of AI for psychological diagnostics, it can be concluded that the results obtained indicate positive prospects for this direction. It is similarly important to note the necessity of including the results of psychiatric diagnostics at subsequent stages when evaluating the feasibility of the method in clinical practice.

5. Conclusion

Our work considers the potential of using generative images for psychological diagnostics in relation to the tasks of assessing symptoms of anxiety states and characterological traits of personality. For the diagnosis of subjective well-being, life satisfaction, and depressive symptoms, additional refinement of the material is required – the potential of using generated images in the case of special selection on the basis of concepts semantically close in content to these constructs is shown.

Acknowledgment

We thank Ekaterina Kochetkova, Egor Torsky (Yuri), other anons; all volunteers who took part in the online experiments; social media communities.

Declaration of generative AI and AI-assisted technologies

During the preparation of this work the author used the generative model text-to-image [Kandinsky] in order to preparing experimental materials.

References

- [1] Garb HN, Wood JM, Lilienfeld SO, Nezworski MT. Roots of the Rorschach controversy. *Clin Psychol Rev*. 2005;25(1):97-118. doi: <https://doi.org/10.1016/j.cpr.2004.09.002>.
- [2] GALA KLIMT. Digital art collection of pictorial poems. Ridero; 2022. 162 p. ISBN: 978-5-0059-2526-8. Available from: <https://www.amazon.com/dp/B0BN...CY4>.
- [3] Vlasov A. GALA inspired by Klimt's art: image-to-text processing with implementation in interaction and perception studies: library and case examples. *Annu Rev CyberTher Telemed*. 2023;21:200-5.
- [4] Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction with life scale. *J Pers Assess*. 1985;49(1):71-5. doi: https://doi.org/10.1207/s15327752jpa4901_13. Available from: https://www.tandfonline.com/doi/abs/10.1207/s15327752jpa4901_13.
- [5] The World Health Organization – Five Well-Being Index (WHO-5). Available from: <https://www.corc.uk.net/outcome-experience-measures/the-world-health-organisation-five-well-beingindex-who-5>.
- [6] Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. *Psychother Psychosom*. 2015;84(3):167-76. doi: <https://doi.org/10.1159/000376585>.
- [7] Schmieschek H. Fragebogen zur Ermittlung akzentuiert er Persönlichkeiten [Questionnaire for the determination of accentuated personalities]. *Psychiatr Neurol Med Psychol (Leipz)*. 1970;22(10):378-81. German.
- [8] Leonhard K. Accentuated personalities. Kiev; 1989. 375 p. Russian.
- [9] Leonhard K. Akcentuirovannye lichnosti [Accentuated personalities]. Moscow: Eksmo-Press; 2002. 358 p. ISBN: 5-04-008782-9. Russian.
- [10] Snaith RP, Zigmond AS. The hospital anxiety and depression scale. *BMJ Clin Res Ed*. 1986;292(6516):344. doi: <https://doi.org/10.1136/bmj.292.6516.344>.
- [11] Diener E. Subjective well-being. *Psychol Bull*. 1984;95(3):542-75. doi: <https://doi.org/10.1037/0033-295X.95.3.542>.
- [12] Beck AT. Depression: clinical, experimental and theoretical aspects. New York: Harper and Row; 1967. 370 p.
- [13] Beck AT, Emery G, Greenberg R. Anxiety disorders and phobias: a cognitive perspective. New York: Basic Books; 1985. 368 p.

Exploring the Relaxing Effects of Hallucinatory Simulated Experiences in Virtual Reality: A Pilot Study

Giulia BRIZZI^{a,b,1}, Chiara PUPILLO^{a,c,l}, Clara RASTELLI^{d,f}, Antonino GRECO^{e,f,g}, Luca BERNARDELLI^h, Anna Flavia DI NATALE^h, Silvia Francesca Maria PIZZOLI^h, Elena SAJNO^{a,c}, Fabio FRISONE^{a,i}, Andrea UBALDI^b, Daniele DI LERNIA^{a,b,i,l}, and Giuseppe RIVA^{a,b,i,l}

^a*Humane Technology Lab, Università Cattolica del Sacro Cuore, Milan, Italy*

^b*Applied Technology for Neuro-Psychology Lab, IRCCS Istituto Auxologico Italiano, Milan, Italy*

^c*Department of Computer Science, University of Pisa, Pisa, Italy*

^d*Department of Psychology and Cognitive Science, University of Trento, Rovereto, Italy*

^e*Department of Neural Dynamics and Magnetoencephalography, Hertie Institute for Clinical Brain Research, University of Tübingen, Germany*

^f*MEG Center, University of Tübingen, Tübingen, Germany*

^g*Center for Integrative Neuroscience, University of Tübingen, Tübingen, Germany*

^h*Become-Hub, Milan, Italy*

ⁱ*Department of Psychology, Università Cattolica del Sacro Cuore, Milan, Italy*

ORCID ID: Giulia Brizzi <https://orcid.org/0009-0000-7472-742X>

Abstract. Anxiety is a phenomenon affecting individuals' psychological well-being, often leading to seeking relief through substances such as psychedelics. DeepDream algorithm enabled the simulation of visual hallucinatory experiences often associated with substance use. These experiences are called Hallucinatory Visual Virtual Experiences (HVVEs). Previous studies have found that some participants find such experiences calming and relaxing, but the specific effects on autonomic activation remain still unexplored. This pilot study investigated whether HVVEs influenced self-reported anxiety and autonomic response. 20 healthy participants assessed their state level of anxiety, and subsequently, their Heart Rate Variability. Participants were then exposed to an immersive naturalistic computer-generated landscape and its "DeepDream" version. After each viewing, they had to rate their current anxiety levels, and autonomic activation was recorded. Results revealed that artificial hallucination had an anxiolytic effect while maintaining physiological activation. This suggests HVVEs might promote attention shifting from ruminating thoughts and a form of eustress, mimicking post-psychadelic effects. Future research should validate these results. This opens new avenues for HVVEs use for anxiety-related disorders.

Keywords. Virtual reality, hallucination, anxiety, cyberdelics, hallucination machine

1. Introduction

Anxiety is among the most prevalent symptoms worldwide, significantly affecting individuals' physical and emotional health and causing difficulties across different life domains [1]. Symptoms can vary among individuals and can include persistent fear, nausea, and palpitations, ranging from mild to severe. Current treatments, including psychological therapy and pharmacologic agents like selective serotonin reuptake inhibitors, are not always effective or well-tolerated [2]. Consequently, many individuals do not respond adequately, leading to the persistence of anxiety and stress-related conditions, as well as treatment-resistant anxiety disorders.

¹ Corresponding Author: Giulia Brizzi, giulia.brizzi@unicatt.it

¹ These authors contributed equally.

1.1. Psychedelics and Cyberdelics to Alleviate Anxiety

Psychedelics are substances capable of altering cognition, mood, and sensory perception. These substances encompass a wide range of mechanisms of action [3]. Psychedelics refer to substances that produce their psychotropic effects primarily by acting as agonists or partial agonists of 5-hydroxytryptamine type 2A (5-HT2A) receptors. However, psychedelics have also been shown to interact with other receptors such as 5-HT1A, 5HT2B, and 5-HT2C [2]. Research has shown that psychedelics can produce visual and auditory hallucinations, dissociation, and mystical experiences that can alter mood, thinking patterns, and sensory perception, potentially contributing to therapeutic effects in dealing with anxiety. A recent review indeed reported the efficacy of these substances in reducing anxiety symptoms, increasing self-perception, and increasing social function in patients with generalized anxiety disorder, social anxiety disorder, or anxiety attributable to another medical condition [2]. However, the use of psychedelics involves inherent risks and faces legal restrictions in many countries. To address these challenges, recent research simulated psychedelic-like experiences through the combination of Immersive Virtual Reality (IVR) and Artificial Intelligence (AI), which were named “Cyberdelics” [3]. Previous research showed that artificial hallucinations generated using DeepDream algorithm – i.e., Hallucinatory Visual Virtual Experiences (HVVEs) [4] – were able to improve cognitive flexibility [5] and elicit positive emotional states [6].

1.2. The Present Study

Cyberdelics field presents promising opportunities for treating various psychopathological conditions, including mood and anxiety disorders. Indeed, it offers the unique opportunity to exploit psychedelic's therapeutic effects while reducing side effects. The current study aimed to investigate if and how HVVEs affect self-reported anxiety and autonomic activation, specifically Heart Rate Variability (HRV).

2. Methods

2.1. Participants

20 healthy participants (9 females, 10 males, 1 non-binary; mean age = 24.10, SD = 2.55) took part in this pilot study. Only participants without self-reported neurological conditions, prior visual hallucinatory experiences, and the habitual use of psychoactive substances were included. The study was approved by the Ethical Committee of Catholic University of Sacred Heart.

2.2. Procedure

Participants completed a questionnaire to collect sociodemographic information, followed by the State Anxiety Inventory (STAI-S). Subsequently, HRV was measured for 2.5 minutes. Participants were then presented with the first IVR experience. Afterward, they filled out the STAI-S, and HRV measurements were recorded a second time. Lastly, the second IVR experience was proposed following the same procedure. The order of the video presentation was counterbalanced. The study was a within-subjects experimental design. The entire procedure lasted approximately 50 minutes. This sample is part of an ongoing study with a larger sample size.

2.3. Measures and Instruments

- STAI-S [7, 8]: The STAI-S is a self-report questionnaire to measure current anxiety levels. It comprises 20 items and respondents must rate on a 4-point Likert scale (from not at all to very) their level of agreement with each statement. Higher values indicate greater state anxiety.

- HRV: HRV was recorded using a CorSense from EliteHRV, which acquires pulse oximetry data from the finger (sampling at 500 Hz). Data were automatically collected and processed by the KubiosHRV smartphone app, extracting time and frequency indexes of autonomic activity. In this study, Heart Rate (HR), mean RR, High Frequency (HF), Low Frequency (LF), LF/HF ratio, Stress Index (SI), parasympathetic nervous system (PNS) index, and sympathetic nervous system (SNS) index were specifically considered [9].
- Videos: "The Secret Garden" is a 10-minute VR video that immerses users in a Japanese garden. It has been used in previous research within a relaxing protocol to contrast work stress and burnout, showing promising results in inducing relaxation [10]. This video served as a control condition (CC). Its hallucinated version was developed by applying the DeepDream algorithm to the original version (experimental condition (EC); [5]; **Figure 1**).



Figure 1. Immersive experiences were used in the current study. On the left is the experimental condition (hallucinated condition obtained through DeepDream), and on the right is the control condition (original version "The Secret Garden").

3. Results

Statistical analyses were performed in Jasp (v 0.18). Natural Logarithm (NL) transformation was applied to HF and LF measures (NL HF and NL LF) to normalize the data and reduce variance [11]. For physiological measures, outliers were removed using the InterQuartile Range. The normality of the variables was assessed using the Shapiro-Wilk test ($p = <0.05$). We conduct repeated measures (RMs) ANOVA to test differences across conditions (i.e., baseline, CC, EC) in STAI-S scores and autonomic activity indexes [12, 13]. Greenhouse–Geisser correction was applied when the assumption of sphericity was violated, and post hoc contrast was performed applying Bonferroni correction. In the case of non-normal distribution, the Friedman test was used.

3.1. STAI-S

RM-ANOVA revealed a significant main effect of Condition ($F(2,38) = 49.39, p < 0.001$, $n^2 = 0.72$) on state anxiety level. Post hoc comparison showed that self-perceived state anxiety significantly diminished after IVR videos as compared to the baseline, with no significant differences between EC and CC.

3.2. HR and Mean RR

RM-ANOVA revealed a significant main effect of Condition ($F(2, 30) = 5.29, p = 0.01$, $n^2 = 0.26$) on HR. Post hoc comparison showed that HR significantly diminished only after CC compared to the baseline, and no other significant differences emerged. These results were corroborated by mean RR results ($F(1.44, 24.58) = 6.78, p = 0.009$, $n^2 = 0.28$).

3.3. SNS Index

RM-ANOVA revealed a significant main effect of Condition ($F(1.33, 0.19) = 6.96, p =$

0.003, $n_2 = 0.30$) on SNS index. Post hoc comparison showed that the SNS index decreased after CC as compared to the baseline, and no other significant differences emerged.

Table 1 reports comparisons in detail.

No significant results were observed instead concerning other physiological indexes – specifically, NL LF ($p= 0.82$), N HF ($p = 0.10$), SI ($p = 0.38$), PNS index ($p= 0.06$), LF/HF ratio ($p= 0.22$).

Table 1. This table presents the results of the post hoc Bonferroni tests following an RM ANOVA for a within-subjects design. B= baseline; CC= Control Condition; EC= Experimental Condition; MD = Mean Difference; SE = Standard error; t = t-statistic for each comparison; Cohen's d = Cohen's d effect size; p-value (adj) = Bonferroni-adjusted p-value for each comparison * $=0.05$, ** $=0.01$, *** $=<0.001$.

	Comparison	MD	SE	t	Cohen's d	P-value (adj)
STAI-S	B - EC	11.150	1.528	7.29	2.117	< 0.001***
	B - CC	14.500	1.528	9.49	9.492	<0.001***
	EC- CC	3.350	1.528	2.19	2.193	0.10
HR	B - EC	2.813	1.346	2.09	2.090	0.13
	B - CC	4.313	1.346	3.20	0.450	0.01**
	EC- CC	1.500	1.346	1.11	0.157	0.82
Mean RR	B - EC	-37.123	15.018	-2.47	-2.472	0.056
	B - CC	-54.092	15.018	-3.60	-3.602	0.003**
	EC - CC	-16.969	15.018	-1.13	-1.130	0.799
SNS Index	B - EC	0.278	0.150	1.85	0.241	0.21
	B - CC	0.559	0.150	3.73	0.484	0.002**
	EC- CC	0.281	0.150	0.24	0.243	0.20

4. Discussion

This pilot study investigated the effects of HVVEs on anxiety levels and autonomic activation. Our results showed that simulated hallucinatory experiences had a similar calming effect as relaxing videos, especially when considering self-report measures. This aligns with psychedelics research, suggesting that a course of psychedelic drugs can be as effective as existing treatments in reducing anxiety symptoms [2]. The underlying mechanisms seem to be related to changes at a cortical level: specifically, psychedelics decrease activity in the Default Mode Network, a brain network associated with self-referential thinking, rumination, and worry. Thus, by quieting this network, psychedelics may dissolve mental barriers and promote a sense of calm and openness [18, 19]. It might be that artificial hallucinatory experiences similarly affect brain functioning, even though more research is needed to corroborate this hypothesis. Another possibility might also be that IVR technology has a relaxing effect per se. Indeed, IVR induces a sense of presence in the virtual environment able to switch attentional focus, potentially disrupting patterns of anxious and ruminative cognition [20]. If this is the case, a hallucinatory experience might contrast with such a calming effect, leading to an intermediate level of relaxation as compared to the baseline and the control conditions. This perspective aligns with the linear trend observed in our results, where the hallucinated condition showed a medium level between baseline and control conditions. Future research might better isolate the effect of IVR, simulated hallucination, and relaxing experience by manipulating video content and technology. Alternatively, it might be that hallucinatory-like experiences induce eustress, namely a positive stress response, like what was observed in psychedelic research [19]. This would account for the contrast between physiological arousal and self-report anxiety. Our study has several limitations. This is a pilot study, so a larger sample size is needed to further validate our findings. Additionally, our experience lasted 10 minutes, and more time might be needed to observe noticeable effects. Third, we consider as baseline the assessment at the

beginning of the experiment, but future research might consider the possibility of having a different baseline for each condition to better characterize changes. Lastly, we used only HRV indexes as autonomic measures, but other indexes such as ECG and EEG might be used to provide a deeper understanding of the cardiovascular responses to HVVEs as well as possible changes in brain wave patterns.

5. Conclusion

This study aimed to investigate the effects of HVVEs on self-reported anxiety and autonomic activity. Our findings provide preliminary evidence for the anxiolytic potential of HVVEs. Even though further research is needed to corroborate results, this might open new avenues for the management and treatment of anxiety-related disorders, providing the opportunity to exploit psychedelic advantages while avoiding their negative effects.

References

- [1] Yang X, Fang Y, Chen H, Zhang T, Yin X, Man J, et al. Global, regional and national burden of anxiety disorders from 1990 to 2019: results from the Global Burden of Disease Study 2019. *Epidemiol Psychiatr Sci.* 2021;30:e36. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8157816/>.
- [2] Feulner L, Sermchaiwong T, Rodland N, Galarneau D. Efficacy and safety of psychedelics in treating anxiety disorders. Ochsner J. 2023;23:315-28. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10741816/>.
- [3] Hartogsohn I. Cyberdelics in context: on the prospects and challenges of mind-manifesting technologies. *Front Psychol.* 2023;13:1073235.
- [4] Suzuki K, Roseboom W, Schwartzman DJ, Seth AK. A deep-dream virtual reality platform for studying altered perceptual phenomenology. *Sci Rep.* 2017;7:15982.
- [5] Rastelli C, Greco A, Kenett YN, Finocchiaro C, De Pisapia N. Simulated visual hallucinations in virtual reality enhance cognitive flexibility. *Sci Rep.* 2022;12:4027.
- [6] Casu M, Farrauto C, Farruggio G, Bellissima S, Battiato S, Caponnetto P. Exploring the therapeutic potential of virtual reality: a review on the simulation of psychedelic effects for treating psychological disorders. *Psychol Int.* 2024;6:603-17. Available from: <https://www.mdpi.com/2813-9844/6/2/36>.
- [7] Pedrabissi L, Santinello M. Verifica della validità dello STAI forma Y di Spielberger [Verification of the validity of the STAI, form Y, by Spielberger]. *Giunti Organ Spec.* 1989;191-2:11-4.
- [8] Spielberger CD. State-Trait Anxiety Inventory for Adults. APA PsycTests. 1983.
- [9] Tarvainen MP, Niskanen J, Lipponen JA, Ranta-Aho PO, Karjalainen PA. Kubios HRV—heart rate variability analysis software. *Comput Methods Programs Biomed.* 2014;113:210-20.
- [10] Riva G, Bernardelli L, Browning MHEM, Castelnovo G, Cavedoni S, Chirico A, et al. COVID Feel Good—an easy self-help virtual reality protocol to overcome the psychological burden of coronavirus. *Front Psychiatry.* 2020;11:563319.
- [11] Laborde S, Mosley E, Thayer JF. Heart rate variability and cardiac vagal tone in psychophysiological research—recommendations for experiment planning, data analysis, and data reporting. *Front Psychol.* 2017;8:213.
- [12] Tabachnick BG, Fidell LS. Using multivariate statistics. 3rd ed. Boston: Pearson; 2007.
- [13] Blanca MJ, Alarcón R, Arnau J, Bono R, Bendayan R. Effect of variance ratio on ANOVA robustness: might 1.5 be the limit? *Behav Res Methods.* 2018;50:937-62.
- [14] Thayer JF, Ahs F, Fredrikson M, Sollers JJ, Wager TD. A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. *Neurosci Biobehav Rev.* 2012;36:747-56.
- [15] Hölzel BK, Carmody J, Vangel M, Congleton C, Yerramsetti SM, Gard T, Lazar SW. Mindfulness practice leads to increases in regional brain gray matter density. *Psychiatry Res.* 2011;191:36-43. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3004979/>.
- [16] Carhart-Harris RL, Friston KJ. REBUS and the anarchic brain: toward a unified model of the brain action of psychedelics. *Pharmacol Rev.* 2019;71:316-44.
- [17] Carhart-Harris RL, Erritzoe D, Williams T, Stone JM, Reed LJ, Colasanti A, et al. Neural correlates of the psychedelic state as determined by fMRI studies with psilocybin. *Proc Natl Acad Sci U S A.* 2012;109:2138-43.
- [18] Brewer JA, Worhunsky PD, Gray JR, Tang Y, Weber J, Kober H. Meditation experience is associated with differences in default mode network activity and connectivity. *Proc Natl Acad Sci U S A.* 2011;108:20254-9.
- [19] Millière R, Carhart-Harris RL, Roseman L, Trautwein FM, Berkovich-Ohana A. Psychedelics, Meditation, and Self-Consciousness. *Front Psychol.* 2018 Sep 4;9:1475. doi: <https://doi.org/10.3389/fpsyg.2018.01475>. PMID: 30245648; PMCID: PMC6137697.
- [20] Riva G, Mantovani F, Capideville CS, Preziosa A, Morganti F, Villani D, Gaggioli A, Botella C, Alcañiz M. Affective interactions using virtual reality: the link between presence and emotions. *Cyberpsychol Behav.* 2007 Feb;10(1):45-56. doi: <https://doi.org/10.1089/cpb.2006.9993>. PMID: 17305448.

CALMA Online Psychological First Aid Program in Peru: Personal and Sociodemographic Characteristics and Challenges

Karen PÉREZ-MARAVÍ^a, Giovanna OSORIO^a and Iván MONTES-ITURRIZAGA^{a,1}

^a*School of Psychology, Universidad Continental, Huancayo, Perú*

ORCiD ID: Karen Pérez-Maraví <https://orcid.org/0000-0003-0068-4708>

Abstract. The COVID-19 pandemic in Peru led to the closure of almost all mental health services. In addition, the authorities forced citizens to stay at home for prolonged periods. Universidad Continental designed and implemented the CALMA online psychological first aid program at the beginning of the COVID-19 pandemic, conceived within the World Health Organization's technical guide for psychological first aid (2012). This initiative has offered free telecare through digital media to more than 10,000 users under a diverse, inclusive, and multicultural approach to mental health services. This study characterizes the population served ($N=10096$) over these 4 years. Of this group, 74.6% were women ($n=7567$), and 25.4% were men ($n=2527$). Regarding the most represented nationalities, 9137 were Peruvians (90.5%), 318 Mexicans (3.1%), 132 Ecuadorians (1.3%), 109 Colombians (1.1%) and 97 Bolivians (1.1%). The 3 main reasons for seeking care were similar among men and women: lack of apathy and initiative, emotional pain and sadness, and anger, irritability, and fear (30.8% and 30.7%).

Keywords. Psychological first aid, telepsychology, COVID-19

1. Introduction

The context of the COVID-19 pandemic generated the response of multiple organizations to offer emotional support and medical care and continue educational activities at all levels [1]. In this context, specifically in psychology, protocols were developed for using digital platforms for tele-consultations, evaluations, and crisis care.

The context of the COVID-19 pandemic generated the response of multiple organizations to offer emotional support and medical care and continue educational activities at all levels [1]. In this context, specifically in psychology, protocols were developed for using digital platforms for tele-consultations, evaluations, and crisis care [1–4]. Thus, and in this last deployment, listening lines or channels emerged internationally - and to a lesser extent - to offer support to people suffering the consequences of social isolation, the loss of a family member, or the after-effects of different illnesses. Similarly, psychological care, usually provided in private clinics, began to be transferred to digital environments. In summary, the COVID-19 pandemic signified a dizzying worldwide change in how we understand many facets of our lives through digital platforms, the use of cell phones, and messaging systems (text and voice) available on social networks [5,6].

In Peru, the pandemic above had a very negative impact on the lives of its citizens. We were considered the country with the highest fatality rate worldwide. This caused great suffering in the population, mental health problems (anxiety, fear of contagion, and depression, among others), and a deterioration in living conditions, as confinement and unemployment increased poverty considerably [6–9]. In response to this, the Universidad Continental (Peru) carried out a university social responsibility initiative aimed at providing online psychological first aid through a program called CALMA [10].

It is essential to mention that this program (in force until today), initially thought to attend Peruvian citizens, was naturally extended to the international context. A

¹ Corresponding Author: Iván Montes-Iturrizaga, imontes@uc.cl

tremendous national and global demand was generated for this completely free service with the participation, at the beginning, of 160 volunteers (internal psychologists and professors of our school of psychology); at present, 39 volunteers. The training of these volunteers covered topics such as telepsychology, online psychological first aid, ABCDE model of psychological first aid, gender and health course of the Pan American Health Organization, mental health management in the context of COVID-19, training in suicide and crisis care, among others.

CALMA was created from a primary healthcare perspective to intervene in this crisis and national emergency. In coherence with this, CALMA was supported from the beginning by the psychological first aid guide proposed by the World Health Organization in 2012 [4] and by the scope of the American Psychological Association (APA) in terms of tele-psychological care (non-face-to-face) [11]. The ABCDE model developed by Paula Cortez and Rodrigo Figueroa of the Pontificia Universidad Católica de Chile was also used as an essential reference [10].

It is important to note that the Ministry of Health of Peru issued a Technical Guide at the beginning of the COVID-19 pandemic and established a resolution for mental health care in the context of COVID-19, which pointed out the importance of promoting mental health care through technological resources in various groups in situations of greater vulnerability. Likewise, 166 days after the beginning of the pandemic, this same ministry published a resolution that approved the Technical Guide for Psychological First Aid. CALMA began 15 days after the health emergency was declared.

Thus, this article analyzes the personal and sociodemographic characteristics and reasons for appointments (with emphasis on pre- and post-pandemic times) in the 10096 people who attended this service.

2. Methods

2.1. Participants

Individuals who attended the Online Psychological First Aid Program (CALMA) of the Universidad Continental (Peru) completed a form (Google Forms) to schedule an appointment ($N=10096$). These participants were 2567 men (25.4%) and 7567 (74.6%). In this report, we consider the entire population served by CALMA since its inception on March 30, 2020, and up to the present (April 17, 2024).

2.2. Instrument

Under informed consent, participants had to complete an anonymous form (Google Forms) to schedule an appointment, which was available in Spanish, Quechua (native language), and English. This form includes the following questions: gender, age, marital status, religion, education, city and country of residence, the reason for the appointment, how they learned about CALMA, and if they have any neurodiversity condition or disability. In any case, the instrument likely did not realistically capture the reasons for consultation, given that the specific question had the following problems: use of technical-academic language, many alternatives or response options, and alternatives of diverse nature. Therefore, given the difficulty of handling this complex information, we had to reduce the response options by grouping them according to similarity.

2.3. Statistical Analyses

Statistical analyses were performed in SPSS version 28 for Windows and JAMOVI version 2.3.28. Analyses are reported considering pandemic vs. post-pandemic time and the sex variable.

3. Results

Figure 1 shows the evaluation of the CALMA Program's performance since its inception and to date. Thus, in 2020 and 2021, the most significant number of users attended (and

in specific months), and just at the time of maximum health tension and economic recession that forced millions of Peruvians and Latin Americans to stay off the streets.

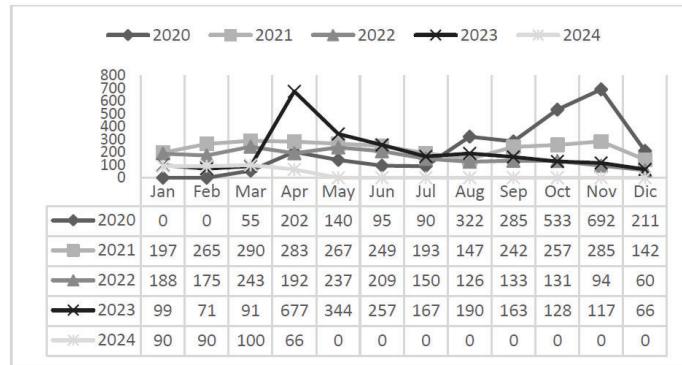


Figure 1. CALMA Program services since its creation.

Table 1 shows the composition of the population served, with a predominance of women and those with higher education. Likewise, there is a predominance of singles and people who profess the Catholic religion. It is relevant to mention, in terms of religion, that a high number of those who marked the "other" option are people who profess the so-called Christian cults that makeup Peru's varied religious spectrum.

Table 1. Socio-demographic distribution of the population served by CALMA.

Variable	N	%
Sex		
<i>Male</i>	2567	25.4
<i>Female</i>	7529	76.6
Education		
<i>Primary</i>	243	2.4
<i>Secondary</i>	2875	28.5
<i>Technical</i>	1907	18.9
<i>Incomplete university</i>	2486	24.6
<i>Universitaria complete</i>	2585	25.6
Marital status		
<i>Singel</i>	8171	81.5
<i>Cohabitan - Married</i>	1601	16.0
<i>Separated - Divorce</i>	219	2.2
<i>Widowed</i>	36	0.4
Religion		
<i>Catholic</i>	5764	57.1
<i>Evangelical</i>	1362	13.5
<i>Jewish</i>	8	0.1
<i>Muslim</i>	7	0.1
<i>Buddhist</i>	26	0.3
<i>Other</i>	1802	17.8
<i>Agnostic</i>	977	9.7
<i>Atheist</i>	150	1.5

Table 2 shows that 397 beneficiaries have some disability and neurodiversity, the most frequent being psychosocial, visual, and motor.

Table 2. Population distribution by disability and neurodiversity.

	n	%
<i>Other</i>	47	11.8
<i>Autism</i>	2	0.2
Hearing Impairment	11	2.8
Speech disability	3	0.8
Intellectual disability	8	2.0
Motor disability	57	14.4
Multiple disabilities	2	0.5
Organic or visceral disability	2	0.5
Psychosocial disability	227	57.2

Visual disability	35	8.8
Dyslexia	3	0.8
TOTAL	397	100

Table 3 presents the reasons for requesting this service during the pandemic by COVID-19 and post-pandemic time. The most considered reasons were the same in these two pandemic times; however, in the pandemic, anger, irritability, and fear scored higher, and, in the post-pandemic, lack of initiative and apathy were the most frequent. These reasons are repeated when we analyze the reasons according to men and women: lack of apathy and initiative (20.1% and 21.0%); emotional pain and sadness (25.0% and 20.7%); and anger, irritability, and fear (30.8% and 30.7%).

Table 3. Distribution of reasons for care by the time of pandemic.

Pandemic momentum	n	%
Pandemic		
Lack of initiative and apathy	1218	17.8
Emotional pain and sadness	1662	24.3
Lack of adaptation to work or school	77	1.1
Other reasons	833	12.2
Interpersonal conflicts	44	0.6
Physical symptoms	149	2.2
Anger, irritability, and fear	2429	35.5
Concern for survival	292	4.3
Feelings of abandonment and rejection	26	0.4
Feelings of guilt	106	1.6
TOTAL	6836	100
Post pandemia		
Lack of initiative and apathy	702	25.8
Emotional pain and sadness	623	22.9
Lack of adaptation to work or school	84	3.1
Other reasons	233	8.6
Interpersonal conflicts	52	1.9
Physical symptoms	130	4.8
Anger, irritability, and fear	552	20.3
Concern for survival	151	5.6
Feelings of abandonment and rejection	43	1.6
Feelings of guilt	146	5.4
TOTAL	2716	100

1. Discussion

The composition of the population served by CALMA since the beginning of the pandemic has unique characteristics given that the online component expects some communication device (laptop, Tablet, or cell phone) from the beneficiaries, a fact that possibly made it possible to serve mostly vulnerable populations, but from urban areas and with better services. Likewise, it reveals that the reasons for care are the same according to gender or time of the pandemic, although with different intensities.

References

- [1] Sandoval-Obando E, Sandoval-Díaz J. Psicología de la emergencia en contexto de pandemia. Tesis Psicol. 2020;15(2):254–279. doi: <https://doi.org/10.37511/tesis.v15n2a14>.
- [2] Rojas-Jara C, Polanco-Carrasco R, Caycho-Rodríguez T, Muñoz-Vega C, Muñoz-Marabolí M, Luma-Gómez T, et al. Telepsychology for psychotherapists: lessons learned in times of Covid-19. Interam J Psychol. 2022;56(2):e1733. doi: <https://doi.org/10.30849/ripijp.v56i2.1733>.
- [3] Feijt M, de Kort Y, Bongers I, Bierbooms J, Westerink J, IJsselsteijn J. Mental health care goes online: mental health during COVID-19. Cyberpsychol Behav Soc Netw. 2020;23(12):860–864. doi: <https://doi.org/10.1089/cyber.2020.0370>.
- [4] World Health Organization, War Trauma Foundation, World Vision International. Psychological first aid: Guide for field workers. Geneva: WHO [Internet]. 2011. Available from: <https://www.who.int/publications/i/item/9789241548205>.
- [5] Argüero-Fonseca A, Aguirre-Ojeda DP, Reynoso-González OU, Girón-Pérez MI, Espinosa-Parra IM, Sierra-Murguía MA. Impacto de la telepsicología en la satisfacción de la atención a pacientes con Covid-19. Psicol Iberoam. 2021;29(3):e293325. doi: <https://doi.org/10.48102/pi.v29i3.325>.

- [6] Moreno-Méndez JH. Los beneficios de la telepsicología para la relación docencia-servicio en el contexto de pandemia. *Univ Psychol.* 2022;21:1–18. doi: <https://doi.org/10.11144/Javeriana.upsy21.btrd>.
- [7] Escobar-Arango F, Florez-Torres AM, Estrada-Jiménez MJ, Agudelo-Valencia A. Panorama de las intervenciones psicológicas virtuales en Hispanoamérica y España antes de la pandemia COVID-19: revisión documental. *Rev Virtual Univ Catol Norte.* 2022;19(67):217–248. doi: <https://doi.org/10.35575/rvucn.n67a9>.
- [8] World Health Organization. Mental health and COVID-19: Scientific brief. Geneva: WHO [Internet]. 2022. Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-Sci_Brief_Mental_health-2022_1.
- [9] Chuquimarca-González ME, Torres-Quinchanegua V. Los primeros auxilios psicológicos en el manejo de ataques de ansiedad en adultos jóvenes de edades entre 20 y 25 años, debido a la crisis sanitaria por el COVID-19 entre los meses de julio de 2021 al enero del 2022 [undergraduate thesis]. Quito: Universidad Politécnica Salesiana; 2022. Available from: <http://dspace.ups.edu.ec/handle/123456789/22057>.
- [10] Ramos del Río B, Martí-Noguera JJ, coordinators. Experiencias en ciberpsicología hacia una nueva era de la psicología [Internet]. Ciudad de México: Universidad Nacional Autónoma de México; 2022. Available from: <https://www.zaragoza.unam.mx/experiencias-ciberpsicologia/>
- [11] Pierce BS, Perrin PB, Tyler CM, McKee GB, Watson JD. The COVID-19 telepsychology revolution: a national study of pandemic-based changes in U.S. mental health care delivery. *Am Psychol.* 2021;76(1):14–25. doi: <https://doi.org/10.1037/amp0000722>

Dynamics of Online Relationships Through Virtual Reality Experiences

Aimée ARGÜERO-FONSECA^{a,1} and Davide MARCHIORO^{b,a}

^a*Autonomous University of Nayarit*

^b*University Institute of Saletian Studies of Venice*

ORCID ID: Aimée Argüero-Fonseca <https://orcid.org/0000-0002-3864-5299>

Abstract. This exploratory study investigates the dynamics of online relationships within virtual reality (VR) experiences to understand the intersection of human connections and technology. The primary objective is to examine how VR influences the initiation, development, and sustainability of online relationships. Employing a mixed-methods approach, the study combines qualitative analysis of user experiences in VR environments with quantitative measures of relationship satisfaction, authenticity, and psychological well-being. Participants were recruited via social media. Inclusion criteria specified individuals aged 18 to 30 who had met their partners online, been in a relationship for less than three months, and were willing to acquire the VR device. A total of 24 participants met these criteria and engaged in simulated social interactions within VR settings. Qualitative data, including participant narratives, underwent thematic analysis to identify recurrent themes and patterns. Concurrently, quantitative data were gathered through surveys evaluating participants' perceived authenticity of interactions, emotional connection, and overall satisfaction with VR-facilitated online relationships. The qualitative analysis revealed rich insights into the immersive nature of VR experiences, highlighting elements that significantly contribute to meaningful connections. Participants reported heightened feelings of authenticity and emotional connection within the virtual realm. Quantitative findings indicated a positive correlation between the depth of engagement in VR-mediated interactions and participants' satisfaction with the authenticity of online relationships. The study concludes that VR experiences profoundly influence the dynamics of online relationships, providing a unique and enriched platform for social interactions. The heightened sense of authenticity and emotional connection reported by participants underscores VR's potential to enhance the quality of online relationships. These findings contribute to a broader understanding of the synergies between online connections and VR experiences, offering implications for researchers and practitioners navigating the evolving landscape of digital relationships.

Keywords. Virtual Reality, online relationships, digital interactions

1. Introduction

The advent of digital technology has profoundly transformed interpersonal relationships, with virtual reality (VR) emerging as a groundbreaking platform for social interaction. VR technology offers immersive environments that simulate real-world experiences, enabling users to engage in lifelike interactions despite physical distances. Unlike traditional online communication platforms, VR provides a sense of presence and embodiment, making interactions more engaging and realistic. As VR becomes more integrated into daily life, understanding its impact on relationships is crucial. This research explores the intersection of human connections and VR technology, aiming to understand how VR influences the initiation, development, and sustainability of online relationships.

1.1. Brief Literature Review

The exploration of online relationships through VR experiences reveals mixed findings regarding authenticity and emotional connection compared to in-person interactions. Dailey-Hebert notes that VR can enhance informal connections and engagement among online learners, indicating a positive impact on satisfaction and motivation by providing a sense of social presence [1]. Roth et al. found that avatars can evoke a similar level of

¹ Corresponding Author: Aimée Argüero Fonseca, NayaAimee.arguero@uan.edu.mz

empathy and perceived authenticity as human interactions, suggesting that avatars can effectively mediate self-disclosure and foster pro-social intentions [2]. Matsui highlights that co-presence and visual elements in 3D virtual learning environments can enhance emotional connections, although the impact varies based on language proficiency and the topics discussed [3]. Herrschaft discusses the challenges and opportunities in creating meaningful interpersonal connections in virtual spaces, underscoring the importance of designing welcoming and inclusive virtual environments [4]. Overall, these studies suggest that while VR can facilitate authentic and emotionally rich interactions, the effectiveness varies based on design, context, and user perceptions.

VR has the potential to support a wide variety of interactions, including romantic, platonic, and familial relationships. Ferdig et al. explore the recursive relationship between virtual and real interactions, emphasizing that many people build and sustain both romantic and platonic relationships through technology, including virtual worlds [5]. Freeman and Acena focus on social VR systems, which provide high-fidelity, immersive experiences that can enhance the way people meet, interact, and build interpersonal relationships, highlighting activities that foster these relationships in social VR [6]. These studies collectively show that VR can effectively support various types of relationships by offering immersive, interactive, and supportive environments for users.

VR can enhance the quality of romantic relationships by providing immersive, interactive, and emotionally engaging experiences that can strengthen bonds and create a sense of presence. Gordon discusses how VR-mediated love can withstand philosophical critiques, suggesting that VR can offer significant value in maintaining romantic relationships by creating a shared space for interaction [7]. Koike et al. introduce the concept of romantic anthropomorphism, where non-human agents in romantic contexts can provide authentic feelings and positive affect, potentially fulfilling emotional needs in virtual romantic relationships [8]. Gilbert et al. found that communication in VR environments like Second Life can be perceived as higher quality compared to real-life interactions, leading to greater satisfaction with virtual partners [9]. Huang and Bailenson explore the intersection of close relationships and VR, emphasizing the potential of VR to facilitate attachment processes and enhance relationship formation and maintenance [10]. These studies collectively highlight that VR can offer unique opportunities to enhance romantic relationships by providing novel ways to connect, communicate, and share experiences.

1.2. Identification of the Problem

While VR offers several benefits for maintaining long-distance relationships, it also presents limitations that can impact its effectiveness. Key challenges include technological, psychological, and practical aspects. Technological limitations involve hardware accessibility, cost, technical difficulties, and limited content availability [11][12]. Psychological limitations include physical discomfort, health concerns, and emotional disconnect due to the lack of physical touch and subtle non-verbal cues [13][14]. Practical limitations encompass time and scheduling conflicts, as well as adaptation and learning curves associated with using VR technology [11].

1.3. Objectives and Purpose

The primary objective of this research is to investigate how VR influences the initiation, development, and sustainability of online relationships. Specifically, the study aims to assess relationship satisfaction, perceived authenticity, and psychological well-being among individuals using VR. By employing a mixed-methods approach, this research seeks to achieve a nuanced understanding of the dynamics at play within VR environments across different types of relationships.

1.4. Justification and Relevance

This research is necessary to advance the understanding of VR's role in modern social interactions. By examining how VR affects relationship dynamics in romantic, platonic, and familial contexts, the study contributes to both theoretical and practical knowledge in the field. The findings will inform the development of VR platforms, guiding ethical

considerations and enhancing their effectiveness in fostering authentic connections. Understanding these dynamics is crucial for developing technologies that support healthy and meaningful online interactions.

2. Methods

The study employs a mixed-methods approach, integrating qualitative analysis of user experiences in VR environments with quantitative measures of relationship satisfaction, authenticity, and psychological well-being. Participants were recruited via social media through a call for participation at the “PsicoCiencias” Laboratory of the Autonomous University of Nayarit. The inclusion criteria specified individuals aged 18 to 30 who had met their partners online, been in a relationship for less than three months, and were willing to acquire the VR device. A total of 24 participants met these criteria and engaged in simulated social interactions within VR settings, providing valuable insights into the nuances of relationship formation and maintenance.

To analyze the results, a mixed-methods approach was employed. This included thematic analysis of qualitative data, where participant narratives were coded and grouped into themes such as immersive experiences, authenticity of interactions, and emotional connections. Concurrently, quantitative data from surveys were analyzed using descriptive statistics to calculate mean scores and standard deviations for perceived authenticity, emotional connection, overall satisfaction, and engagement depth. Pearson correlation coefficients were computed to examine the relationships between engagement depth and satisfaction metrics. Integration of qualitative themes with quantitative scores through triangulation ensured comprehensive understanding. NVivo for qualitative analysis and SPSS for quantitative analysis were utilized to facilitate this process.

3. Results

The qualitative analysis revealed rich insights into the immersive nature of VR experiences, highlighting elements that significantly contribute to meaningful connections. Participants reported heightened feelings of authenticity and emotional connection within the virtual realm. Quantitative findings indicated a positive correlation between the depth of engagement in VR-mediated interactions and participants' satisfaction with the authenticity of online relationships.

Table 1. Results of Thematic and Quantitative Analysis of VR-Facilitated Online Relationships

Theme	Participant Narratives (Qualitative)	Survey (Quantitative)	Metrics	Correlation (r-value)
Immersive Nature of VR Experiences	- "I felt like I was truly there, it was incredibly lifelike." - "The surroundings and interactions were so realistic, I often forgot I was in a virtual environment." "VR made me feel fully present and engaged."	-	-	-
Authenticity of Interactions	- "Conversations felt more genuine and personal." "I could see and feel the authenticity in the interactions, unlike text-based chats." "The body language and expressions added a layer of authenticity."	Perceived Score: 4.7/5	Authenticity	r = 0.76
Emotional Connection	"I could feel the emotions of the other person, it was very real." - "The emotional bond I formed was unlike any other online experience." - "VR allowed us to connect on a deeper emotional level."	Emotional Connection Score: 4.5/5		r = 0.82
Overall Satisfaction	- "I am very satisfied with my relationships in VR, it feels like real life." - "My overall satisfaction with VR interactions is extremely high." - "These relationships are as fulfilling as physical ones."	Overall Satisfaction Score: 4.8/5		r = 0.79

Depth Engagement	of	- "The more I engaged, the more real it felt, enhancing my satisfaction." - "Deep engagement in VR was key to forming real connections." - "In-depth interactions made a significant difference in my experience."	Engagement Depth Score: 4.6/5	r = 0.81
Meaningful Connections		- "VR has allowed me to form deeper, more meaningful relationships." - "The connections I've made in VR are some of the most meaningful in my life." - "Meaningful relationships were easier to build in VR."	Meaningful Connection Score: 4.9/5	r = 0.84

Notes Perceived Authenticity Score: Average rating given by participants on the authenticity of interactions in VR; Emotional Connection Score: Average rating given by participants on the emotional connection experienced in VR; Overall Satisfaction Score: Average rating given by participants on their overall satisfaction with VR-facilitated relationships; Engagement Depth Score: Average rating given by participants on the depth of their engagement in VR interactions; Meaningful Connection Score: Average rating given by participants on the meaningfulness of connections formed in VR; Correlation (r-value): Represents the strength of the relationship between depth of engagement and satisfaction/authenticity of online relationships.

4. Discussion

The findings from this study provide compelling evidence that VR experiences can profoundly influence the dynamics of online relationships. The heightened sense of authenticity and emotional connection reported by participants underscores VR's potential to enhance the quality of online interactions. Participants indicated that the immersive nature of VR allowed for deeper engagement and more genuine connections, which were reflected in both qualitative narratives and quantitative scores. The strong correlations between the depth of engagement and satisfaction with authenticity and emotional connection highlight the importance of immersive and interactive elements in fostering meaningful relationships in VR. These results align with previous research suggesting that VR can enhance informal connections [1]. However, this study also brings to light several challenges associated with VR-mediated relationships. Technological limitations such as hardware accessibility, cost, and technical difficulties remain significant barriers to widespread adoption. Psychological issues, including physical discomfort and the potential for emotional disconnect due to the lack of physical touch and subtle non-verbal cues, also present obstacles. Practical challenges such as time and scheduling conflicts, along with the steep learning curve for new users, further complicate the effective use of VR for maintaining relationships. Addressing these challenges requires ongoing technological advancements, user education, and the development of more tailored VR content that can better support relationship maintenance. Ethical considerations and privacy safeguards must also be prioritized to ensure safe and positive VR experiences.

5. Conclusion

This exploratory study concludes that VR experiences significantly impact the dynamics of online relationships, offering a unique and enriched platform for social interactions. The heightened sense of authenticity and emotional connection reported by participants suggests that VR has the potential to enhance the quality of online relationships. These findings contribute to a broader understanding of the synergies between online connections and VR experiences, offering valuable insights for researchers and practitioners navigating the evolving landscape of digital relationships. Future research should continue to explore the ethical and practical implications of VR-mediated interactions, focusing on developing solutions to the challenges identified in this study. By addressing these issues, VR technology can become an even more effective tool for fostering meaningful and authentic online relationships.

References

- [1] Dailey-Hebert A. Student perspectives on using virtual reality to create informal connection and engagement. InSight: A Journal of Scholarly Teaching. 2022;17:28-46. doi: <https://doi.org/10.46504/17202202da>.
- [2] Roth D, Bloch C, Schmitt JB, Frischlich L, Latoschik ME, Bente G. Perceived authenticity, empathy, and pro-social intentions evoked through avatar-mediated self-disclosures. 2019;21-30. doi: <https://doi.org/10.1145/3340764.3340797>.
- [3] Matsui H. The Impact of Co-Presence and Visual Elements in 3D VLEs on Interpersonal Emotional Connection in Telecollaboration. 2013.
- [4] Herrschaft G. Facilitating Meaningful Interpersonal Connections Through a Virtual Space. 2017.
- [5] Ferdig RE, Pytash KE, Muschert GW. The recursive relationship between virtual and real in relationships. J Virtual Worlds Res. 2017;10(2). doi: <https://doi.org/10.4101/JVWR.V10I2.7286>.
- [6] Freeman G, Acena D. Hugging from a distance: building interpersonal relationships in social virtual reality. 2021;84-95. doi: <https://doi.org/10.1145/3452918.3458805>.
- [7] Gordon EC. Virtual reality and technologically mediated love. Kennedy Inst Ethics J. 2022;32(4):329-57. doi: <https://doi.org/10.1353/ken.2022.0027>.
- [8] Koike M, Loughman S, Stanton SCE. Virtually in love: the role of anthropomorphism in virtual romantic relationships. Br J Soc Psychol. 2022;62(1):600-16. doi: <https://doi.org/10.1111/bjso.12564>.
- [9] Gilbert RL, Murphy NA, Ávalos MC. Communication patterns and satisfaction levels in three-dimensional versus real-life intimate relationships. Cyberpsychol Behav Soc Netw. 2011;14(10):585-9. doi: <https://doi.org/10.1089/CYBER.2010.0468>.
- [10] Huang SA, Bailenson JN. Close relationships and virtual reality. 2018;49-65. doi: https://doi.org/10.1007/978-3-030-02631-8_4.
- [11] Han E, Nowak KL, Bailenson JN. Prerequisites for learning in networked immersive virtual reality. Technol Mind Behav. 2022;3(4). doi: <https://doi.org/10.1037/tmb0000094>.
- [12] Rukangu A, Franzluebbers A, Tuttle A, Mattingly K, Walker O'Neal C, Robinson DT, Ahn SJG, Johnsen K. Virtual family room: bridging the long distance. 2020;52-6. doi: <https://doi.org/10.1109/VRW50115.2020.00017>.
- [13] Елесин СС, Фещенко АВ. Виртуальная реальность в образовании: сомнения и надежды. 2015.
- [14] Zamanifard S, Freeman G. "The togetherness that we crave": experiencing social VR in long-distance relationships. 2019;438-42. doi: <https://doi.org/10.1145/3311957.3359453>.

Using Eye Gaze to Differentiate Internal Feelings of Familiarity in Virtual Reality Environments: Challenges and Opportunities

Trevor CHARTIER^{a,1}, Iliana CASTILLON^a, Videep VENKATESHA^a, Anne M. CLEARY^{a,2}, and Nathaniel T. BLANCHARD^a

^a*Colorado State University, Fort Collins, CO, 80523, USA*

ORCID ID: Trevor Chartier <https://orcid.org/0009-0003-6751-1905>

Abstract. Our group previously reported a feasible approach to detect the internal state of familiarity with eye-gaze features [1]. Utilizing an existing paradigm [2], we examined participants' feelings of familiarity during immersion within virtual reality (VR) scenes, some of which had had their spatial layout familiarized through prior presentation of a different scene with the same configuration. While immersed in a test scene, participants indicated the onset of familiarity via a button press on a handheld controller, then verbally indicated whether they could state the source of the familiarity or not. A potential issue is that machine learning models may have detected eye-gaze features reflecting the act of pressing the button rather than features associated with the internal state of familiarity. Although in [1] we addressed this challenge by including a buffer period between the button press and the window of data used for model training, it remains uncertain within what time frame features associated with the button press may persist. Here, we introduce an approach for potentially overcoming the confounding effects of the button-press by holding it constant. We examine machine learning models' ability to detect whether a scene's layout had been experimentally familiarized among only instances where subjective familiarity was reported. We then repeat this method for instances where no familiarity was reported. Finally, we examine experimentally familiarized scenes where familiarity was reported to detect recall-success vs. recall-failure for the familiarity's source.

Keywords. Machine learning, internal state detection, familiarity, virtual reality, intelligent virtual tutoring systems

1. Introduction

Research on internal state detection centers on determining the reliability of machine learning models at identifying whether an individual is in a particular cognitive state based on the external physiological qualities they exhibit. Much of this research examines the automatic detection of mind wandering, an internal state often characterized by task-unrelated-thought. These methods have been mostly successful using data from various physiological features, including eye-gaze, skin conductance, and heart rate, with eye-gaze features being the most effective [3,4].

Some research has examined the detection of another internal state—the subjective sense of familiarity—from eye-gaze features [1]. Familiarity is a common subjective experience that is tied closely to recognition. It is thought to trigger a recall search in memory where individuals will attempt to identify the source of familiarity [5]. Our group previously employed popular machine learning algorithms to detect when participants were consciously experiencing familiarity within virtual reality (VR) scenes and found above-chance accuracy results [1]. The current work aims to expand upon our prior findings, including addressing some methodological challenges we previously encountered.

¹ Corresponding Author: Trevor Chartier, trevor.chartier@colostate.edu

² Corresponding Author: Anne M. Cleary, anne.cleary@colostate.edu

Establishing a ground-truth for familiarity experiences is a non-trivial task that requires some form of subjective report from participants. Often used in mind wandering studies [6], probe-based reports prompt participants throughout the experiment to periodically answer whether they are currently experiencing the given cognitive state. However, this method fails to capture the precise moment of onset for a cognitive state. We aimed to capture the moment of onset of familiarity by having participants indicate, via a button-press on a VR hand-controller, the moment something feels familiar in a VR scene. This study design posed unique challenges. Namely, eye-gaze features may correlate with the action of pressing the hand-controller button, thereby introducing a confounding variable to extracting the eye-gaze data that reflect the sensation of familiarity. Training models on this data set initially produced accuracy levels too high to fit in with prior research regarding cognitive state detection, suggesting that the models may have been picking up on differentiable eye-gaze patterns of intent to press the button in the moments preceding a button press. To negate this effect in our previous study, we introduced a 500 millisecond (ms) buffer period between the button press and the two second window of eye-gaze data collected. While this is a viable approach, there is currently little research to inform the correct buffer size to use in VR settings for removing eye-gaze features related to the button-press. As such, it is possible that residual gaze features correlated with the button press persisted beyond the 500ms buffer period and into the window used for model training; it is also possible that valuable eye-gaze data reflective of the state of familiarity were removed from model training.

In the current work we present another approach to dealing with this issue. We hold the VR hand-controller button-press constant while training the model on experimentally familiarized vs. unfamiliarized VR scenes. Based on prior research [2,7,8], a VR scene was considered experimentally familiarized if an otherwise novel scene containing its same spatial layout had been presented earlier in the experiment. First, we conducted a model search using eye-gaze features coming from the moments preceding a familiarity report in a scene to predict the experimentally-familiarized status of the given scene. Next we conduct a model search using eye-gaze features coming from randomly selected moments from within scenes that no familiarity was reported to detect, again, the experimentally-familiarized status of the given scene. Finally, we trained the models on experimentally familiarized test scenes where familiarity was reported via the button-press to compare instances where recall of the corresponding configurally similar study scene succeeded vs failed.

Beyond addressing the button-press challenges inherent in using a button-press to detect internal states via eye-gaze patterns in VR, we aim to show that distinct patterns in eye-gaze features emerge from different levels of familiarity when the subjective state is reported. Familiarity strengths are known to vary, where “a high degree of overlap between the features of the current situation and the features of previous experiences in memory produces a relatively strong familiarity signal” and “a low degree of overlap produces a relatively weak familiarity signal” [9,10]. As such, we might expect higher levels of familiarity with experimentally familiarized scenes than those that have not been experimentally familiarized, even among VR scenes for which the button was pressed to indicate a sense of familiarity.

Machine learning models have been shown to be able to identify previously seen images from unseen images using eye-gaze features, even in cases where participants failed to consciously recognize the previously studied images [11]. However, unlike previous research that examined when a participant viewed an image identical to one previously seen, we aim to detect when participants are immersed in novel scenes containing the same spatial layout as a previously viewed scene. Additionally, the current approach allows us to assess whether machine learning models are able to make this separation not only when participants consciously felt that the scene was familiar, but also when participants felt that it was unfamiliar, allowing for a direct comparison of model performance in each case.

2. Methods

We used the same dataset as [1], in whose study 26 undergraduate students from Colorado State University participated for course credit. As described in [1], the procedure from which the data were obtained closely resembled that of [2].

2.1. Data Collection

Eye-gaze data was measured at 120hz by the HTC Vive Pro Eye headset and collected using the SRanipal software development kit (SDK) version 1.3.6.8. Data was stored in comma-separated-values (CSV) files along with participants' recorded verbal responses and annotations, which included columns for familiarization/study status (whether a configurally identical scene was studied) and recall status (whether the studied scene was recalled or not following a report of familiarity).

2.2. Data Preprocessing and Model Training

We extracted eye-gaze data from 1, 2, and 3 second windows before each familiarity indication. Four buffer periods, ranging from 0 ms to 1000 ms, were included between the familiarity indication and the window of data extracted, consistent with [1]. Trials in which a familiarity indication occurred too soon into scene's onset to extract the given buffer and window size were excluded. For each positive report of familiarity, a negative instance of the same window size is extracted, from the same participant, from a scene in which no familiarity was reported. No buffer period is included for the negative instances, as no button press occurs during these. Eye-gaze features of interest were extracted using PyTrack. The result was 24 distinct datasets of eye-gaze features, with 12 datasets for positive instances and 12 for negative instances, for model training and comparison.

Across all detection tasks, model search was conducted with Hyperopt for hyperparameter optimization using random search over 300 training evaluations. Model algorithms used included AdaBoost, Naive Bayes, Logistic Regression, Support Vector Classifier, Random Forest, and K-Nearest Neighbors. Training and evaluation was done using Leave-One-Participant-Out-Cross-Validation. Cohen's Kappa was used to evaluate the model accuracy; specifically, we report the average across the held participants.

3. Results

During instances of positive familiarity reports, the Ada Boost model algorithm preformed best at differentiating between scenes that were experimentally familiarized and scenes that were not experimentally familiarized. This model resulted in a Cohen's Kappa score of 0.16 from training on 3 seconds of eye-gaze data, temporally separated from the button press by a 500ms buffer. Model results for all buffer and window sizes tested can be found in **Table 1**.

The same detection task, preformed on eye-gaze data extracted from scenes where no familiarity was reported, also found the best performance from the Ada Boost algorithm combined with a 3 second window of eye-gaze data. However, the model achieved a lower Cohen's Kappa score of 0.09. Model results for all window sizes tested can be found at the bottom of **Table 1**.

During experimentally familiarized test scenes where familiarity was reported, the Naive Bayes model was able to detect a participant's recall status with a Cohen's Kappa score of 0.25. This best performing model resulted from training on 3 seconds of eye-gaze data, temporally separated from the button press by a 250ms buffer. Model results for all buffer and window sizes tested can be found in **Table 2**.

4. Discussion

Among instances where participants reported consciously feeling a sense of familiarity with a VR scene (thus the button-press was held constant as it was always pressed in these cases), detectable differences in eye-gaze patterns occurred between experimentally familiarized and unfamiliarized scenes. In contrast, when participants did not consciously experience familiarity with a scene (thus the button-press was held constant in being unpressed), our models were unable to detect any significant differences in the eye-gaze patterns among experimentally familiarized vs. unfamiliarized scenes. These patterns may suggest that eye-gaze patterns reflective of experimental familiarization of scene layout is only detectable when participants are experiencing a sense of familiarity for the scene. If so, an implication might be that a subjectively detectable sense of familiarity must be present in the experiencer in order for variations in familiarity intensity to be externally detectable. However, the lower model performance among negative instances might reflect not having a precise moment (the button-press) to reference. Unlike in positive instances where data corresponds to the time point of the button-press, for negative instances, we sampled a random window of data from within the scene. Future research should continue to try to tackle this challenge. One method might be to yoke each positive instance with a random negative instance from the same participant, using the same time point in the negative instance as when the button was pressed in the positive instance.

Table 1. Detecting the experimentally familiarized status of scenes using eye-gaze features: a comparison of model performance with various buffer and window sizes (Standard Deviation in parenthesis).

Buffer	Window	Model	Cohen's Kappa	F1 Score	Familiarity Reported
0 ms	1 sec	Naive Bayes	0.11 (0.17)	0.49 (0.13)	✓
	2 sec	Ada Boost	0.08 (0.16)	0.61 (0.11)	✓
	3 sec	Ada Boost	0.06 (0.17)	0.60 (0.10)	✓
	1 sec	Random Forest	0.08 (0.18)	0.62 (0.13)	✓
	2 sec	Ada Boost	0.13 (0.18)	0.63 (0.15)	✓
	3 sec	Ada Boost	0.13 (0.21)	0.60 (0.13)	✓
250 ms	1 sec	Ada Boost	0.08 (0.25)	0.62 (0.15)	✓
	2 sec	SVC	0.13 (0.16)	0.63 (0.12)	✓
	3 sec	Ada Boost	0.16 (0.22)	0.64 (0.12)	✓
	1 sec	Ada Boost	0.08 (0.17)	0.62 (0.14)	✓
	2 sec	Random Forest	0.10 (0.18)	0.62 (0.15)	✓
	3 sec	Ada Boost	0.09 (0.18)	0.60 (0.11)	✓
500 ms	1 sec	SVC	0.05 (0.14)	0.56 (0.17)	✗
	2 sec	Ada Boost	0.01 (0.14)	0.56 (0.14)	✗
	3 sec	Ada Boost	0.09 (0.21)	0.63 (0.15)	✗
	1 sec	Ada Boost	0.08 (0.17)	0.62 (0.14)	✓
	2 sec	Random Forest	0.10 (0.18)	0.62 (0.15)	✓
	3 sec	Ada Boost	0.09 (0.18)	0.60 (0.11)	✓
1000 ms	1 sec	SVC	0.05 (0.14)	0.56 (0.17)	✗
	2 sec	Ada Boost	0.01 (0.14)	0.56 (0.14)	✗
	3 sec	Ada Boost	0.09 (0.21)	0.63 (0.15)	✗
	1 sec	Ada Boost	0.08 (0.17)	0.62 (0.14)	✓
	2 sec	Random Forest	0.10 (0.18)	0.62 (0.15)	✓
	3 sec	Ada Boost	0.09 (0.18)	0.60 (0.11)	✓
N/A	1 sec	SVC	0.05 (0.14)	0.56 (0.17)	✗
	2 sec	Ada Boost	0.01 (0.14)	0.56 (0.14)	✗
	3 sec	Ada Boost	0.09 (0.21)	0.63 (0.15)	✗
	1 sec	Ada Boost	0.08 (0.17)	0.62 (0.14)	✓
	2 sec	Random Forest	0.10 (0.18)	0.62 (0.15)	✓
	3 sec	Ada Boost	0.09 (0.18)	0.60 (0.11)	✓

Table 2. Detecting participant's recall status preceding positive reports among experimentally familiarized scenes: a comparison of model performance with various buffer and window sizes (Standard Deviation in parenthesis).

Buffer	Window	Model	Cohen's Kappa	F1 Score
0 ms	1 sec	Ada Boost	0.16 (0.27)	0.64 (0.17)
	2 sec	Ada Boost	0.14 (0.26)	0.53 (0.21)
	3 sec	Naive Bayes	0.23 (0.18)	0.64 (0.12)

	1 sec	Ada Boost	0.14 (0.24)	0.64 (0.13)
250 ms	2 sec	Naive Bayes	0.10 (0.24)	0.54 (0.19)
	3 sec	Naive Bayes	0.25 (0.21)	0.66 (0.12)
500 ms	1 sec	Naive Bayes	0.06 (0.18)	0.45 (0.16)
	2 sec	Ada Boost	0.19 (0.18)	0.66 (0.12)
	3 sec	Naive Bayes	0.23 (0.28)	0.65 (0.14)
1000 ms	1 sec	Ada Boost	0.16 (0.23)	0.65 (0.15)
	2 sec	Naive Bayes	0.18 (0.25)	0.60 (0.15)
	3 sec	Random Forest	0.17 (0.23)	0.66 (0.15)

Finally, among experimentally familiarized scenes for which the button was pressed, we also observed differences in eye-gaze patterns between instances of familiarity accompanied by recall success vs. failure. This suggests that when a VR scene feels familiar, eye-gaze patterns differ in some fundamental way when the source of the familiarity is identified compared to when it is not, or, alternatively, that there is a detectable gaze pattern difference between recognizing a scene based on recall vs. based on familiarity alone.

Methodologically, this work presents a potential means of addressing issues surrounding measuring eye-gaze in VR when there is a hand-controller buttonpress to indicate a subjective mental state. An alternative method for research could be designing the experiment so that a button-press is always required (both for positive and negative instances).

References

- [1] Castillon I, Chartier T, Venkatesha V, Okada NS, Davis A, Cleary AM, et al. Automatically Identifying the Human Sense of Familiarity Using Eye Gaze Features. In: International Conference on Human-Computer Interaction. Springer; 2024. p. 291-310.
- [2] Okada NS, McNeely-White KL, Cleary AM, Carlaw BN, Drane DL, Parsons TD, et al. A virtual reality paradigm with dynamic scene stimuli for use in memory research. Behavior Research Methods. 2023;1-24.
- [3] Bixler R, Blanchard N, Garrison L, D'Mello S. Automatic detection of mind wandering during reading using gaze and physiology. In: Proceedings of the 2015 ACM on international Conference on Multimodal Interaction; 2015. p. 299-306.
- [4] Blanchard N, Bixler R, Joyce T, D'Mello S. Automated physiological-based detection of mind wandering during learning. In: 12th International Conference on Intelligent Tutoring Systems (ITS). Springer; 2014. p. 55-60.
- [5] Cleary AM, Irving ZC, Mills C. What Flips Attention? Cognitive Science. 2023;47(4):e13274.
- [6] Bixler R, D'Mello S. Automatic gaze-based user-independent detection of mind wandering during computerized reading. User Model User-Adap Interact. 2016;26:33-68.
- [7] Cleary AM, Brown AS, Sawyer BD, Nomi JS, Ajoku AC, Ryals AJ. Familiarity from the configuration of objects in 3-dimensional space and its relation to *déjà vu*: A virtual reality investigation. Conscious Cogn. 2012;21(2):969-75.
- [8] Cleary AM, Claxton AB. *Déjà vu*: An illusion of prediction. Psychol Sci. 2018;29(4):635-44.
- [9] Cleary AM. Recognition memory, familiarity, and *déjà vu* experiences. Curr Dir Psychol Sci. 2008;17(5):353-7.
- [10] Cleary AM, McNeely-White KL, Huebert AM, Claxton AB. *Déjà vu* and the feeling of prediction: An association with familiarity strength. Memory. 2021;29(7):904-20.
- [11] Nishimura G, Faisal A. *Déjà vu*: Classification of memory using eye movements. 2015.

Satisfaction with the CALMA Online Psychological First Aid Program in Peru

Karen PÉREZ-MARAVÍ^a, Giovanna OSORIO^a and Iván MONTES-ITURRIZAGA^{a,1}

^a*School of Psychology, Universidad Continental, Huancayo, Perú*

ORCID ID: Karen Pérez-Maraví <https://orcid.org/0000-0003-0068-4708>

Abstract. During the COVID-19 pandemic, the Universidad Continental (Peru) initiated an online Psychological First Aid program (CALMA). This free and inclusive program was designed with the World Health Organization's Psychological First Aid Technical Guide (2012) in mind. This study offers a quantitative, observational (non-experimental) approach and, through the survey technique, to the knowledge of the satisfaction of 1078 beneficiaries (out of 10096) who received care between March 2020 and April 2024. Therefore, we found a response rate of 10.7% to this form (Google Forms) anonymous and provided under informed consent. The results indicate that 96.6% of the subjects considered that their reason for consultation was attended to. Regarding the level of satisfaction with the care provided, 81.6% thought they were delighted. The main advantages perceived by the participants concerning the service were free, immediate, online, inclusive, and respectful of human diversity 40.9%. The results also suggest differences in satisfaction and timing of care (pandemic, transition, and post-pandemic).

Keywords. Psychological first aid, satisfaction, telepsychology, COVID-19, accessibility, quality

1. Introduction

During the COVID-19 pandemic, the Universidad Continental (Peru) initiated an online psychological first aid program (CALMA Program). This program was designed with the World Health Organization's technical guide for psychological first aid in mind [1] and has as an essential reference the contributions of the ABCDE model developed by Paula Cortez and Rodrigo Figueroa of the Pontificia Universidad Católica de Chile [2].

In this framework, online psychological first aid is a human and professional response of support to another person suffering and needing help in an unexpected situation, i.e., an untimely crisis. Similarly, this initiative was developed within the framework of telepsychology as an inclusive, easily accessible service that enables close access to mental health services through diverse digital means (Zhou et al., 2020; Goldschmidt et al., 2021; APA, 2023) [3]. However, it is essential to highlight that the pandemic caused a vertiginous revolution in the way of providing professional care in the field of psychology (Taylor, 2020), where teleconsultations through digital platforms predominated, being, to a lesser extent, the experiences of crisis care under an inclusive, free, and accessible scheme.

Thus, the CALMA has been offering care under a highly protocolized scheme and with specially elaborate standards to take care of processes such as the selection of volunteer psychologists (and psychology interns), their training, and strict supervision. In the same way, CALMA established a working system attentive to diversity in all its forms and within reasonable accommodations to conditions of visual, hearing, motor, and speech disabilities, among other conditions [4]. In this way, more than 10,000 services have been offered through digital media under the ABCDE model adapted for online application [5]. In addition, the high acceptance and urgent need for psychological help during the COVID-19 pandemic led us—in its beginnings—to an average of 14 hours a day of attention, thanks to the work of 160 volunteers (Continental University) among students, interns, and psychologists.

¹ Corresponding Author: Iván Montes-Iturrizaga, imontes@uc.cl

However, very little is known about satisfaction with online psychological first aid services. And there is even less research that tells us about the impacts of these programs on the lives of the beneficiaries [6]. Thus, it is known that telepsychology services satisfy populations that cannot move normally because of geographical and economic reasons or disability situations. In the same way, the immediacy and identity reserved for the consultant or patient who, for some reason (such as shyness, fear, or embarrassment), does not want to show his or her face are highlighted.

On the other hand, in the online psychological first aid field, there are few studies on its impact and satisfaction [7–9]. In this way and given that it is a service in the field of telepsychology, previous studies (in this broader framework) report that the consultants or beneficiaries value the punctuality of the appointments, the technological resources of the psychologist, the warmth of the treatment and the protocolized management of the session and the reasonable adjustments to the diversities and situations of disability of diverse nature. In this sense, it is an important challenge to be able to carry out much more in-depth studies to learn how online psychological first aid (and also face-to-face) allowed the crisis or emergency that placed the patient in a situation of extreme vulnerability to be attended [6,8,10–12]. In particular, this study offers a quantitative, observational (non-experimental) approach through the survey technique to determine the satisfaction of 1078 beneficiaries (out of a total of 10096) who received care between March 2020 and April 2024.

2. Methods

2.1. Participants

The sample comprises 1078 people who, after receiving care in the Program (CALMA), voluntarily decided to answer the satisfaction questionnaire. Thus, the response rate is 10.7% if we consider that 10096 people requested this service and received the satisfaction questionnaire. It is important to mention that we are considering in this report the entire population served by CALMA since its inception (March 30, 2020) and up to the present (April 17, 2024) who responded to the satisfaction questionnaire.

2.2. Instrument

To measure satisfaction with the care received in the CALMA, a questionnaire was prepared (anonymous and answered under informed consent) consisting of the following items: *Was your appointment scheduled on the day and time requested?* (item 1); *Was your reason for consultation attended to in the online Psychological First Aid service?* (item 2); *How many times did you access the online Psychological First Aid service?* (item 3); *How satisfied or dissatisfied do you feel with the care provided by the online Psychological First Aid Volunteer Psychologist?* (item 4) *Do you consider that online Psychological First Aid helps people to reduce their psychological manifestations in a period of a health emergency?* (item 5); and, *what were the advantages you found in the online Psychological First Aid (CALMA) care?* (item 6).

2.3. Statistical Analyses

Statistical analyses were performed in SPSS version 28 for Windows and JAMOVI version 2.3.28. Descriptive analyses and Pearson's Chi-square (χ^2) analysis were performed in order to analyze the possible relationship between our categorical variables. Analyses are established considering the pandemic time, transition time and post-pandemic time.

3. Results

The results show that 96.6% considered that their reason for consultation was attended (item 2). Likewise, regarding the number of times they have used this program, it was reported that 83.1% went once, 10.7% twice, 2.8% three times, and 3.4% more than three

times (item 3). Regarding the level of satisfaction with the care provided by the volunteer, 2.8% were very dissatisfied, 1.5% were somewhat dissatisfied, 2.4% were neither satisfied nor dissatisfied, 11.8% were somewhat satisfied, and 81.6% were very satisfied (item 4). Regarding the advantages perceived by the participants with respect to online psychological first aid (and from their experience in the CALMA Program, the 3 main responses were the following: 1) free, immediate, online, inclusive, and respectful of human diversity care in 28.1%; 2) free, immediate and online psychological care in 12.8%; and 3) free and immediate psychological care in 8.0% (item 6). The least chosen options appealed to empathy, active listening, ethical aspects, and the diversity of schedules for attending to consultants (item 6). The results also suggest differences in satisfaction and timing of care, whether pandemic (March 30, 2020, to July 31, 2022), transition (August 1 to December 31, 2022), and post-pandemic (January 1, 2024, to present). Finally, in a Likert-formatted opinion question regarding "whether psychological first aid services help people in times of health crisis," we found that the two majority response categories were "strongly agree" with 67.6% and "agree" with 30.1% (item 5).

Specifically, **Table 1** shows satisfaction results at the time of the pandemic. Here it can be seen that satisfaction concerning the care received by the volunteer psychologist remained high at all three times related to the pandemic by COVID-19. Likewise, a significant association was found between satisfaction concerning the volunteer psychologist and the time or moment of the pandemic ($\chi^2(8, N=1078) = 27.8, p < .001$).

Table 1. How satisfied or dissatisfied do you feel with the care provided by the "Psychological First Aid Online" Volunteer Psychologist?

	n	%
Pandemic momentum		
<i>Very dissatisfied</i>	17	2.7%
<i>Somewhat dissatisfied</i>	9	1.4%
<i>Neither satisfied nor dissatisfied</i>	23	3.6%
<i>Somewhat satisfied</i>	95	15.0%
<i>Very satisfied</i>	491	77.3%
Total	635	100%
Transition		
<i>Very dissatisfied</i>	0	0%
<i>Somewhat dissatisfied</i>	1	2.5%
<i>Neither satisfied nor dissatisfied</i>	0	0%
<i>Somewhat satisfied</i>	3	7.5%
<i>Very satisfied</i>	36	90%
Total	40	100%
Post pandemic momentum		
<i>Very dissatisfied</i>	12	3%
<i>Somewhat dissatisfied</i>	6	1.5%
<i>Neither satisfied nor dissatisfied</i>	3	0.7%
<i>Somewhat satisfied</i>	29	7.2%
<i>Very satisfied</i>	353	87.6%
Total	403	100%

Table 2 shows the evaluative perception of the importance of online psychological first aid according to the moments of the pandemic. It can thus be seen that the considerations were very favorable at all times of the pandemic, and where the perception of importance was higher at the post-pandemic time ($\chi^2(8, N=1078) = 29.3, p < .001$).

Table 2. Do you consider that Psychological First Aid Online helps people to reduce their psychological manifestations in a period of a health emergency?

	n	%
Pandemia		
Totally agree	399	62.8%
agree	207	32.6%
<i>Indifferent</i>	10	1.6%
Disagree	5	0.8%
<i>Totally disagree</i>	14	2.2%
Total	635	100%
Transition		
Totally agree	23	57.5%
agree	17	42.5%
<i>Indifferent</i>	0	0.0%
Disagree	0	0.0%
<i>Totally disagree</i>	0	0.0%
Total	40	100%
Post pandemic momentum		
<i>Totally agree</i>	299	74.2%
agree	97	24.1%
<i>Indifferent</i>	3	0.7%
Disagree	2	0.5%
<i>Totally disagree</i>	2	0.5%
Total	403	100%

4. Discussion

The results indicate high satisfaction and perception of the importance of the psychological first aid service offered by CALMA, evidence that would be primarily associated with the fact that more than 95% considered that the reason for the consultation had been attended. Likewise, there is evidence of high satisfaction in the three moments of a pandemic, but in the transition time and post-pandemic, satisfaction was higher. This result does show that the service was pertinently focused on the reason that initially troubled the beneficiaries. It is also worth noting that the sample size represents only 10.7% of the total number of people served over the years and who, most likely, would be the most satisfied. Finally, we consider that a limitation of the study is that the questionnaire did not include data on sex, age, and reason for consultation. This did not allow for inferential analysis.

References

- [1] World Health Organization, War Trauma Foundation, World Vision International. Psychological first aid: Guide for field workers. Geneva: WHO; 2011. Available from: <https://www.who.int/publications/item/9789241548205>.
- [2] Cortés-Montenegro P, Figueroa-Cabello R. Manual ABCDE para la aplicación de Primeros Auxilios Psicológicos. Chile: Pontificia Universidad Católica de Chile y Centro Nacional de Investigación para la Gestión Integrada de Desastres Naturales (CIGIDEN); 2016. Available from: <https://medicina.uc.cl/publicacion/manual-abcd-e-la-aplicacion-primeros-auxilios-psicologicos/>
- [3] Pierce BS, Perrin PB, Tyler CM, McKee GB, Watson JD. The COVID-19 telepsychology revolution: A national study of pandemic-based changes in U.S. mental health care delivery. Am Psychol. 2021;76(1):14–25. doi: <https://doi.org/10.1037/amp0000722>.
- [4] Ramos del Río B, Martí-Noguera JJ, coordinators. Experiencias en ciberpsicología hacia una nueva era de la psicología [En línea]. Ciudad de México: Universidad Nacional Autónoma de México; 2022. Available from: <https://www.zaragoza.unam.mx/experiencias-ciberpsicologia/>
- [5] Carrion-Lalangui IN, Molina-Velasco AZ, Saavedra-Vilchez JA. Adaptación de la Guía ABCDE de primeros auxilios psicológicos a un contexto virtual Perú 2022 [undergraduate thesis]. Lima: Universidad Continental; 2023. Available from: <https://hdl.handle.net/20.500.12394/14027>
- [6] Argüero-Fonseca A, Aguirre-Ojeda DP, Reynoso-González OU, Girón-Pérez MI, Espinosa-Parra IM, Sierra-Murguía MA. Impacto de la telepsicología en la satisfacción de la atención a pacientes con Covid-19. Psicol Iberoam. 2021;29(3):e293325. doi: <https://doi.org/10.48102/pi.v29i3.325>.
- [7] Colón-Llamas T, Escobar-Altare A, Ramírez-Castro MF, Bermúdez-Jaimes ME, López-Sánchez MC, et al. Telepsicología en un centro de atención psicológica universitario durante el COVID-19: Experiencias

- de consultantes y estudiantes. Univ Psychol. 2022;21:1–18. doi: <https://doi.org/10.11144/Javeriana.upsy21.tcap>.
- [8] Rojas-Jara C, Polanco-Carrasco R, Caycho-Rodríguez T, Muñoz-Vega C, Muñoz-Marabolí M, Luna-Gómez T, et al. Telepsychology for psychotherapists: lessons learned in times of Covid-19. Interam J Psychol. 2022;56(2):e1733. doi: <https://doi.org/10.30849/ripijp.v56i2.1733>.
- [9] Comité Permanente entre Organismos. Consideraciones operacionales para programas multisectoriales de salud mental y apoyo psicosocial durante la pandemia de la COVID-19 VERSIÓN 1.1. IASC; 2020. Available from: <https://www.paho.org/es/documentos/consideraciones-operacionales-para-programas-multisectoriales-salud-mental-apoyo>
- [10] Feijt M, de Kort Y, Bongers I, Bierbooms J, Westerink J, IJsselsteijn J. Mental health care goes online: mental health during COVID-19. Cyberpsychol Behav Soc Netw. 2020;23(12):860–4. doi: <https://doi.org/10.1089/cyber.2020.0370>.
- [11] Sandoval-Obando E, Sandoval-Díaz J. Psicología de la emergencia en contexto de pandemia. Tesis Psicol. 2020;15(2):254–79. doi: <https://doi.org/10.37511/tesis.v15n2a14>.
- [12] Balluerka N, Gómez J, Hidalgo D, Gorostiaga A, Espada JP, Padilla JL, et al. Las consecuencias psicológicas de la COVID-19 y el confinamiento. Informe de Investigación. España: Universidad del País Vasco, UNED, Universitat de Barcelona, Universidad de Granada, Universidad Miguel Hernández y Universidad de Murcia; 2020. Available from: <https://www.cedid.es/es/documentacion/catalogo/Record/551534>

Exploring the Impact of Gamified Virtual Reality Cognitive Training on Improving Anxiety and Depressive Symptoms in Spanish Adolescent Girls

Aikaterini AMPATZOGLOU^a, Anna CARBALLO-MARQUEZ^a, Anna GARCIA-CASANOVASA, Juliana ROJAS^a, Maria FERNANDEZ-CAPO^a, Maria GAMIZ-SANFELIU^a, Maite GAROLERA^{a,b} and Bruno PORRAS-GARCIA^{a,b,1}

^a*BrainXR Lab, Area of Psychology and Mental Health, Department of Basic Sciences.*

Faculty of Medicine and Health Science, International University of Catalonia, Spain.

^b*Brain, Cognition, and Behavior Research Group, Consorci Sanitari de Terrassa-Hospital Universitari, Spain.*

ORCID ID: Aikaterini Ampatzoglou <https://orcid.org/0009-0004-3048-3557>

Abstract. Anxiety and depressive symptoms during adolescence have been linked to adverse outcomes in academic, social, and health domains. Interventions promoting emotion regulation (ER) serve as a preventive strategy for developing those symptoms. A key aspect to consider in the development and regulation of ER in adolescents is executive function (EF). Immersive virtual reality (IVR) has emerged as an innovative tool to enhance the effectiveness of EF training interventions. This study aims to assess the preliminary effect of IVR-gamified cognitive training on improving ER and reducing internalizing symptoms in young adolescent girls. Methodology: This is a longitudinal, parallel, single-blind, randomized pilot study with a sample of 45 Spanish-speaking adolescent girls. The sample was randomly allocated into an experimental group receiving a gamified IVR cognitive training program and a control group receiving a comparable IVR relaxation experience. Participants underwent a baseline and post-intervention assessment with the Revised Child Anxiety and Depression scale (RCADS) and the emotional control subscale from the Behavior Rating Inventory of Executive Functions (BRIEF-2). Results: Two-way mixed ANOVAs were conducted, revealing no significant group*time interactions for internalizing symptoms, total anxiety symptoms, depressive symptoms, and emotional control. Main effects of time were significant for internalizing symptoms and total anxiety symptoms. Pairwise comparison tests indicated statistically significant reductions in internalizing symptoms and total anxiety symptoms. These reductions were particularly noticeable in the experimental group. Conclusions: The results suggest that both interventions had a positive effect on participants' internalizing symptoms, with the experimental group showing a more pronounced reduction.

Keywords. Immersive virtual reality, emotion regulation, executive functions, adolescence, internalizing symptoms

1. Introduction

Adolescents are increasingly recognized as one of the most vulnerable groups affected by mental health problems globally. According to the World Health Organization (WHO) [1] depression and anxiety pose significant challenges to mental health initiatives worldwide among this population. In addition, a gender gap in depression and anxiety becomes apparent during adolescence. From approximately ages 11 to 15, girls exhibit a higher likelihood of experiencing clinical depression and elevated depressive symptoms compared to boys, and this trend persists well into adulthood [2]. Adolescent depression

¹ Corresponding Author: Bruno-Porras Garcia, bporras@uic.es.

and anxiety can significantly impact various aspects of life into adulthood, including education, employment and relationships. This can lead to feelings of loneliness, sadness, and diminished self-esteem, particularly affecting girls whose self-concept is more intertwined with their interpersonal relationships [3].

Emotion regulation (ER) plays a crucial role in preventive treatments targeting depressive and anxiety symptoms in adolescents⁴. ER involves adjusting emotions in response to environmental stimuli, employing both automatic and deliberate cognitive processes. According to Compas and colleagues [4] individuals experiencing emotional dysregulation struggle with intentionally focusing and shifting attention, as well as controlling or initiating behaviors to regulate their emotional responses.

Deficits in executive function (EF) are seen as pivotal in the formation and control of emotions. EF refers to a series of higher-order cognitive processes linked with the brain's prefrontal cortex that support goal-directed activities, including working memory, planning, set shifting, and inhibitory control [5]. Key functions such as working memory, inhibition, and cognitive flexibility are considered essential for effective ER. According to research, children and adolescents with more effective EF skills may better regulate their emotions in response to stressful situations and have fewer internalizing symptoms [6].

Early identification, monitoring and preventive transdiagnostic interventions for emotion dysregulation symptoms in children and adolescents are important in reducing the development of disorders later in life. Consequently, targeting EF abilities may operate as a preventive measure against emotion dysregulation and the emergence of Internalizing symptoms in adolescents.

In recent times, immersive virtual reality (IVR) has emerged as an innovative tool for enhancing the accessibility and efficacy of EF training interventions. IVR enables users to immerse themselves in three-dimensional environments, offering highly engaging and realistic interactions. A systematic review and meta-analysis found that people who used IVR exergames experienced improved global cognition and memory, as well as reduced depressive symptoms [7-8].

This study aims to assess the preliminary effects of IVR-gamified cognitive training on improving ER and reducing anxiety and depressive symptoms in young adolescent girls. We hypothesize that both interventions will have a positive effect on participants' anxiety and internalizing symptoms.

2. Methods

This is a longitudinal, parallel, single-blind, randomized pilot study with a sample of 45 Spanish-speaking adolescent girls, aged 11-14 years old. These preliminary results belong to a larger study adhered to CONSORT guidelines, which was registered on clinicaltrials.gov (NCT06236919). This study has been approved by the Ethics Committee of the International University of Catalonia (PSI-2023-04).

2.1. Participants

Spanish school-aged female participants (aged 11 to 14 years old; Mage = 12.57 years; SD= 0.39 years), whose parents agreed to participate, were recruited in the study. Participants were recruited from a private school located in the metropolitan area of Barcelona, Catalonia.

Participants were included based on the following criteria: (a) between the ages of 10 years and 16 years; (b) able to read, understand and speak Spanish; and c) individuals who provide consent to participate (and their legal guardians). Participants were excluded in case they had: (a) a prior diagnosis of severe psychiatric conditions (e.g., manic or psychotic symptoms) or serious neurodevelopmental disorders (e.g., severe autism spectrum disorders, intellectual disabilities), whether self-reported or reported by family or school, (b) other physical, motor, or sensory impairments that could interfere with the examination or the IVR program.

2.2. Measures

Participants underwent a baseline and post-intervention assessment with the Revised Child Anxiety and Depression scale (RCADS-47) [9] and the Behavior Rating Inventory of Executive Functions, Second Edition (BRIEF-2) [10].

The RCADS-47 is a 47-item, youth self-report questionnaire that assesses anxiety and depression symptoms including six subscales. In our analysis, we focused specifically on the total anxiety scale, the major depressive disorder (MDD) subscale, and the total internalizing scale. Higher scores indicate higher levels of anxiety and/or depression.

The BRIEF-2 questionnaire is a test intended for the evaluation of the behavioral and observable aspects of the EFs of children and adolescents between 5 and 18 years old. It consists of 63 items, and it evaluates different scales related to EFs such as inhibition, flexibility and emotional control. Higher scores reflect greater difficulty in EF. For our purposes, we only used the emotional control subscale.

2.3. Procedure

Both Enhance IVR (experimental group) and the Control IVR (control group) interventions lasted five weeks, twice a week for 30 minutes (10 sessions per participant), with intervention parameters based on similar research in this area [11]. Each participant was provided with a single VR headset, and the intervention was delivered as a face-to-face activity in groups of up to six participants, supervised by MSc and PhD level psychologists. Study completion was defined as participants attending 80% or more of the scheduled sessions.

The Enhance IVR group was engaged in a subset of six games (Enhance VR, Virtuleaps) designed to target specific cognitive functions (see Table 1). The control IVR group was immersed in ten different nature-based VR environments (NatureTreks,

Greener Games) where they were able to tailor the content and perform relaxation exercises (deep breathing). Some examples of the environments included: Blue Ocean (a tropical beach with palm trees) and Orange Sunset (a wooded forest with red sunset skies).

Table 1. Enhance VR games.

Game name	Category	Neuropsychological test	Game description
Pizza builder	Attention & planification	Dual task paradigms [12]	Assemble multiple pizzas according to continuously incoming and changing orders
Shuffled	Attention	Moving boxes task [13]	Track a single jellyfish and ignore the rest as they move around.
React	Cognitive flexibility & impulsivity	Wisconsin card sorting test [14] & Stroop task [15]	Throw objects into portals according to their shape and colour.
Assembly	Processing speed & impulsivity	Trail making test [16]	Select, one by one, gears from a group in ascending size.
Maestro	Working memory	N-back test [17]	Memorize light patterns and report when the patterns repeat.
Memory Wall	Memory	Visual patterns test [18]	Memorize and reproduce a pattern of cubes.

Notes: The Enhance VR library is composed of 15 games (cognitive exercises).

2.4. Hardware

For the intervention, we utilized the Meta Quest 3, a popular standalone VR headset (developed by Meta, California). The subject interacts with the virtual environment using two hand motion controllers. It features a Qualcomm Snapdragon XR2 Gen 2 chipset, 2064x2208 per eye resolution and 8 GB memory.

2.5. Data Analyses

Analyses were conducted using SPSS v.25. Two-way mixed ANOVAs were conducted to examine the effects of group (experimental vs. control) and time (pre- vs. post-intervention) on internalizing symptoms, total anxiety symptoms, depressive symptoms, and emotional control.

3. Results

Two-way mixed ANOVAs analysis did not reveal any significant group*time interactions for any of the variables. However, there were significant main effects of time on both internalizing symptoms and total anxiety symptoms (see **Table 2**).

Table 2. Two-way mixed ANOVA results for ER and mental health outcomes.

Measures	Group*time			Main effect time		
	F	p	np ²	F	p	np ²
BRIEF-2-SR Emotional control	0.30	.587	.007	0.16	.688	.004
RCADS total anxiety	1.16	.287	0.26	7.37	.009	.146
RCADS total depression	0.94	.336	.022	1.78	.188	.40
RCADS total internalizing	1.23	.272	0.28	6.38	.015	.129

Notes: BRIEF-2-SR = Behavior Rating Inventory of Executive Functions, Second Edition; RCADS = Revised Child Anxiety and Depression scale.

Pairwise comparisons indicated a statistically significant mean reduction in internalizing symptoms from pre- to post-assessment ($MD = 5.440$, $SE = 2.152$, $p = .015$). Similarly, total anxiety symptoms also showed a significant decrease over time, with pairwise comparisons revealing a significant mean reduction from pre- to post-assessment ($MD = 4.741$, $SE = 1.746$, $p = .009$). These reductions were more pronounced in the experimental group compared to the control group (see **Figure 1**). No significant main effects of time were found for depressive symptoms or emotional control.

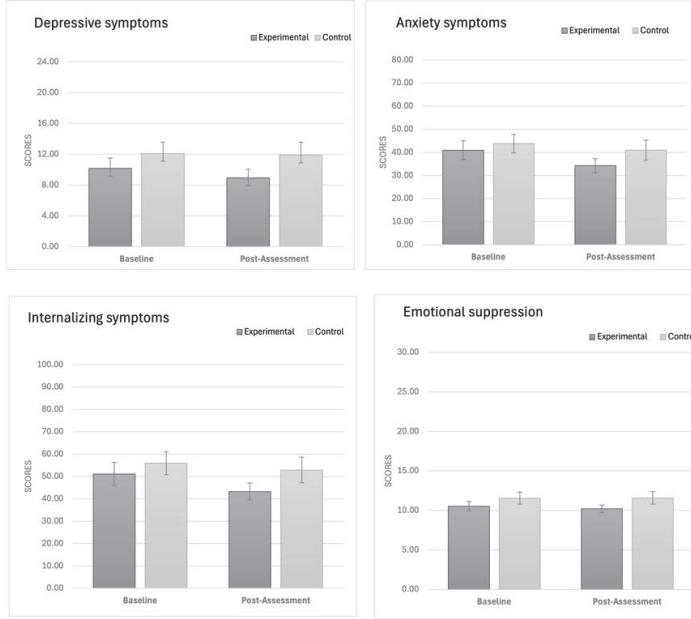


Figure 1. Baseline and post assessment outcomes for the mental health and ER measures

4. Discussion

The objective of the present was to assess the preliminary effect of an IVR-gamified cognitive training on improving ER and reducing anxiety and depressive symptoms in young adolescent girls. The study focused on adolescent girls since they usually show higher vulnerability to developing anxiety and mood disorders compared to adolescent boys [3].

Both interventions resulted in a reduction of participants' internalizing symptoms. Specifically, the experimental group, which received the IVR-gamified cognitive training, exhibited a more noticeable reduction in anxiety and depressive symptoms compared to the control group. These findings are consistent with previous research [19], indicating the effectiveness of cognitive training in alleviating internalizing symptoms such as anxiety and depression. A recent study [20], found that working memory training can decrease anxiety symptoms and improve ER. In addition, cognitive training that targets attention has been shown to reduce internalizing symptoms in adolescents at risk for anxiety [21]. A key distinction between our research and previous studies is the use of IVR cognitive training programs, instead if traditional cognitive training methods. IVR training offers several advantages, making our approach distinct and potentially more effective. Studies have indicated that IVR enhances the effectiveness and accessibility of cognitive training for children and adolescents [22], while the gamification parameter within IVR significantly boosts engagement and motivation [23]. This combination of increased engagement, realistic scenarios, and self-regulated learning makes IVR a superior method for cognitive training compared to traditional approaches.

In conclusion, the findings from this preliminary study suggest that IVR-gamified cognitive training holds promise as an effective intervention for improving ER and reducing anxiety and depressive symptoms in adolescent girls. Future research should encompass a larger sample size that includes adolescent boys, longer intervention duration and long-term follow-up assessment.

Acknowledgments

This study was funded and supported by a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation (ID: 31607) and by the Catalan Agency for Management of University and Research Grants (AGAUR, ID:2023IMPAC00019) from Generalitat de Catalunya. We would also like to express our sincere gratitude to Mr. Amir Bozorgzadeh and Dr Bebiana Moura, Head of Partnerships at Virtuleaps, for their

crucial technical support in designing the Enhance IVR virtual games research protocol. Their expertise and dedication, together with the invaluable support of the VirtuLeaps team, were instrumental in tackling the complex technical issues before and during recruitment.

References

- [1] World Health Organization. COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide [Internet]. 2022 Mar 2 [cited 2022 Nov 14]. Available from: <https://www.who.int/news-room/detail/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide>.
- [2] Chaplin TM, Hong K, Bergquist K, Sinha R. Gender differences in response to emotional stress: An assessment across subjective, behavioral, and physiological domains and relations to alcohol craving. *Alcohol Clin Exp Res* [Internet]. 2008;32(7):1242–50. Available from: <http://dx.doi.org/10.1111/j.1530-0277.2008.00679.x>
- [3] Rose AJ, Rudolph KD. A review of sex differences in peer relationship processes: Potential trade-offs for the emotional and behavioral development of girls and boys. *Psychol Bull* [Internet]. 2006;132(1):98–131. Available from: <http://dx.doi.org/10.1037/0033-2909.132.1.98>
- [4] Compas BE, Jaser SS, Bettis AH, Watson KH, Gruhn MA, Dunbar JP, et al. Coping, emotion regulation, and psychopathology in childhood and adolescence: A meta-analysis and narrative review. *Psychol Bull* [Internet]. 2017;143(9):939–91. Available from: <http://dx.doi.org/10.1037/bul0000110>
- [5] Garon N, Bryson SE, Smith IM. Executive function in preschoolers: A review using an integrative framework. *Psychol Bull* [Internet]. 2008;134(1):31–60. Available from: <http://dx.doi.org/10.1037/0033-2909.134.1.31>
- [6] Lantrip C, Isquith PK, Koven NS, Welsh K, Roth RM. Executive function and emotion regulation strategy use in adolescents. *Appl Neuropsychol Child* [Internet]. 2016;5(1):50–5. Available from: <http://dx.doi.org/10.1080/21622965.2014.960567>
- [7] Li J, Theng Y-L, Foo S. Effect of exergames on depression: A systematic review and meta-analysis. *Cyberpsychol Behav Soc Netw* [Internet]. 2016;19(1):34–42. Available from: <http://dx.doi.org/10.1089/cyber.2015.0366>
- [8] Cugusi L, Prosperi L, Mura G. Exergaming for quality of life in persons living with chronic diseases: A systematic review and meta-analysis. *PM R* [Internet]. 2021;13(7):756–80. Available from: <http://dx.doi.org/10.1002/pmrj.12444>
- [9] Sandín B, Valiente RM, Chorot P. RCADS : evaluación de los síntomas de los trastornos de ansiedad y depresión en niños y adolescentes. *Rev psicopatol psicol clín* [Internet]. 2009;14(3). Available from: <http://dx.doi.org/10.5944/rppc.vol.14.num.3.2009.4078>
- [10] Gioia G, Isquith A, Guy K, Kenworthy C. Behavior Rating Inventory of Executive Function, Second Edition (BRIEF2). Psychological Assessment Resources. 2015.
- [11] Jahn FS, Skovbyre M, Obenhausen K, Jespersen AE, Miskowiak KW. Cognitive training with fully immersive virtual reality in patients with neurological and psychiatric disorders: A systematic review of randomized controlled trials. *Psychiatry Res* [Internet]. 2021;300(113928):113928. Available from: <http://dx.doi.org/10.1016/j.psychres.2021.113928>
- [12] Tedstone D, Coyle K. Cognitive impairments in sober alcoholics: performance on selective and divided attention tasks. *Drug Alcohol Depend* [Internet]. 2004;75(3):277–86. Available from: <http://dx.doi.org/10.1016/j.drugalcdep.2004.03.005>
- [13] Rivière J, Brisson J. Space-based visual attention: A marker of immature selective attention in toddlers? *Dev Psychobiol* [Internet]. 2014;56(7):1589–94. Available from: <http://dx.doi.org/10.1002/dev.21240>
- [14] Grant DA, Berg E. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *J Exp Psychol* [Internet]. 1948;38(4):404–11. Available from: <http://dx.doi.org/10.1037/h0059831>
- [15] Stroop JR. Studies of interference in serial verbal reactions. *J Exp Psychol* [Internet]. 1935;18(6):643–62. Available from: <http://dx.doi.org/10.1037/h0054651>
- [16] Reitan RM. Validity of the trail making test as an indicator of organic brain damage. *Percept Mot Skills* [Internet]. 1958;8(7):271. Available from: <http://dx.doi.org/10.2466/pms.8.7.271-276>
- [17] Kirchner WK. Age differences in short-term retention of rapidly changing information. *J Exp Psychol* [Internet]. 1958;55(4):352–8. Available from: <http://dx.doi.org/10.1037/h0043688>
- [18] Gray C, Baddeley A, Wilson L. Visual Patterns Test. A Test of Short-Term Visual Recall. Bury St Edmunds: Thames Valley Test Company; 1997.
- [19] Bowie CR, Gupta M, Holshausen K. Cognitive remediation therapy for mood disorders: rationale, early evidence, and future directions. *Can J Psychiatry* [Internet]. 2013;58(6):319–25. Available from: <http://dx.doi.org/10.1177/070674371305800603>
- [20] Cui X, Zhang S, Yu S, Ding Q, Li X. Does working memory training improve emotion regulation and reduce internalizing symptoms? A pair of three-level meta-analyses. *Behav Res Ther*. 2024 Aug;179:104549. doi: <http://dx.doi.org/10.1016/j.brat.2024.104549>. Epub 2024 Apr 28. PMID: 38761555
- [21] Eldar S, Aptek A, Lotan D, Edgar KP, Naim R, Fox NA, et al. Attention bias modification treatment for pediatric anxiety disorders: A randomized controlled trial. *Am J Psychiatry* [Internet]. 2012;169(2):213–30. Available from: <http://dx.doi.org/10.1176/appi.ajp.2011.11060886>
- [22] Romero-Ayuso D, Toledano-González A, Rodríguez-Martínez M del C, Arroyo-Castillo P, Triviño-Juárez JM, González P, et al. Effectiveness of virtual reality-based interventions for children and adolescents with ADHD: A systematic review and meta-analysis. *Children (Basel)* [Internet]. 2021;8(2):70. Available from: <http://dx.doi.org/10.3390/children8020070>

- [23] Zhu S, Sui Y, Shen Y, Zhu Y, Ali N, Guo C, et al. Effects of virtual reality intervention on cognition and motor function in older adults with mild cognitive impairment or dementia: A systematic review and meta-analysis. *Front Aging Neurosci* [Internet]. 2021;13. Available from: <http://dx.doi.org/10.3389/fnagi.2021.586999>

Development of a Virtual Reality Driving Simulator for Cognitive Research: a Motion Sickness Analysis

Maria Eleonora MINISSI^{a,1}, Luna MADDALON^a, Juan Miguel ROCHINA^a, Pablo ESCRIBANO-BASTANTE^a and Mariano ALCAÑIZ^a

^a*Instituto Universitario de Investigación en Tecnología Centrada en el Ser Humano (HUMAN-tech), Universitat Politècnica de Valencia, Valencia, Spain*

ORCiD ID: Maria Eleonora Minissi <https://orcid.org/0000-0001-6326-0609>

Abstract. This paper presents the development of a Head Mounted Display (HMD)-based Virtual Reality (VR) driving simulator designed for cognitive and behavioral research, offering two distinct versions: a virtual driving experience and a virtual passenger experience. The aim is to create an immersive, standardized environment to investigate cognitive and behavioral effects related to driving. The virtual driving experience involved active participation using a steering wheel and pedals, while the virtual passenger experience simulated an autonomous vehicle ride without user control. To reduce the risk of motion sickness in users, two rounds of testing were conducted. The first round involved 18 participants, divided into the two simulator versions, with results indicating significant motion discomfort primarily due to car braking rather than turning. Based on these findings, adjustments were made to reduce braking speed. A subsequent round of testing with 10 participants evaluated the effectiveness of these modifications. Results showed reduced motion discomfort compared to the first round of testing and satisfactory use of the applications by users. Future research on cognitive and behavioral effects during driving can benefit from this type of VR driving simulator. Additionally, exploring the use of more precise and reliable measures for motion sickness in modern VR systems is necessary for advancing the application of VR in cognitive and psychological research.

Keywords. Virtual reality; driving simulator; motion sickness; HMDs

1. Introduction

Virtual reality (VR) has emerged as a promising tool in cognitive research, offering unparalleled opportunities to replicate complex, real-world tasks within controlled experimental settings. It can address the need for ecological validity in experimental procedures, as it prompts users to exhibit realistic reactions while performing virtual tasks as if they were in the real world [1].

Among the complex real-world tasks that can be studied through VR, there is driving. In recent decades, many traffic studies have used 3D driving simulators composed of a high-performance computer connected to a driver's seat, a steering wheel, pedals, and a gear shifter located in front of three high-resolution screens [e.g., 2-4]. Nevertheless, while these systems are widely used, they are not fully immersive and do not guarantee the manifestation of intuitive reactions in drivers [5]. Conversely, fully immersive VR, such as Head Mounted Displays (HMDs), provides a significant sense of presence in the virtual environment and completely isolates users from the real world [6]. Moreover, the field of view in HMDs is similar to the real world, in contrast to 3D driving simulators. Many HMD-based driving simulators have been developed, although they are used mostly for commercial and entertainment purposes rather than research. These commercial driving simulators allow users a high degree of freedom in driving behavior, reducing the chances of providing standardized situations necessary for rigorous scientific research to objectively study cognitive and behavioral phenomena. Moreover,

¹ Corresponding Author: Maria Eleonora Minissi, meminissi@i3b.upv.es.

to reduce the risk of fostering unpleasant reactions and/or drop-off from studies employing VR driving simulators, pre-tests of the application are encouraged to mitigate the risk of motion sickness in users.

Motion sickness in VR (also called cybersickness) refers to a variety of symptoms, such as nausea and dizziness, that may appear due to sustained visual exposure to motion, while no physical motion is felt [5]. Reducing the risk of motion sickness in a VR driving simulator is pivotal for conducting effective ecological research. Consequently, in this study, we describe a newly developed VR driving simulator for conducting cognitive research, which is divided into two versions, and we also present our findings from two rounds of testing to reduce the risk of motion sickness in users.

2. Methods

2.1. VR driving simulator

The VR driving simulator setup included a computer, an HP Reverb G2 Omnicept Edition HMD for projecting the virtual environments, and a Subsonic SV250 Superdrive wheel and pedal unit, whose pedals corresponded to brake and accelerator. Nevertheless, the VR driving simulator is compatible with any type of steering wheel.

Two versions of a driving simulator were developed using Unity 3D software, one in which the participant drives a car using the steering wheel and pedals (i.e., the virtual driving experience), and another in which the participant is a passenger in an autonomous vehicle, without the need for steering wheel and pedals (i.e., the virtual passenger experience). The introduction of the virtual passenger experience relies on the recent advent of autonomous cars.

In both versions, the participant is located in the driver's seat (see Figure 1). To standardize the experience for all participants and across versions, the car followed a pre-set route based on GPS directions, which included turns and red lights where the participant's car would stop. GPS directions were visible on a screen located in the upper part of the car dashboard (see Figure 1), and participants received auditory instructions when the car needed to turn, like common GPS directions. The pre-set route consists of a 3-km urban track that participants travel at a maximum speed of 40 km/h.

To ensure a realistic environment, details such as pedestrians, cars traveling on the streets, a moving ambulance and commercial stores were modeled in the virtual environment. In addition, the 3-km urban track included: 32 zebra crossings, 27 traffic lights (6 of which are red, with 2 red lights per kilometer), 1 zebra crossing with a pedestrian crossing, and 45 traffic signs.

In the virtual driving experience, participants received visual and auditory real-time error messages on the GPS screen if they failed to adhere to the GPS guidance and traffic rules, such as braking at red traffic lights or when a pedestrian was crossing. When the car was approaching a turn, there was a specific portion of the street before the turn where the participant had to turn the steering wheel to the right side to record a correct turn. Conversely, if a correct turn of the steering wheel was not recorded within the defined street portion, the car turned autonomously, and an error message was displayed. To record a turn, participants had to turn the steering wheel more than 10% of its maximum rotation angle. Similarly, at red traffic lights, there was a portion of the street before the stop line where participants had to brake. If they failed to do so, the car braked autonomously, and an error message was displayed.

On the other hand, in the virtual passenger experience, participants passively experienced the virtual environment without needing to control the vehicle or the possibility of making errors, as if they were traveling in an autonomous car. GPS directions were provided visually and auditorily as in the virtual driving experience.



Figure 1: Virtual environment from the participant's perspective

2.2. Participants

The initial round of testing conducted to assess motion sickness involved 18 adult participants divided into two distinct groups. Thirteen individuals (6 females) were assigned to the virtual driving experience group ($M = 25.61$; $sd = 3.75$), while five individuals (2 females) were assigned to the virtual passenger experience group ($M = 27.8$; $sd = 6.34$). The second round of testing, aimed to evaluate if motion sickness was mitigated after speed adaptation due to the first testing findings, involved 10 adult participants allocated into two groups. Five individuals (4 females) were assigned to the virtual driving experience group ($M = 25.20$; $sd = 1.64$), and the remaining five (2 females) were assigned to the virtual passenger experience group ($M = 26.40$; $sd = 2.30$). Participation was voluntary and uncompensated. Written informed consent was obtained prior to participation.

2.3. Procedure

For each round of testing, the same procedure was applied to both experimental groups using the two versions of the virtual application. Participants first completed a demographic questionnaire and control questions concerning their driving ability and prior experience with VR. Subsequently, participants completed the Simulator Sickness Questionnaire (SSQ) [7] to measure motion sickness prior to testing. Participants then underwent a VR familiarization step, which consisted of a brief introduction of car's characteristics and a simulation mirroring the experimental phase. This simulation presented the same type of application (driving or passenger) that participants would have experienced afterward. In the driving experience group, this step could be repeated multiple times in case participants did not understand how to drive the car.

After the experimental VR phase, participants completed the SSQ once again to measure motion sickness induced by the virtual experience. Additionally, self-report data were collected regarding the source of potential motion sickness, specifically whether it was primarily due to braking or turning, the intensity of the discomfort, and whether participants wished to terminate the experience prematurely.

2.4. Data analysis

Data analyses were performed using JASP 0.18.3 (University of Amsterdam, 2013). Normality assumption was checked by Shapiro-Wilk's test. A paired-sample t-test was used in both rounds of testing for each experimental group to assess differences between

the raw scores of the SSQ pre-testing and the SSQ post testing. In case normality assumption was violated, Wilcoxon signed-rank test was conducted.

3. Results

3.1. Virtual driving experience

In the first round of testing, there was not a statistically significant difference in the SSQ scores in the driving group before and after testing, $t(12) = -1.932, p = .077$ (see Table 1). Nevertheless, 5 participants reported significant to bad motion sickness in the SSQ scoring after testing, and 10 out of 13 participants reported that braking at all traffic lights induced moderate motion sickness, while turning speed was satisfactory. No participants reported that motion sickness was enough to terminate the experience prematurely.

Braking speed was reduced prior to the second round of testing. In this second round, there was no statistically significant difference in the SSQ scores before and after the virtual experience, $t(4) = -2.346, p = .079$, while SSQ scoring after testing suggested bad motion sickness in 3 out of 5 participants (see Table 1). Nevertheless, 4 out of 5 participants reported through self-reported measures that they did not feel motion discomfort, and nobody wanted to terminate the experience prematurely.

3.2. Virtual passenger experience

In the first round of testing, there was a statistically significant difference in the SSQ scores before and after the virtual experience, $t(4) = -3.959, p = .017$, while SSQ scoring after testing suggested negligible motion sickness in most of them (see Table 1). Self-report measures confirmed that 3 out of 5 participants felt motion sickness primarily due to the car's braking and two specific turns. Consequently, despite the SSQ scoring after testing, the braking speed was reduced in the same amount as the virtual driving version, and the speed of the car at the two identified turns was slightly reduced.

The findings from the second round of testing indicated no statistically significant differences in SSQ scores before and after the virtual experience, $z = -1.095, p = .345$, which was also confirmed by SSQ scoring after testing (see Table 1), and participants' self-report who reported no discomfort.

Table 1. SSQ raw score means and standard deviations, and SSQ scoring after testing split by type of VR application tested, round of testing, and SSQ testing phase.

VR application	Round of testing	SSQ phase	N	SSQ mean (sd)	SSQ scoring after testing (N)
Driving experience	First	Pre testing	13	3.70 (2.60)	Negligible (8) Significant (1) Concerning (2) Bad (2)
		Post testing	13	6.15 (5.60)	
	Second	Pre testing	5	1.60 (1.14)	Negligible (2) Bad (3)
		Post testing	5	8.60 (6.58)	
Passenger experience	First	Pre testing	5	2.80 (1.48)	Negligible (4) Bad (1)
		Post testing	5	8.20 (2.95)	
	Second	Pre testing	5	1.20 (.83)	Negligible (4) Bad (1)
		Post testing	5	3.60 5.41)	

4. Discussion

This study presents a novel VR driving simulator that can be useful in research to investigate complex cognitive and behavioral effects while driving. The objective of this paper is twofold: to describe how the VR driving simulator works in its two versions, and to conduct pre-tests regarding the induced motion sickness of the two experiences.

Regarding the design and development of the VR driving simulator, two versions have been developed due to the recent advent of autonomous cars. In addition to a VR

driving experience, a passenger experience was developed, in which users were located in the driver's seat with no need to drive the vehicle. To ensure standardization between participants and across VR conditions, the cars in both driving and passenger experiences followed a pre-set route based on GPS directions. This is particularly relevant in cognitive attentional research, where there is a need to reduce stimuli variability and provide users the similar experiences to allow comparisons. Additionally, this VR driving simulator could aid in investigating cognitive effects on driving while participants are engaged in multiple tasks at the same time, such as driving and talking on the phone, or listening to music on the car radio. Moreover, to target specific cognitive and behavioral effects where it is important for individuals to have a high degree of freedom in driving choices, it is recommended to jointly use VR driving simulators like the one presented here with VR driving simulators resembling commercial adventure/racing games where players can drive as they prefer. This would ensure ecological driving behaviors whilst controlling for stimuli variability.

Finally, regarding pre-tests on motion sickness induced by the two VR experiences, both versions have been improved after two rounds of testing by reducing braking speed and turn speed in the two turns of the passenger experience. These adjustments were made despite the incongruence between SSQ findings and participants' self-reports in both rounds and conditions. Indeed, to reduce the risk of discomfort and motion sickness in future participants, we decided to rely mostly on participants' self-reports on their individual experiences rather than SSQ scores. This decision also stems from the limitations of the SSQ reported in the literature, such as its validation based on data collected from military professionals (who are not representative of the general public) using early 1990s simulators that may not fully align with modern VR systems, the absence of relevant motion symptoms in the questionnaire (e.g., vomiting), inadequate psychometric properties for evaluating cybersickness in VR, and variability in SSQ test-retest reliability [8,9]. The present study joins the substantial portion of literature advocating for precise and reliable questionnaires to assess motion sickness in modern VR systems [10].

5. Conclusion

This study describes two versions of a novel VR driving simulator designed for use in cognitive and psychological research, posing a reduced risk of inducing motion sickness in participants. This paper can assist researchers in developing similar applications.

References

- [1] Clemente M, Rey B, Rodríguez-Pujadas A, Barros-Loscertales A, Baños RM, Botella C, et al. An fMRI study to analyze neural correlates of presence during virtual reality experiences. *Interact Comput*. 2014;26(3):269–84. doi: <https://doi.org/10.1093/iwc/iwt040>
- [2] Markkula G, Romano R, Jamson AH, Pariota L, Bean A, Boer ER. Using driver control models to understand and evaluate behavioral validity of driving simulators. *IEEE Trans Hum Mach Syst*. 2018;48(6):592–603. doi: <https://doi.org/10.1109/THMS.2018.2865666>
- [3] Rendon-Velez E, Van Leeuwen PM, Happee R, Horvath I, van der Vegte WF, de Winter JC. The effects of time pressure on driver performance and physiological activity: a driving simulator study. *Transp Res Part F Traffic Psychol Behav*. 2016;41:150–69. doi: <https://doi.org/10.1016/j.trf.2016.06.013>
- [4] Maxwell H, Weaver B, Gagnon S, Marshall S, Bédard M. The validity of three new driving simulator scenarios: detecting differences in driving performance by difficulty and driver gender and age. *Hum Factors*. 2021;63(8):1449–64. doi: <https://doi.org/10.1177/0018720820977135>
- [5] Ihemedu-Steinke QC, Erbach R, Halady P, Meixner G, Weber M. Virtual reality driving simulator based on head-mounted displays. In: *Automotive User Interfaces: Creating Interactive Experiences in the Car*. Cham: Springer; 2017. p. 401–28. doi: https://doi.org/10.1007/978-3-319-49448-7_17
- [6] Cipresso P, Giglioli IAC, Raya MA, Riva G. The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature. *Front Psychol*. 2018;9:309500. doi: <https://doi.org/10.3389/fpsyg.2018.02086>
- [7] Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *Int J Aviat Psychol*. 1993;3(3):203–20. doi: https://doi.org/10.1207/s15327108ijap0303_3
- [8] Sevinc V, Berkman MI. Psychometric evaluation of simulator sickness questionnaire and its variants as a measure of cybersickness in consumer virtual environments. *Appl Ergon*. 2020;82:102958. doi: <https://doi.org/10.1016/j.apergo.2019.102958>

- [9] Kourtesis P, Linnell J, Amir R, Argelaguet F, MacPherson SE. Cybersickness in virtual reality questionnaire (CSQ-VR): a validation and comparison against SSQ and VRSQ. *Virtual Worlds*. 2023;2(1):16–35. doi: <https://doi.org/10.3390/virtualworlds2010002>
- [10] Davis S, Nesbitt K, Nalivaiko E. A systematic review of cybersickness. In: Proceedings of the 2014 Conference on Interactive Entertainment; 2014 Dec; Newcastle, Australia. p. 1–9. doi: <https://doi.org/10.1145/2677758.2677780>

SECTION IV

ORIGINAL RESEARCH

Health care is one of the areas that could be most dramatically reshaped by these new technologies.

Distributed communication media could become a significant enabler of consumer health initiatives. In fact they provide an increasingly accessible communications channel for a growing segment of the population.

Moreover, in comparison to traditional communication technologies, shared media offer greater interactivity and better tailoring of information to individual needs.

Wiederhold & Riva, 2004

Social Media: Positive or Negative for Health? User Profiles in the Context of Exposure to Body Image-Positive Content on Social Media

Vincent LEMIEUX^{a,1}, Ariane CASSISTA^a and Marie-Pierre GAGNON-GIROUARD^a

^aAPIC (*Équipe de recherche sur les attitudes et les perceptions liées à l'image du corporelle, Department of Psychology, Université du Québec à Trois-Rivières, Québec, Canada*)

ORCID ID: Vincent Lemieux <https://orcid.org/0000-0002-2195-225X>

Abstract. The relationship between social media and body image is contradictory. On one hand, virtual interactions about physical appearance (whether positive or negative) tend to increase body dissatisfaction. On the other hand, viewing images imbued with body diversity, featuring self-compassionate messages or unaltered photos promotes body satisfaction. The relationship one has with his own body image is linked to the reasons why one practice physical exercise, either to modify one's body or to move in a more intuitive way. This study explores the links between social media use, body image and intuitive exercise among 262 participants (86.3% women, with an average age of 41.35 ± 12.89) who were recruited to complete online questionnaires. Ascending hierarchical cluster analyses were used to create different profiles of social media users, which were compared on body satisfaction and intuitive exercise. Group 1 uses social media less often and more briefly. Group 2 and Group 3 use it similarly, but Group 3 is more active, creating posts, compared to Group 2's passive viewing. Group 1 had higher appearance-related body satisfaction ($M = 2.08$, $SD = 0.756$) than Group 3 ($M = 1.823$, $SD = 0.752$) ($p = 0.072$). Group 3 showed higher mindful exercise ($M = 3.679$, $SD = 1.033$) than Group 1 ($M = 3.246$, $SD = 1.066$) ($p = 0.019$). Participants who use social media less often show better appearance-related body satisfaction. For people who use social media more, those who use it in an active way seem to benefit on their exercise habits.

Keywords. Social media, body satisfaction, cluster analyses.

1. Introduction

Social media use increased following the COVID-19 pandemic [1]. In 2022, the number of active social media users in Canada was approximately 87.1% of the total Canadian population. Literature shows that the use of social media that focuses on body image (such as Instagram or TikTok) tends to increase body dissatisfaction in both boys and girls [2]. Whether comments on social media are positive or negative, if they focus on physical appearance, they are associated, in the short term, with higher body dissatisfaction [3]. For some time now, the physical ideal of thinness ("thinspiration") has been gradually supplanted by ideals of beauty that place greater emphasis on muscularity ("fitspiration"). The latter are more socially accepted, as they embody an alternative that seems to promote healthy habits such as an active lifestyle. However, literature shows that women who consume this kind of content on social media are at greater risk of compulsive exercise [4]. Moreover, the literature indicates that the type of content viewed or shared, such as altered photos, selfies, or pages focused on both "thinspiration" and "fitspiration", had a negative association with body image, anxiety, self-esteem and may even affect the development of an eating disorder [5–10].

In parallel, social media is increasingly a place to fight grossophobia (i.e. discriminatory behaviors and attitudes that stigmatize people with a higher weight) [11–

¹ Corresponding Author: Vincent Lemieux, Vincent.lemieux1@uqtr.ca.

12]. Indeed, viewing images imbued with body diversity (and therefore a more humanizing reality), featuring self-compassionate messages or unaltered photos promotes body satisfaction and self-compassion [11]. Whether it's the World Health Organization [13] or the meta-analysis by Stice and colleagues [14], all recommend the implementation of person-and community-centered prevention and health promotion programs in the living environment. Several studies have shown that the use of social media seems to be promising for health promotion and prevention [15–18]. In parallel to the promotion of a positive body image, social media can also be a way to promote intuitive and healthy physical exercise, in contrast to exercise that is performed to modify the body, as in the fitspiration movement. Unlike engaging in a rigid exercise practice, intuitive exercise involves listening to one's body signals and becoming aware of one's senses to know what to do and when to start and stop exercising [19].

Social media are neither positive nor negative in themselves, but it is how one uses them that makes them positive or negative [20]. The scientific literature has highlighted a number of these behaviors, such as the type of use (more passive [i.e., they only watch without interacting with others] or more active [i.e., they interact with others]), the type of content viewed (pages advocating body diversity and mental and physical well-being, pages promoting diet culture, an unattainable body ideal [of the “fitspiration” type], altered photos, self-portraits), frequency of use (i.e., the number of times we go on each social medium each day), duration in minutes spent on each social medium and the number of different social media that are used [5, 21–26].

The main objective of this study is to explore the links between social media use, body satisfaction and intuitive exercise among participants engaged in a study on social media exposure to positive body image content.

2. Methods

2.1. Participants and procedure

Data from this study were collected as part of a larger study with a randomized longitudinal design, which aims to evaluate the impact of online exposure to positive body image content on body image. This study is approved by the Ethics Committee of the Université du Québec à Trois-Rivières (CERPPE-23-25-08-01.03). Participants were recruited via social media advertisements (mainly on Facebook) and completed an anonymous online survey hosted by a secure website (Qualtrics). Before completing the questionnaire, participants electronically signed a consent form. A total of 299 participants completed the questionnaire; 37 participants were did not fully complete the social media questionnaire and were excluded. The remaining 262 participants were mainly women (86.3%), with an average age of 41.35 ± 12.89 . Most participants reported using social media several times daily (87.4%).

2.2. Measures

Sociodemographic data. Participants were asked about general information (e.g., age, sex, gender identity and scolarity) and specific information (e.g., intention with their weight).

Social media use (clustering variable). A home questionnaire was used to measure social media use (frequency, duration and type of use) and perception (whether social media helps or harms body image, etc.).

Body satisfaction. The Body-Esteem Scale [27] was used to assess body satisfaction on a 5-point Likert scale. The questionnaire comprises 23 items and two subscales: appearance and weight. The instrument has excellent reliability and validity ($\alpha = 0.91$ and 0.92). Participants rated their perceptions of body satisfaction attitude and behaviors on a 5-point Likert-type scale.

Intuitive exercise. The Intuitive Exercise Scale is a 14-item instrument designed to measure an individual's relationship with exercise, along four subscales: mindful exercise, emotional exercise, rigidity and body confidence. Participants rated their perceptions of personal exercise behaviors on a 5-point Likert-type scale.

2.3. Statistical Analyses

Statistical analysis was carried out using IBM SPSS Statistics V.29. To create different profiles of social media users, frequency, duration and mode of social media utilization (passive and active) were introduced within an ascending hierarchical cluster analysis (Ward method [28]). This method is an algorithm for grouping two classes of a subset to obtain a more aggregated subset. Once the number of clusters has been obtained using Ward's method, a two-step cluster analysis is performed (i.e., a distance measure to separate the clusters and a probabilistic approach) that enables the selection of the optimal subgroup model [29]. Once the profiles were obtained, an ANOVA ($p \leq 0.05$) was performed to compare the different profiles on body satisfaction and intuitive exercise. Preliminary analyses were performed to ensure that the principles of the ANOVA were respected. Bonferroni post-hoc tests were carried out to identify between-group differences. Considering the very small amount of missing data, no replacement strategy was used.

3. Results

3.1. Clustering analysis

Cluster analysis led to the identification of three subgroups. Participants from Group 1 used social media less often and for shorter periods of time. Group 2 and Group 3 were similar in terms of frequency and duration of social media use, but Group 3 used social media more actively (interacting with others) than Group 2 who played a more passive role (watching content only).

The ANOVA reveals marginally significant differences between clusters on appearance-related body esteem, marginally significant differences on the rigidity in exercise and significant differences on mindful exercise. Post hoc analyses indicate that Group 1 had higher appearance-related body satisfaction than Group 3 ($p = 0.072$, marginally significant) and Group 2 ($p = 0.247$, not significant). Group 1 showed marginally more rigidity in exercise routine than Group 3 ($p = 0.08$) and significantly more mindful exercise than Group 1 ($p = 0.019$).

Table 1. Descriptive statistics and differences between the three groups on main outcomes.

	Group 1 (n = 65)	Group 2 (n = 59)	Group 3 (n = 138)	F (2.252)	p value
Age	M (S.D.) %	M (S.D.) %	M (S.D.) %		
	41.66 (12.09)	40.34 (12.52)	41.64 (13.47)		
Sex					
Men	15.4%	10.2%	13.8%		
Women	84.6%	88.1%	86.2%		
Prefers not to answer	0%	1.7%	0%		
Social media use Frequency					
	2.54 (0.69)	2.93 (0.37)	2.94 (0.24)	22.38	<0.001
Duration	2.40 (0.86)	2.31 (0.62)	3.99 (0.12)	312.53	<0.001
Passive use	2.63 (0.58)	4.25 (0.44)	4.14 (0.78)	131.80	<0.001
Active use	2.02 (0.53)	2.15 (0.52)	2.52 (0.80)	14.22	<0.001
Body Esteem					
Appearance related	2.08 (0.76)	1.84 (0.75)	1.82 (0.76)	2.729	0.067
Weight related	1.79 (0.85)	1.69 (0.94)	1.57 (0.92)	1.39	0.250
Intuitive physical activity					
Mindful exercise	3.25 (1.07)	3.51 (1.02)	3.68 (1.03)	3.83	0.023*
Emotional exercise	2.69 (0.92)	2.33 (1.06)	2.36 (1.11)	2.54	0.081
Rigidity	3.13 (1.10)	2.69 (1.23)	2.73 (1.23)	2.90	0.057
Body confidence	2.57 (0.93)	2.49 (1.16)	2.63 (1.11)	0.33	0.719

Note: Body Esteem = Body-esteem Scale; Intuitive physical activity = Intuitive exercise scale; italic = marginally significant; * = significant; bold = post hoc difference between these groups.

4. Discussion

This study explored the links between social media use, body satisfaction and intuitive exercise among participants in a study focusing on social media exposure to positive body image content. Using cluster analyses on dimensions of social media use (i.e., frequency, duration, passive use and active use), three groups of participants emerged.

Group 1 used social media the least often. However, they seemed to use it longer each time than Group 2. This group used social media the least actively and the least passively. Group 2 and Group 3 used social media almost equally often. On the other hand, Group 2 used social media for the shortest time per use and used them the most passively. Group 3 used social media most often and for the longest time per use. Moreover, this group made the most interactive use of social media, even if they used it more passively than actively.

This study highlights the complexity of the association of social media use with body image. It is possible to observe several differences between the three groups in this study, mainly between Group 1 (i.e., the group that appears to have more favorable use) and Group 3 (i.e., the group that appears to have the least favorable use). Indeed, Group 3 scored lower on body esteem (satisfaction with appearance) than Group 1. This result highlights that the less one uses social media, the higher one's body satisfaction related to appearance satisfaction seems to be. These results go in the same direction as Marques & al. [9], which explains that greater use of social media is associated with greater body dissatisfaction. Moreover, the results demonstrate that Group 3 scored higher than Group 1 on mindfulness in intuitive exercise. These results allow us to hypothesize that people who use social media to interact with others seem to have a healthier relationship with exercise (i.e., they move more mindfully, listening to their body). In contrast, results also seem to show that the use of social media is negatively associated with intuitive physical activity. Indeed, Group 1 is less rigid in their physical activity practice (i.e., they have a more flexible and varied practice) than Group 3. This study has a few limitations. First, a larger sample size would have allowed us to find more significant differences between groups, or vice versa. The second limitation is the representativeness of the sample. As the sample consists almost exclusively of Quebec women in the same age bracket, the results of this study cannot be applied to other parts of the population.

5. Conclusion

This study highlights several important points to bear in mind when it comes to social media. Firstly, the more one uses social media, the lower appearance-related body satisfaction is. However, using social media actively, as opposed to passively, is associated with a more intuitive way of exercising. People who interactively use social media seem to move their bodies in a healthier way by engaging in mindful physical activity.

References

- [1] Kemp S. Digital 2023. Global overview report. [Online]. We Are Social & Meltwater; 2019 May [cited 2024 Jun 20]. Available from: <https://wearesocial.com/wp-content/uploads/2023/03/Digital-2023-GlobalOverview-Report.pdf>.
- [2] Griffiths S, Castle D, Cunningham M, Murray SB, Bastian B, Barlow FK. How does exposure to thinspiration and fitspiration relate to symptom severity among individuals with eating disorders? Evaluation of a proposed model. *Body Image*. 2018;27:187-95.
- [3] Tiggemann M, Barbato I. "You look great!": The effect of viewing appearance-related Instagram comments on women's body image. *Body Image*. 2018;27:61-6.
- [4] Holland G, Tiggemann M. "Strong beats skinny every time": Disordered eating and compulsive exercise in women who post fitspiration on Instagram. *International Journal of Eating Disorders*. 2017;50(1):769.
- [5] Appel M, Marker C, Gnams T. Are Social Media Ruining Our Lives? A Review of Meta-Analytic Evidence. *Review of General Psychology*. 2020;24(1):60-74.
- [6] Cingel DP, Carter MC, Krause H-V. Social media and self-esteem. *Current opinion in psychology*. 2022;45:101304.
- [7] Course-Choi J, Hammond L. Social media use and adolescent well-being: A narrative review of longitudinal studies. *Cyberpsychology, Behavior, and Social Networking*. 2021;24(4):223-36.
- [8] Deighton-Smith N, Bell BT. Objectifying fitness: A content and thematic analysis of #fitspiration images on social media. *Psychology of Popular Media Culture*. 2018;7(4):467-83.

- [9] Marques MD, Paxton SJ, McLean SA, Jarman HK, Sibley CG. A prospective examination of relationships between social media use and body dissatisfaction in a representative sample of adults. *Body Image*. 2022;40:1-11.
- [10] Steinsbekk S, Wichstrøm L, Stenseng F, Nesi J, Hygen BW, Skalická V. The impact of social media use on appearance self-esteem from childhood to adolescence – A 3-wave community study. *Computers in Human Behavior*. 2021;114:106528.
- [11] Folope V. Grossophobia in the care sector, a reality to be fought. *Soins*. 2021;66(858):22-4.
- [12] Tiggemann M, Zinovieva K. The effect of #enhancement-free Instagram images and hashtags on women's body image. *Body Image*. 2019;31:131-8.
- [13] World Health Organization. Health promotion and disease prevention through population-based interventions, including action to address social determinants and health inequity. [cited 2024 Jun 20]. Available from: <https://www.emro.who.int/aboutwho/public-health-functions/health-promotion-disease-prevention.html>.
- [14] Stice E, Marti CN, Shaw H, Rohde P. Meta-analytic review of dissonance-based eating disorder prevention programs: Intervention, participant, and facilitator features that predict larger effects. *Clinical Psychology Review*. 2019;70:91-107.
- [15] García Del Castillo JA, García Del Castillo-López Á, Dias PC, García-Castillo F. Social networks as tools for the prevention and promotion of health among youth. *Psicología: Reflexão e Crítica*. 2020;33(1).
- [16] Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. The future of mental health care: peer-to-peer support and social media. *Epidemiology and Psychiatric Sciences*. 2016;25(2):113-22.
- [17] O'Reilly M, Dogra N, Hughes J, Reilly P, George R, Whiteman N. Potential of social media in promoting mental health in adolescents. *Health Promotion International*. 2019;34(5):981-91.
- [18] Paxton SJ, McLean SA, Rodgers RF. "My critical filter buffers your app filter": Social media literacy as a protective factor for body image. *Body Image*. 2022;40:158-64.
- [19] Reel J, Miyairi M. The right 'dose' of activity: Health educators should promote mindful and intuitive exercise. *Community Med Health Educ*. 2012;2:1000-111.
- [20] Kross E, Verdun P, Sheppes G, Costello CK, Jonides J, Ybarra O. Social Media and Well-Being: Pitfalls, Progress, and Next Steps. *Trends in Cognitive Sciences*. 2021;25(1):55–66.
- [21] Lemieux V, Coulombe AS. Les médias sociaux et les algorithmes, nos amis? In: Dontigny F, editor. *Apprécier son corps à sa juste valeur*. Montréal: JFD éditions; 2024. p. 65-79.
- [22] Cohen R, Newton-John T, Slater A. The case for body positivity on social media: Perspectives on current advances and future directions. *Journal of Health Psychology*. 2021;26(13):2365-73.
- [23] Maheux AJ, Nesi J, Galla BM, Roberts SR, Choukas-Bradley S. #Grateful: Longitudinal Associations Between Adolescents' Social Media Use and Gratitude During the COVID-19 Pandemic. *Journal of Research on Adolescence*. 2021;31(3):734-47.
- [24] Matzen PL. The impact of social media influencers on male body image: How influencers shape male body image. 2020.
- [25] Valkenburg PM. Social media use and well-being: What we know and what we need to know. *Current Opinion in Psychology*. 2022;45:101294.
- [26] Valkenburg PM, Van Driel II, Beyens I. The associations of active and passive social media use with well-being: A critical scoping review. *New Media & Society*. 2022;24(2):530-49.
- [27] Mendelson BK, Mendelson MJ, White DR. Body-Esteem Scale for Adolescents and Adults. *Journal of Personality Assessment*. 2001;76(1):90–106.
- [28] Ward JH, Hook ME. Application of an Hierarchical Grouping Procedure to a Problem of Grouping Profiles. *Educational and Psychological Measurement*. 1963;23(1):69–81.
- [29] Kent P, Jensen RK, Kongsted A. A comparison of three clustering methods for finding subgroups in MRI, SMS or clinical data: SPSS TwoStep Cluster analysis, Latent Gold and SNOB. *BMC Medical Research Methodology*. 2014;14(1):113.

Human Body Odors Modulation on Affective Processing of Social-Emotional Virtual Environments

Luna MADDALON^{a,1}, Sergio CERVERA-TORRES^a, Maria Eleonora MINISSI^a, Alberto GRECO^b, Luca CITI^c, Javier MARIN-MORALES^a, Mariano ALCAÑIZ^a

^a*Human-Centered Technology Research Institute (HUMAN-Tech). Polytechnic University of Valencia, Valencia, Spain*

^b*Research Center “E. Piaggio”, University of Pisa, Pisa, Italy*

^c*School of Computer Science and Electronic Engineering, University of Essex, Colchester, UK*

ORCID ID: Luna MADDALON <https://orcid.org/0000-0002-2789-6703>

Abstract. The link between the olfactory and affective systems in humans has been explored in clinical and non-clinical populations. Specifically, there is evidence indicating that both artificial scents and human body odors (HBOs) can influence individual affective processing. Research on HBOs has primarily concentrated on the appraisal of emotional faces, while their effects on other social stimuli are still being investigated. In particular, social-emotional situations can evoke various reactions in humans, and it is essential to examine HBOs' role in their affective processing. Virtual reality facilitates this research by allowing the reproduction of standardized and realistic social-emotional situations. In this study, we investigated whether exposure to emotion-related HBOs (happy, fear, and neutral state) modulates the affective processing of social-emotional situations compared to the absence of exposure. Three social-emotional virtual environments (VEs) designed to elicit positive, negative, and neutral affectivity were tested. During the VE experiences, the experimental group was exposed to HBOs ($n=77$), while the control group experienced them without HBOs ($n=72$). Participants were required to rate individual valence and arousal after each social-emotional VE. Findings revealed that HBO exposure overall resulted in lower arousal ratings and higher valence ratings than the control group. The difference in valence rating was particularly evident in the negative VE. This demonstrates the modulating effect of HBOs on the affective processing of social-emotional VEs. Future directions include investigating differences across emotion-related HBOs in the affective processing of social-emotional situations.

Keywords. Human body odors; virtual environment; emotion; social stimulation; affective processing

1. Introduction

An expanding corpus of research underscores the influential role of olfactory stimuli in shaping the dynamics and outcomes of social interactions among humans. Investigations have delved into the intricate mechanisms through which human behavioral responses within social-emotional contexts are modulated by both synthetic [1] and naturally occurring odors [2]. This body of literature elucidates the interplay between olfactory cues and social behavior, shedding light on how scent perception modulates human social dynamics and decision-making processes. By directing attention toward human body odors (HBOs), individuals can convey socially pertinent information, including indications of cooperation, aggression, and competition, through chemosensory signals [3, 4]. For example, a study suggests that HBOs associated with fear can induce heightened anxiety among recipients by intricately modulating their negative perceptions of social stimuli experienced within immersive virtual reality environments [5]. Given the olfactory system's capacity to exert influence over individual affective processing, HBOs emerge as a pivotal point within the realm of social neuroscience, as they can convey numerous social cues. Thus far, attention has predominantly been directed toward investigating emotion elicitation based on socially pertinent stimuli such as facial

¹ Corresponding Author: Luna Maddalon, lmaddal@htech.upv.es

images and videos [4]. However, little has been studied on how the HBOs modulate affective processing in relation to other social-emotional stimuli. Within this lacuna, the virtual reality methodology assumes prominence, offering a means to replicate social-emotional scenarios within controlled environments that afford participants a heightened sense of presence and immersion.

Considering the above, we have developed social-emotional virtual environments (VEs) that seek to elicit positive, negative, or neutral affectivity [6, 7]. This study aims to investigate the potential influence of HBOs on affective processing within social-emotional VEs, with emotional self-reports as the primary assessment metric.

2. Methods

2.1. Participants

The study included 149 adults aged 18 to 55 years in two groups. Participants took part in the study in exchange for monetary rewards. Participants were randomly split into two groups: the experimental group exposed to HBOs ($n=77$) and the control group not exposed to HBOs ($n=72$). Participants belonging to the experimental group were 36 males and 41 females ($M = 34.02$; $SD = 10.89$). Participants belonging to the control group were 32 males and 40 females ($M = 33.68$; $SD = 9.38$).

Participants were Spanish-speaking, not pregnant, non-smokers, and were not taking any psychiatric medication. Before participating in the study, informed written consent was given. The study has been approved by the Ethical Committee of the Polytechnic University of Valencia (P2_18_06_19).

2.2. Human body odors

The HBOs used for the present study were previously sampled through donors' armpit sweat at the Institute of Applied Psychology (Lisbon, Portugal). Specifically, pads were applied directly to donors' bodies to soak up the sweat they released while watching emotional-laden videos. The exposure to these videos was part of an emotional induction procedure intended to elicit emotional states of happiness, fear, and neutrality, respectively. The induction of these emotional states was analyzed by filling out self-report questionnaires, which assessed their levels of anger, fear, happiness, sadness, disgust, neutrality, surprise, calmness, and amusement using 7-point Likert scales, ranging from 1, indicating "not at all," to 7, indicating "very much." After that, HBOs were created by combining different pads related to the same emotion and stored in a -80°C freezer. Previous studies have been conducted on the sweat sample collection and conservation method [8].

During the odor exposure procedure, HBOs were released through a purposely developed 3-channel olfactometer (with airflow distributed through nose-positioned vials), which dispensed odors from pads at a controlled and gradual flow rate of 50ml/min, which was also used in previous studies [8]. The olfactometer was connected to an air compressor. Clean air was proportioned for a few seconds between odor administration to remove residual odor remaining in the machine.

2.3. Virtual environments

The VEs in this study were specifically adapted for a semi-immersive CAVE system. Concretely, three social-emotional VEs were modeled with Unity 3D® software to convey positive, negative, or neutral affectivity. The VEs had the same layout with a stable number of park elements and virtual agents between park conditions. They consisted of three different "emotional" parks trying to convey affectivity through visual and acoustic stimuli (see Figure 1). Each emotional VE lasted 60 seconds, and social stimulation was presented for 30 seconds. This social stimulation used virtual agents to

convey social meaning. The virtual agents were pre-programmed avatars with human appearances. Further details regarding the VEs are specified in prior publications [6, 7].



Figure 1: Positive VE. a. non-social; b. social

2.4. Procedure

Within a broader experimental methodology, useful data were collected for this study. Before the experimental testing, the Liebowitz Social Anxiety Scale (LSAS) and the Beck Depression Inventory (BDI) were administered to rule out the risk for depression and social anxiety in both groups. Moreover, the olfactory sensitivity of participants in the experimental group was screened via the Discrimination and Thresholds subtests of the Sniffin' Sticks Test [9]. Afterward, participants were individually exposed to the three VEs in a counterbalanced order. Only the experimental group ($n=77$) was exposed to three fear, happiness, and neutral HBOs while watching the three VEs. HBOs were administrated in a pseudorandomized order across participants. After experiencing each VE, participants were asked to self-report two canonical dimensions of affect, emotional valence and arousal, on a visual 9-point Likert scale. These measures were taken from the Self-Assessment Manikin [10]. In the valence scale, the "0" pole represented the highest negativity, and the "9" pole represented the highest positivity. In the arousal scale, "0" represented minimal activation, and "9" represented maximum activation.

2.5. Data analysis

Data analyses were performed using IBM SPSS Statistics v25. A generalized linear mixed model (GLMM) approach was used to compare self-reported valence and arousal responses. Analyses specified a Gaussian distribution with a log-link function and a robust standard estimator. After running the Kolmogorov-Smirnov test, a log-link function was chosen to be applied, which is suitable for modeling data that deviate from normality as the current data. The subjects' ID was included as a random factor with random intercepts. The analysis accounted for VE and group (HBO, non-HBO) as fixed factors for valence and arousal reports ($1 + \text{VE} + \text{GROUP} + \text{VE} * \text{GROUP} + (1|\text{subject})$). Post hoc analyses were run using Bonferroni correction.

3. Results

Table 1: Means and standard deviations of valence and arousal reported by groups in the three emotional VEs.

Affect dimension	VE	GROUP	N	Mean	SD
Valence	Positive	HBO	77	7.299	1.940
		non-HBO	72	7.111	1.968
		HBO	77	4.130	2.041
	Negative	non-HBO	72	3.264	2.295
		HBO	77	6.026	1.899
	Neutral	non-HBO	72	5.570	1.774
		HBO	77	3.312	2.358
	Arousal	non-HBO	72	4.292	2.417
		HBO	77	4.909	2.202
		non-HBO	72	5.597	2.448
Arousal	Positive	HBO	77	2.857	1.890
		non-HBO	72	3.194	1.933

Table 1 presents the means and standard deviations of valence and arousal as reported by the groups across the three emotional VEs.

Results showed significant main effects of VE in valence, $F(2, 443) = 225.84, p < .001$ and arousal, $F(2, 443) = 149.78, p < .001$. In particular, participants' self-reported valence matched the emotional VE types, which was interpreted as a successful manipulation check. Consequently, on average, positive VEs received higher valence ratings compared to the others (negative – neutral: $MD = .603; SE = .0334, p < .001$; negative – positive: $MD = .486; SE = .0259, p < .001$; positive – neutral: $MD = 1.242; SE = .0564, p < .001$). On the other hand, on average, participants reported greater arousal in the negative VE, with the positive and neutral VEs following in that order (negative – neutral: $MD = 1.75; SE = .1029, p < .001$; negative – positive: $MD = 1.42; SE = .0783, p < .001$; positive – neutral: $MD = 1.23; SE = .0779, p = .003$).

There was also a significant effect of the main group factor on valence, $F(1, 147) = 10.61, p = .001$, and arousal, $F(1, 147) = 6.75, p = .009$. Concretely, on average, the HBOs group reported higher valence than the non-HBOs group ($MD = .891; SE = .0375, p = .001$). The HBOs group also showed, on average, lower arousal ratings than the non-HBOs group ($MD = 1.19; SE = .0837, p = .009$).

Interestingly, there was also a significant interaction between groups and VE, $F(2, 294) = 8.74, p = .013$, indicating emotional valence differences in the negative VE. Indeed, the post hoc comparison showed that the group exposed to HBOs during the negative VE reported a higher positive valence than the non-HBOs group ($MD = .739; SE = .0661, p = .011$). The group * VE interaction regarding arousal did not yield statistical significance, $F(2, 294) = 1.48, p = .47$.

4. Discussion

Regardless of group type, the results demonstrate a direct self-reported valence modulation depending on the type of the social-emotional VE and a consistent arousal modulation in the negative and neutral VEs (but inconsistent in the positive VE). This effect was also found in a similar study that did not implement the HBOs [7]. This is consistent with affect attribution theories [11], demonstrating the activation of corresponding emotional responses upon exposure to emotional stimuli. It is in line with findings from affect research indicating a tendency towards negativity bias, where negative stimuli typically evoke greater perceived arousal levels according to self-reports [11]. Furthermore, regardless of social-emotional VEs, findings revealed that exposure to HBOs resulted in lower arousal ratings and higher valence ratings compared to the control group, elucidating the modulating effect of HBOs on the affective processing of social-emotional VEs. This observation aligns with broader studies on HBOs and emotional affective processing, extending beyond the sphere of VEs [12]. Thus, HBOs may mitigate arousal activation by promoting positive responses. Finally, the results indicate that the group exposed to HBOs during the negative VE rated a greater positive valence than the control group. It has been posited that exposure to emotion-related HBOs promotes the illusory presence of others during decision-making [13]. This might indicate that HBOs could activate neural circuits associated with positive social experiences or memories of pleasant interactions, even in negative contexts, favouring a more positive evaluation. This could occur through several mechanisms, such as the unconscious association of odor with positive past experiences or their impact on the limbic system, which regulates emotions and affective responses [1]. A further possible interpretation might be that exposure to HBOs could influence the perception of social-emotional VEs through associative conditioning, attenuating the perception of negativity and inducing greater attention to positive features of the context. Indeed, individuals might unconsciously associate HBOs with rewarding social experiences or feelings of safety and familiarity. As a result, during exposure to HBOs, positive connections might form between odors and social experiences, leading to a more positive evaluation of VE. In alignment with this notion, empirical findings suggest that HBOs influence social-emotional perception, supporting the idea that exposure to HBOs may affect the perception of VEs through associative learning and emotional responses. For example, a study supports that HBOs, a meaningful and rewarding social signal, can influence social-emotional perception through positive associative learning between HBOs and

social experiences [14]. Another study showed that fear-related HBO heightened anxiety and reduced trust towards a virtual character, indicating that their impact may influence social information processing and impression formation in VEs [5].

Based on these findings, it is recommended that future research investigate variations in emotion-related HBOs and their impact on the affective processing of social-emotional scenarios. Namely, investigating which type of HBO influences a certain social-emotional situation and why. In general, future studies should explore the influence of HBOs on VEs, employing objective methodologies such as electroencephalography or galvanic skin response. Furthermore, examining potential discrepancies in HBO effects within social and non-social contexts would be pertinent. This study's limitations could encompass subjectivity bias in self-reports, wherein participants may demonstrate response bias or social desirability bias, thereby impacting their responses and potentially distorting the results. Consequently, HBO effects are expected to be more evident when employing implicit measures, such as psychophysiological activity.

5. Conclusion

The extensive research on HBOs underscores their significant role in shaping human social interactions and behavioral responses in general [2-5]. With HBOs linked in social neuroscience, virtual reality methodologies are suggested for understanding how these signals interact with social-emotional stimuli, suggesting their potential to induce affective responses. This work aimed to evaluate how individuals process social-emotional VEs when exposed to non-specific human HBOs. This was accomplished by analyzing self-reported responses on the degree of perceived valence and arousal in each experimental condition.

In synthesis, the social-emotional VEs were found effective in eliciting positive, negative, and neutral affections, as measured by self-reports. In addition, despite the social- emotional VEs, exposure to HBO led to reduced arousal ratings compared to the control group. However, HBOs increased self-reported valence, mitigating, in particular, the negative responses to negative VEs and revealing the role of odors in modulating emotions and affective responses.

References

- [1] Leppänen JM, Hietanen YJK. Affect and face perception: odors modulate the recognition advantage of happy faces. *Emotion*. 2003;3(4):315–26. doi: <https://doi.org/10.1037/1528-3542.3.4.315>
- [2] de Groot JH, Smeets MA, Kaldewij A, Duijndam MJ, Semin GR. Chemo-signals communicate human emotions. *Psychol Sci*. 2012;23(11):1417–24. doi: <https://doi.org/10.1177/0956797612445317>
- [3] Banner A, Frumin I, Shamay-Tsoory SG. Androstadienone, a chemosignal found in human sweat, increases individualistic behavior and decreases cooperative responses in men. *Chem Senses*. 2018;43(3):189–96. doi: <https://doi.org/10.1093/chemse/bjy002>
- [4] Mutic S, Freiherr J, Cavazzana A, Rocha M, Soares SC, Parma V. The scent of the other women: body odor-induced behavioral and physiological effects on face categorization. *Physiol Behav*. 2019;210:112562. doi: <https://doi.org/10.1016/j.physbeh.2019.112562>
- [5] Quintana P, Nolet K, Baus O, Bouchard YS. The effect of exposure to fear-related body odorants on anxiety and interpersonal trust toward a virtual character. *Chem Senses*. 2019;44(9):683–92. doi: <https://doi.org/10.1093/chemse/bjz063>
- [6] Cervera-Torres S, Minissi ME, Greco A, Callara A, Ferdowsi S, Citi L, et al. Modulating virtual affective elicitation by human body odors: advancing research on social signal processing in virtual reality. In: Proceedings of the International Conference on Human-Computer Interaction; 2023; Copenhagen. Cham: Springer Nature Switzerland; 2023. p. 317–27. doi: https://doi.org/10.1007/978-3-031-35017-7_20
- [7] Cervera-Torres S, Gracia Laso DI, Minissi ME, Maddalon L, Chicchi Giglioli IA, Alcañiz M. Social signal processing in affective virtual reality: human-shaped agents increase electrodermal activity in an elicited negative environment. *Cyberpsychol Behav Soc Netw*. 2024;27(4):268–74. doi: <https://doi.org/10.1089/cyber.2023.0273>
- [8] Callara AL, Cecchetto C, Dal Bò E, Citi L, Gentili C, Vanello N, et al. Human body odors of happiness and fear modulate the late positive potential component during neutral face processing: a preliminary ERP study on healthy subjects. In: Proceedings of the 44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2022; Glasgow. IEEE; 2022. p. 4093–6. doi: <https://doi.org/10.1109/EMBC48229.2022.9871495>
- [9] Wolfensberger M. Sniffin' Sticks: a new olfactory test battery. *Acta Otolaryngol*. 2000;120(2):303–6. doi: <https://doi.org/10.1080/000164800750001134>

- [10] Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry*. 1994;25(1):49–59. doi: [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
- [11] Cacioppo JT, Berntson GG. The affect system: architecture and operating characteristics. *Curr Dir Psychol Sci*. 1999;8(5):133–7. doi: <https://doi.org/10.1111/1467-8721.00031>
- [12] Cecchetto C, Fischmeister FPS, Gorkiewicz S, Schuchly W, Bagga D, Parma V, et al. Human body odor increases familiarity for faces during encoding-retrieval task. *Hum Brain Mapp*. 2020;41(7):1904–19. doi: <https://doi.org/10.1002/hbm.24920>
- [13] Alcañiz M, Chicchi Giglioli IA, Carrasco-Ribelles LA, Minissi ME, López CG, Semin GR. How priming with body odors affects decision speeds in consumer behavior. *Sci Rep*. 2023;13(1):609. doi: <https://doi.org/10.1038/s41598-023-27643-y>
- [14] Lübke KT, Croy I, Hoenen M, Gerber J, Pause BM, Hummel T. Does human body odor represent a significant and rewarding social signal to individuals high in social openness? *PLoS One*. 2014;9(4):e94314. doi: <https://doi.org/10.1371/journal.pone.0094314>

Machine Learning Classification Using Emotion Language and Subjective Difficulty to Distinguish Patients with Functional Seizures from Trauma Controls

Natalie N. NEWTON^{a,1}, Nicole A. ROBERTS^a, Estrella M. CONTRERAS^a and Mary H. BURLESON^a

^a*School of social and Behavioral Sciences, Arizona State University, Tempe, Arizona, USA.*

ORCID ID: Natalie N. Newton <https://orcid.org/0000-0002-3905-4423>

Abstract. Functional seizures (FS) is a clinical condition where individuals experience seizure-like symptoms without the expected electrocortical basis for them. Those with FS tend to experience difficulty with negative emotions and have trouble labeling or describing their emotions. We used collected data from a relived emotion task in which FS and a trauma control group (TC) wrote descriptions when prompted to recall memories evoking anger and shame feelings and subsequently rated subjective difficulty for the reliving task. We conducted a sentiment analysis by matching words from descriptions to words in the Affective Norms for English Words (ANEW) database to compute language features of word count, valence, and arousal. Using language features and subjective task difficulty, we tested 45 machine learning models with a logistic regression classification engine to test which features worked best to distinguish FS from TC. For models with language features only, features of the relived shame condition were more accurate at distinguishing FS from TC; models with word count and with valence and/or arousal added were more accurate than with word count alone. Models with language features and difficulty were better at distinguishing FS from TC than models with language features alone. However, the two models with subjective difficulty features alone emerged as the most accurate to distinguish FS from TC. This serves as a valuable first step at demonstrating the utility of language features in combination with subjective ratings at discerning FS from TC, although more work is needed to bolster contributions of language features to this type of ML modeling.

Keywords. Functional seizures, sentiment analysis, machine learning, emotion

1. Introduction

Functional seizures (FS), or psychogenic non-epileptic seizures, is a clinical condition in which patients experience seizure-like symptoms without the expected aberrant electrocortical activity during episodes [1-3]. FS are diagnosed in under 1% of the general population, but FS patients make up ~20% of those seeking epilepsy treatment [4]. FS patients do not respond to anti-seizure medication and often have comorbid psychological conditions [5], making FS difficult to treat because the condition is primarily psychological [2, 3]. Nevertheless, FS is just as debilitating as epileptic seizures [6].

While the cause of FS is unclear, prior work suggests that FS patients tend to struggle with describing and processing emotions [7]. These difficulties, along with possible disruptions in interoception, agency, and greater “fight-or-flight” tendencies, may translate into abnormal motor movement [2, 8-10]. FS patients also tend to exhibit alexithymia (difficulty labeling/describing emotions), report more difficulty with negative emotions, and show more FS symptoms from exposure to negative stimuli [11-12].

¹ Corresponding Author: Natalie N. Newton, nnnewton@asu.edu

13]. Emotional processing issues and alexithymic tendencies make FS language use of particular interest.

A handful of studies have explored language and narratives produced by FS patients.

FS had more trouble recalling details during seizure episodes compared to epilepsy patients [14]. A qualitative study compared narratives produced by FS and epilepsy patients; FS tended to produce shorter, more negative narratives expressing powerlessness primarily stemming from their FS diagnosis [15, 16]. A smaller study found that FS patients commonly disclosed traumatic events but did not recognize them as such [17].

No study to our knowledge has leveraged natural language processing (NLP) to examine language produced by those with FS compared to a clinical control group on their past autobiographical or relived emotional memories. Language features can be used to differentiate normative populations from those with depressive symptoms [18] or those with posttraumatic stress disorder [19], demonstrating the utility of language among those with clinical conditions. Moreover, language use can shape emotional experiences [20, 21].

We investigated systematic differences in language use by FS compared to trauma controls (TC) when prompted to describe memories evoking anger and shame, respectively. We identified language features based on theoretical and empirical grounds to then use in supervised machine learning models as a means of classifying the two groups (FS vs. TC). We also added to the model's participant ratings of subjective (self-reported) difficulty in reliving negative emotions, as these demonstrated a strong group difference effect in our prior work [12]. We hypothesized that 1) models using language features and/or subjective ratings derived from the shame condition would be more accurate than models using only anger features. Additionally, we hypothesized that 2) language features, including word count and two defining characteristics of affective states, namely valence and arousal [25], would yield more accurate models to differentiate FS from TC than models with word count alone, and 3) models with subjective difficulty added as a feature would also yield greater accuracy compared to models with language features alone.

2. Methods

We examined language features and subjective difficulty ratings for a relived emotion task among FS and TC using supervised machine learning (ML) models. We used data collected from participants ($N=60$) with FS ($N=11$) or TC with varying levels of posttraumatic stress (PTS; $N=49$) from a larger study where they completed a relived emotion task [12, 22].

After a *neutral* condition, participants were asked to recall a time they felt strong amounts of *anger*, *shame*, or *happiness* (counterbalanced). For each condition, participants were instructed to write descriptions (~4 sentences) of the memory then verbalize it and rate the subjective difficulty of reliving the memory. Here we focused on written descriptions evoked from *anger* and *shame* conditions and the accompanying subjective difficulty.

2.1. Text and data analysis

We used an NLP technique to identify individual words, irrespective of order, used in written descriptions. Using the Tidytext package in R [23], descriptions were preprocessed by lemmatizing (e.g., *run*, *ran*, *running* all become *run*), removing punctuation and numbers, converting all words to lowercase, and extracting individual words used as tokens (the smallest unit of analysis). We did not account for polysemic words with multiple meanings, negation (e.g., “today was *not* great”), misspelled words, related emotion words (e.g., *monotony* vs. *monotonous*), or initialisms (e.g., *DUI*).

We conducted a sentiment analysis by matching words from descriptions to words found in the Affective Norms for English Words (ANEW) database, comprising 13,000+ words with crowdsourced valence and arousal ratings [24]. Emotions can be classified by valence (how positive or negative the emotion is) and arousal (how physiologically arousing or calming it is) dimensions [25]; single words can also be classified this way [24].

Using only description words found in ANEW, we computed word count, mean valence, and mean arousal per participant per emotion condition of *anger* and *shame*. Using 2 (emotion condition: *shame*, *anger*) x 2 (group: FS, TC) mixed measures ANOVA in SPSS with emotion condition as a repeated measure and group as a between-subject factor, we examined differences in word count, valence, arousal, and difficulty. Of note, 3 FS participants who did not complete the relived emotion task for either *anger* or *shame* did not provide difficulty ratings, and thus were excluded from comparisons of difficulty ratings and from subsequent ML analyses. For the machine learning (ML) classification models, we split data into training and test sets using a proportion split of .65 for the training data [26]. Because the proportion of FS vs. TC was imbalanced, we used stratification to ensure an even proportion of FS was in the training and test data.

Using supervised ML with a binary logistic regression classification engine with 5-fold cross-validation in R with the tidyModels package [27], we trained models and evaluated performance on test data using the area under the curve (AUC) and balanced accuracy (*BA*) values (calculated with the R caret package) [28, 29]. We tested 45 models with features of word count, valence, arousal, and difficulty ratings for *anger* and *shame* conditions individually and in combination (i.e., *anger + shame*).

3. Results

ANOVA results are in **Table 1**. Trend-level effects of emotional condition emerged for valence and word count. A significant group effect also emerged for valence and word count, with FS using fewer and more negative words than TC. For subjective ratings, there was a significant group effect, with FS reporting greater difficulty than TC. No other significant effects emerged.

3.1. Machine learning: Language features

Models that included only language features (word count, valence, and/or arousal) did not exceed a test data AUC of .80, with only some models achieving applicable test data *BA* values due to low sensitivity in identifying FS (see Table 2) [29]. Models with *BA*=.89 or greater emerged as those with the highest predictive accuracy at identifying FS. The model using word count and arousal for the *shame* condition (2 features) performed the best given its high *BA* and acceptable AUC (*BA*=.89, AUC=.72). The model with word count and arousal for the two conditions combined (*anger + shame*; 4 features) had higher sensitivity but poor AUC (*BA*=.92, AUC=.59). Although several models achieved an acceptable AUC of <.70, many of these models had poor sensitivity and low or not applicable *BA* [29, 30].

The other high performing models had equally high *BA* but lower AUC of <.70. The model with word count and valence for *anger* (2 features), valence and arousal for *shame* (2 features), and word count, valence, and arousal for *shame* (3 features) all performed the same (*BA*=.89, AUC=.68). Models with language features of word count, valence, or arousal in isolation (1 feature) for any of the emotion conditions were unable to identify FS participants. Since most higher performing models included features for *shame*, this in part supports our first hypothesis. Our second hypothesis also received support, given that valence and/or arousal language features improved model accuracy versus word count alone; however, neither valence nor arousal alone were adequate models.

Table 1. Comparisons of Valence, Arousal, Word Count, and Difficulty Ratings by Emotion Condition and Group

Outcome Measure	df	F-value	p-value	η^2
Valence				
Emotion	1, 58	3.73	.058	.06
Group	1, 58	4.37	.041	.07
Emotion x Group	1, 58	2.12	.151	.04
Arousal				
Emotion	1, 58	0.80	.375	.01
Group	1, 58	0.13	.723	>.01
Emotion x Group	1, 58	0.82	.369	.01
Word Count				
Emotion	1, 58	2.89	.095	.05
Group	1, 58	11.67	.001	.17
Emotion x Group	1, 58	2.21	.143	.04
Difficulty Ratings				
Emotion	1,55	0.06	.801	.00
Group	1,55	15.52	.001	.22
Emotion x Group	1,55	0.79	.378	.01

Note: df = degrees of freedom; η^2 = effect size measured as *partial eta squared*.

3.2. Machine learning: Language features and subjective difficulty ratings

By adding rated difficulty of reliving emotional experiences to models with language features, predictive accuracy improved for most models, supporting our third hypothesis. Models with $BA=.89$ or greater and $AUC=.70$ emerged with the highest predictive accuracy at distinguishing FS from TC (**Table 2**). The model with difficulty ratings, word count, and arousal for *anger* (3 features) had the highest performance ($BA=.92$, $AUC=.73$). This was closely followed by a model with difficulty, valence, and arousal for *shame* (3 features; $BA=.89$, $AUC=.76$), as well as the model with difficulty and valence for *anger* (2 features; $BA=.92$, $AUC=.70$) and a model with difficulty, word count, valence, and arousal for *anger* (4 features; $BA=.92$, $AUC=.70$). Since most of the highest performing models did not include *shame* features, our first hypothesis was not supported when difficulty ratings were added to language features. Several other difficulty and language feature models fared well at identifying those with FS ($BA=.89$) but had lower AUC values ($AUC<.70$; see **Table 2**).

Using only subjective difficulty yielded the best performing models overall, outperforming those with language features. The model with difficulty for *anger + shame* (2 features) performed the best at distinguishing FS from TC ($BA=.94$, $AUC=.84$). The model with difficulty for *anger* (1 feature) performed similarly ($BA=.94$, $AUC=.83$).

Although the model with difficulty for *shame* had a comparable AUC value, its sensitivity at identifying FS was too low ($BA=NA$, $AUC=.81$). Thus, our first hypothesis (i.e., that *shame* models would outperform *anger* models) was not supported, and our third hypothesis was partially supported, insofar as subjective difficulty, albeit without language features, yielded the two best ML models. Given the small sample, models with 2 features or fewer are more likely to generalize to a larger sample than models with 4 features. The most accurate models generally overfit training data, which also warrants caution.

Table 2. Machine learning models results

Model features	Emotion Condition									
	Anger			Shame			Anger + Shame			
	AUC Training	AUC Test	BA	AUC Training	AUC Test	BA	AUC Training	AUC Test	BA	
Word count	0.81	0.76	NA	0.92	0.76	NA	0.91	0.76	NA	
Valence	0.73	0.52	NA	0.72	0.79	NA	0.75	0.68	NA	
Arousal	0.79	0.39	NA	0.68	0.52	NA	0.93	0.52	NA	
Word count, valence	0.85	0.68	0.89	0.93	0.79	NA	1.00	0.50	NA	
Word count, arousal	0.88	0.57	NA	0.91	0.72	0.89	0.99	0.59	0.92	
Valence, arousal	0.82	0.39	NA	0.68	0.68	0.89	0.93	0.47	0.55	
Word count, valence,	0.88	0.61	NA	0.91	0.68	0.89	1.00	0.63	0.75	
Difficulty	0.94	0.83	0.94	0.58	0.81	NA	0.94	0.84	0.94	
Difficulty, word count	0.95	0.89	0.81	0.93	0.72	NA	0.98	0.84	0.37	
Difficulty, valence	1.00	0.7	0.92	0.75	0.89	NA	1.00	0.77	0.75	
Difficulty, arousal	0.98	0.63	0.92	0.72	0.63	0.89	1.00	0.63	0.89	
Difficulty, word count,	1.00	0.68	0.92	0.73	0.81	NA	1.00	0.67	0.89	
Difficulty, word count,	0.99	0.73	0.92	0.92	0.67	0.89	1.00	0.59	0.55	
Difficulty, valence, arousal	1.00	0.60	0.89	0.73	0.76	0.89	1.00	0.73	0.75	
Difficulty, word count,	1.00	0.70	0.92	0.91	0.64	NA	1.00	0.6	0.64	

Note: AUC = Area under the curve; BA = test data balanced accuracy.

4. Discussion

FS have gained increased empirical attention, given the need to better understand, diagnose and treat this condition. The ML models presented here serve as preliminary evidence that language features in combination with subjective difficulty experiencing emotions can distinguish FS from a comparison group with prior trauma exposure and varying clinical symptom levels. Because those with FS tend to exhibit alexithymia [2, 8] and struggle with emotional processing difficulties for negative emotions [12], we expected language features, particularly emotion words reflecting valence and arousal, in addition to overall word count, to be consequential in ML models. Although language was not more relevant than subjective difficulty ratings in differentiating FS and TC, both sets of features combined – and even subjective task difficulty alone – were predictive

We examined features in the context of *anger* and *shame*, which are theoretically relevant to FS and to traumatic stress conditions. Since prior work has revealed that FS struggle with shame in particular [8, 12], we hypothesized that language features for *shame* would lead to more accurate ML models than those using only *anger* features. This was largely supported, since models with *shame* language features fared better than those using *anger* features. However, this changed when adding subjective difficulty to models with language features. Contrary to our first hypothesis, models using *shame* language and difficulty features were not among the top performing models when difficulty was added. In fact, the two models with subjective difficulty alone for *anger* and *anger + shame* were the best performing models tested, as they were able to distinguish FS from non-FS participants. This suggests that shame may be relevant in how those with FS communicate past events and emotions, whereas reported difficulty reliving anger, in addition to written expressions of anger-inducing experiences, may be relevant to characterizing FS versus TC.

Models with only 2 language features also performed well at identifying FS. While valence and arousal performed adequately combined and with word count, each performed poorly individually. Thus, awareness of both amount and content of text used may be relevant to clinicians and researchers aiming to differentiate those with and without FS. While several of the top performing models had acceptable or excellent AUC

values, all models should ideally have higher sensitivity in identifying FS [30]. High AUC along with high sensitivity and high *BA* are necessary [29].

The results should be interpreted with caution given several limitations. First, as noted earlier, polysemic words, proper nouns, initialisms, negation, misspelled words, and word discrepancies were not accounted for in text preprocessing. Possible cultural, individual stylistic, or other context-based differences in language use, including potential linguistic biases in how words are normed, also were not considered here. Topic content was also not considered but can bring up privacy concerns for patients discussing emotional memories with clinicians. Second, the sample was quite small and highly imbalanced, which makes our results including ML models with 6 features or more less likely to generalize to a larger sample. Third, we only examined language features and difficulty for the negative relived emotions; while these may be particularly challenging for those with FS, they yield salient, unpleasant memories that neither FS [12] nor those with PTS symptoms [31] are keen to experience, potentially minimizing group differences. Other language features may be relevant in differentiating groups, as might verbalizations beyond written text. Finally, other comparison groups may be relevant, such as those with epilepsy, particularly given that not everyone with FS has experienced trauma [2, 11].

With a larger dataset to train ML models and with other comparison groups, this work can be expanded and leveraged to create a training intervention and/or tool to help clinicians better identify whether a patient is experiencing FS (vs. epilepsy) and/or whether a patient is more avoidant of negative emotions. The current study demonstrated that subjective difficulty alone for negative relived emotions created the best ML models. When patients verbalize their emotional experiences, linguistic features may provide additional insight into their social and emotional interactions to distinguish FS from TC.

References

- [1] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. DSM-V. 2013. doi: <https://doi.org/10.1176/appi.books.9780890425596>
- [2] Brown RJ, Reuber M. Towards an integrative theory of psychogenic non-epileptic seizures (PNES). *Clin Psychol Rev*. 2016;47:55-70. doi: <https://doi.org/10.1016/j.cpr.2016.06.003>
- [3] LaFrance WC Jr, Baker GA, Duncan R, Goldstein LH, Reuber M. Minimum requirements for the diagnosis of psychogenic nonepileptic seizures: a staged approach: a report from the International League Against Epilepsy Nonepileptic Seizures Task Force. *Epilepsia*. 2013;54(11):2005-2018. doi: <https://doi.org/10.1111/epi.12356>
- [4] Benbadis SR, Hauser WA. An estimate of the prevalence of psychogenic non-epileptic seizures. *Seizure*. 2000;9(4):280-281. doi: <https://doi.org/10.1053/seiz.2000.0409>
- [5] Fiszman A, Alves-Leon SV, Nunes RG, Isabella DA, Figueira I. Traumatic events and posttraumatic stress disorder in patients with psychogenic nonepileptic seizures. *Epilepsy Behav*. 2004;5(6):818-825. doi: <https://doi.org/10.1016/j.yebeh.2004.08.011>
- [6] Novakova B, Howlett S, Baker R, Reuber M. Emotion processing and psychogenic non-epileptic seizures: a cross-sectional comparison of patients and healthy controls. *Seizure*. 2015;29:4-10. doi: <https://doi.org/10.1016/j.seizure.2015.03.005>
- [7] LaFrance WC Jr, Reuber M, Goldstein LH. Management of psychogenic nonepileptic seizures. *Epilepsia*. 2013;54(suppl 1):53-67. doi: <https://doi.org/10.1111/epi.12106>
- [8] Drane DL, Fani N, Hallett M, Khalsa SS, Perez DL, Roberts NA. A framework for understanding the pathophysiology of functional neurological disorder. *CNS Spectr*. 2021;26(6):555-561. doi: <https://doi.org/10.1017/S109285292000162X>
- [9] Roberts NA, Reuber M. Alterations of consciousness in psychogenic nonepileptic seizures: emotion, emotion regulation, and dissociation. *Epilepsy Behav*. 2014;30:43-49. doi:10.1016/j.yebeh.2013.09.040
- [10] van der Kruis SJ, Bodde NM, Vaessen MJ, Jansen JF. Functional connectivity of dissociation in patients with psychogenic non-epileptic seizures. *J Neurol Neurosurg Psychiatry*. 2012;83(3):239-247. doi: <https://doi.org/10.1136/jnnp-2011-300776>
- [11] Kaplan MJ, Dwivedi AK, Bowman M. Comparisons of childhood trauma, alexithymia, and defensive styles in patients with psychogenic non-epileptic seizures. *J Psychosom Res*. 2013;75(2):142-146. doi: <https://doi.org/10.1016/j.jpsychores.2013.05.007>
- [12] Roberts NA, Burleson MH, Torres DL, Wang NC. Emotional reactivity as a vulnerability for psychogenic nonepileptic seizures. *J Neuropsychiatry Clin Neurosci*. 2020;32(1):95-100. doi: <https://doi.org/10.1176/appi.neuropsych.18070160>
- [13] Pick S, Millman LM, Ward E, David AS. Unravelling the influence of affective stimulation on functional neurological symptoms. *J Neurol Neurosurg Psychiatry*. 2024;95(5):461-470. doi: <https://doi.org/10.1136/jnnp-2023-330123>
- [14] Cornaggia CM, Gugliotta SC, Magaudda A, Alfa R, Beghi M, Polita M. Conversation analysis in the differential diagnosis of Italian patients with epileptic or psychogenic non-epileptic seizures: a blind prospective study. *Epilepsy Behav*. 2012;25(4):598-604. doi: <https://doi.org/10.1016/j.yebeh.2012.09.010>

- [15] Rawlings GH, Brown I, Reuber M. Narrative analysis of written accounts about living with epileptic or psychogenic nonepileptic seizures. *Seizure*. 2018;62:59-65. doi:10.1016/j.seizure.2018.09.017
- [16] Rawlings GH, Brown I, Stone B, Reuber M. Written accounts of living with epilepsy or psychogenic nonepileptic seizures: a thematic comparison. *Qual Health Res*. 2018;28(6):950-962. doi: <https://doi.org/10.1177/1049732318759930>
- [17] Peacock M, Dickson JM, Bissell P, Grunewald R, Reuber M. Beyond the medical encounter: can the free association narrative interview method extend psychosocial understandings of non-epileptic attack disorder? *J Psychosoc Stud*. 2022;15(1):36-51. doi: <https://doi.org/10.1332/147867321X16119281118515>
- [18] Raffaelli Q, Mills C, de Stefano NA, O'Connor MF. The think aloud paradigm reveals differences in the content, dynamics, and conceptual scope. *Sci Rep*. 2021;11(1):19362. doi:10.1038/s41598-021-98876-3
- [19] Sawalha J, Yousefnezhad M, Shah Z, Brown MR, Greenshaw AJ, Greiner R. Detecting presence of PTSD using sentiment analysis from text data. *Front Psychiatry*. 2022;12:811392. doi: <https://doi.org/10.3389/fpsyg.2021.811392>
- [20] Lindquist KA, MacCormack JK, Shablack H. The role of language in emotion: predictions from psychological constructionism. *Front Psychol*. 2015;6:444. doi: <https://doi.org/10.3389/fpsyg.2015.00444>
- [21] Barrett LF, Lindquist KA, Gendron M. Language as context for the perception of emotion. *Trends Cogn Sci*. 2007;11(8):327-332. doi: <https://doi.org/10.1016/j.tics.2007.06.003>
- [22] Levenson RW, Carstensen LL, Friesen WV, Ekman P. Emotion, physiology, and expression in old age. *Psychol Aging*. 1991;6(1):28-35.
- [23] Silge J, Robinson D. tidytext: text mining and analysis using tidy principles. *J Open Source Softw*. 2016;1(3):37. doi: <https://doi.org/10.21105/ioss.00037>
- [24] Warriner AB, Kuperman V, Brysbaert M. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav Res Methods*. 2013;45(4):1191-1207. doi: <https://doi.org/10.3758/s13428-012-0314-x>
- [25] Russell JA. A circumplex model of affect. *J Pers Soc Psychol*. 1980;39(6):1161-1178. doi: <https://doi.org/10.1037/h0077714>
- [26] Rainio O, Teuho J, Klén R. Evaluation metrics and statistical tests for machine learning. *Sci Rep*. 2024;14(1):6086. doi: <https://doi.org/10.1038/s41598-024-06086-9>
- [27] Kuhn M, Wickham H. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. *Tidymodels*. 2022.
- [28] Kuhn M. Building predictive models in R using the caret package. *J Stat Softw*. 2008;28(5):1-26. doi: <https://doi.org/10.18637/jss.v028.i05>
- [29] De Diego IM, Redondo AR, Fernández RR, Navarro J, Moguerza JM. General performance score for classification problems. *Appl Intell*. 2022;52(10):12049-12063. doi: <https://doi.org/10.1007/s10489-022-03311-9>
- [30] Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010;5(9):1315-1316. doi: <https://doi.org/10.1097/JTO.0b013e3181ec173d>
- [31] Badour CL, Resnick HS, Kilpatrick DG. Associations between specific negative emotions and DSM-5 PTSD. *J Interpers Violence*. 2017;32(11):1620-1641. doi: <https://doi.org/10.1177/0886260515589930>

Pursuing Perfection or Striving Towards “Superwoman”? Investigating the Roles of Social Media and Individual Differences in Eating Disorders

Amanda BATISTA^{a,1} and Deborah HALL^a

^a*Arizona State University*

ORCiD ID: Amanda Batista <https://orcid.org/0009-0005-0666-7281>

Abstract. The prevalence of eating disorders continues to increase, particularly among young women. Prior research has identified social media use as one risk factor for eating disorders, along with trait perfectionism and the endorsement of unrealistic ideals pertaining to women’s ability to achieve success across multiple social roles. The present research investigated the extent to which endorsement of unrealistic standards for women across domains (i.e., the “Superwoman” ideal) and social media use interactively predict body dissatisfaction and disordered eating in young women. It also investigated whether endorsement of the Superwoman ideal uniquely predicts these outcomes over and above trait perfectionism. In an online survey of 18-25 year-old cisgender women ($N = 405$), greater social media use was linked with increased body dissatisfaction, and both stronger endorsement of the Superwoman ideal and greater social media use were associated with greater disordered eating. There was no evidence of an interaction between these predictors. Furthermore, whereas trait perfectionism uniquely predicted body dissatisfaction and disordered eating, endorsement of the Superwoman ideal uniquely predicted body dissatisfaction in the opposite direction as predicted and did not uniquely predict disordered eating. Implications, limitations, and exploratory findings of the present study are discussed.

Keywords. Eating disorders, body dissatisfaction, social media, gender role expectations, perfectionism, psychology

1. Introduction

Eating disorders have doubled in frequency throughout the U.S. since the start of the COVID-19 pandemic, disproportionately impacting adolescent girls and young women [1]. One contemporary risk factor for eating disorders is greater social media use [2], which may help to explain gender disparities in prevalence rates, given higher levels of social media use among women compared to men [3]. Indeed, prior research has established a link between social media use and endorsement of the ‘thin is beautiful’ ideal in young women [4], and the detrimental effects of specific social media platforms on the mental health of young women has been documented by the platforms, themselves. Instagram, for instance, found evidence of enhanced body image issues for one in three teen girls who used the platform [5].

An emerging social media trend that has yet to receive empirical attention is the ‘that’ girl trend (i.e., *#thatgirl*), which has garnered over 1.3 million posts on *TikTok* since October 2023 [6]. ‘That’ Girl encompasses a female-focused lifestyle promoting (a) wellness, (b) productivity, (c) beauty, and (d) mindfulness [7]. On one level, key aspects of the trend are positive – e.g., encouraging self-care, balance across life domains, and general female well-being. Yet, the trend’s emphasis on beauty and potentially unrealistic standards for achievement across multiple domains overlaps with a construct that has previously been linked with eating disorders [8]: the “Superwoman” ideal (i.e., the pressures felt by women to simultaneously perform agentic productivity, physical attractiveness, and effective achievement across social roles) [9]. Prior work thus suggests that trends like ‘that’ girl may have unintended negative consequences.

¹ Corresponding Author: Amanda Batista, akbatis1@asu.edu.

We propose eating disorders as one potential negative consequence. Crucially, prior research has yet to examine endorsement of the Superwoman ideal and degree of social media use together, in terms of their potential interactive effect on disordered eating. Furthermore, whereas social media use has been linked to trait perfectionism [10] (i.e., “striving for flawlessness and setting exceedingly high standards for performance, accompanied by tendencies for overly critical evaluations” [11]) — a well-known eating disorder risk factor [12] — and women who endorse the Superwoman ideal tend to be higher in perfectionism [13], the Superwoman ideal notably also captures unique aspects of women’s role expectations [14, 15] that may be especially relevant to current social media trends.

2. Present Research

We investigated the relationships between endorsement of the Superwoman ideal, social media use, trait perfectionism, body dissatisfaction, and disordered eating. We expected that greater endorsement of the Superwoman ideal would predict greater body dissatisfaction and disordered eating. Moreover, we predicted that these associations would be moderated by degree of social media use, such that the positive relations between endorsement of the Superwoman ideal and body dissatisfaction (H1a) and disordered eating (H1b) would be especially strong for those with greater social media use. We also predicted that endorsement of the Superwoman ideal would be a significant unique positive predictor of body dissatisfaction (H2a) and disordered eating (H2b) controlling for perfectionism.

3. Methods

3.1. Participants and Procedure

Cisgender women in the U.S. between the ages of 18-25 years old ($N = 405$; age: $M = 22.7$, $SD = 1.9$) completed an online survey via Prolific.com measuring endorsement of the Superwoman ideal, degree of social media use, trait perfectionism, body dissatisfaction, and disordered eating. The study, which received prior IRB approval, was pre-registered at Open Science Framework [16]. The target sample size was determined a priori based on a G*Power analysis for detecting a small moderation effect ($\alpha = .05$; $\beta = .80$). Additional information about the demographic composition of the sample can be found at [17].

3.2. Measures

Superwoman Ideal. We used the 27-item Superwoman Scale [9], which assesses how well a series of statements about attitudes and behaviors related to standards across multiple life domains reflect participants’ own beliefs (e.g., “I model myself after other people who ‘have it all,’ successful marriages, careers, and great kids”). Responses were measured on a 6-pt scale (*extremely well* to *not at all*) and the items ($\alpha = .80$) were averaged, with higher scores reflecting greater endorsement of the Superwoman ideal.

Social Media Use. We used the 9-item General Social Media Usage Subscale of the Media and Technology Usage and Attitudes Scale [18], which includes statements about frequency of social media use and engagement in specific social media activities (e.g., “liking,” posting). Responses were measured on a 10-pt scale (*never* to *all the time*) and the items ($\alpha = .97$) were averaged, with higher scores reflecting greater social media use.

Trait Perfectionism. We used the 25-item Perfectionism Cognitions Inventory [19], which assesses the frequency of a variety of thoughts concerning self-expectations, negative self-evaluation, and general perfectionism (e.g., “My work should be flawless”). Responses were measured on a 5-pt scale (*not at all* to *all of the time*) and the items ($\alpha = .91$) were averaged, with higher scores reflecting greater perfectionism.

Body Dissatisfaction. We used the 9-item Body Dissatisfaction Subscale of the Eating Disorder Inventory [20], which assesses frequency of identification with different attitudes and feelings about one’s body (e.g., “I feel satisfied with the shape of my body”)

[reverse scored]). Responses were measured on a 6-pt scale (*never* to *always*) and the items ($\alpha = .90$) were averaged, with higher scores reflecting greater body dissatisfaction.

Disordered Eating. We used the 12-item Eating Disorder Examination Questionnaire [21], which assesses individuals' frequency of maladaptive eating behaviors and cognitions within the past week (e.g., "Has thinking about food, eating or calories made it very difficult to concentrate on things you are interested in (such as working, following a conversation or reading)?"). Responses were measured on a 4-pt scale (items 1-10: 0 days to 6-7 days; items 11-12: *not at all* to *markedly*). The items ($\alpha = .91$) were averaged, with higher scores reflecting greater eating disorder symptoms.

Attention Checks. Three attention checks were administered, with failure of two or more checks as the exclusion criteria. Participants who met this criteria ($n = 2$) and one participant who timed out before completing the survey were filtered from the initial sample ($N = 408$), yielding a final sample size of $N = 405$.

4. Results

Descriptive statistics and bivariate correlations for the key study variables can be viewed at [17]. To address a potentially problematic positive skew in eating disorder scores, we employed a natural log transformation on this variable prior to performing the analyses below (see [16]). To test our first hypothesis, regression analyses were performed with endorsement of the Superwoman ideal (SI; mean centered) as the focal predictor, degree of social media use (SMU; mean centered) as the moderator, and either body dissatisfaction (BD; H1a) or eating disorder symptoms (DE; H1b) as the outcome variable. As shown in Table 1, in the model predicting BD, the only significant finding was a main effect of SMU, with greater SMU corresponding with greater BD. In the model predicting DE, stronger SI and greater SMU significantly predicted greater DE, however, there was once again no evidence of moderation. In sum, there was no support for the hypothesized moderation effects (i.e., H1a & H1b).

Table 1. Summary of Moderation Results

* $p < .05$, ** $p < .01$; Predictors: Superwoman ideal (SI), Social media use (SMU); Outcomes: Body dissatisfaction (BD), Disordered eating (DE)

	SI			SMU			SI x SMU		
	b (SE)	t	p	b (SE)	t	p	b (SE)	t	p
BD	-0.06 (0.09)	-0.72	.474	0.12 (0.04)	2.74	.006**	-0.07 (0.06)	-1.09	.275
DE	0.10 (0.03)	3.63	.000**	0.05 (0.01)	3.64	.000**	-0.01 (0.02)	-0.51	.607

The extent to which endorsement of the Superwoman ideal predicted body dissatisfaction and disordered eating over and above perfectionism was tested in separate regression models, with endorsement of the Superwoman ideal (SI) and perfectionism entered as predictors of body dissatisfaction (BD; H2a) and disordered eating (DE; H2b). Both SI ($b = -0.32$, SE = 0.09, $t(402) = -3.74$, $p < .001$) and perfectionism ($b = 0.59$, SE = 0.07, $t(402) = 8.86$, $p < .001$) uniquely predicted BD. Whereas greater perfectionism predicted greater BD, counter to hypotheses, greater SI predicted lower levels of BD. In the model for DE, only perfectionism was a significant unique predictor, with greater perfectionism corresponding with greater DE ($b = 0.23$, SE = 0.02, $t(402) = 11.58$, $p < .001$). Thus, results were largely inconsistent with H2a and H2b.

As noted at preregistration [16], we included some exploratory variables, such as the use of specific platforms, to examine in post-hoc analyses. As shown in Table 2, we tested a series of simple regressions in which daily use of each of eight different platforms was entered as a predictor of endorsement of the Superwoman ideal (SI), body dissatisfaction (BD), and disordered eating (DE), respectively. We found that: (1) greater daily Twitter use predicted greater BD and DE; (2) greater daily LinkedIn use predicted lower BD and DE; (3) greater daily YouTube use predicted lower SI; and (4) greater daily Instagram, LinkedIn, and Pinterest use predicted greater SI.

Table 2. Summary of Simple Regression Analyses for Daily Platform-Specific Social Media. Usage Predicting Superwoman Ideal Endorsement, Body Dissatisfaction, and Disordered Eating (N= 405)

* $p < .05$, ** $p < .01$; Outcomes: Superwoman ideal (SI), Body dissatisfaction (BD), Disordered eating (DE)

	Superwoman Ideal Endorsement (SI)			Body Dissatisfaction (BD)			Disordered Eating (DE)		
Platforms	b (SEb)	β	p	b (SEb)	β	p	b (SEb)	β	p
WhatsApp	0.05 (0.07)	.04	.471	-0.21 (0.12)	-.09	.075	0.03 (0.04)	.04	.440
Instagram	0.31 (0.12)	.12	.014*	0.18 (0.21)	.04	.395	0.13 (0.07)	.09	.076
TikTok	0.05 (0.08)	.03	.501	0.19 (0.13)	.07	.151	0.06 (0.04)	.07	.155
Facebook	0.10 (0.07)	.07	.157	-0.06 (0.12)	-.02	.623	-0.06 (0.04)	-.08	.130
LinkedIn	0.12 (0.06)	.10	.047*	-0.31 (0.11)	-.15	.003**	-0.10 (0.04)	-.14	.007**
Twitter/X	0.02 (0.06)	.02	.738	0.27 (0.11)	.12	.015**	0.11 (0.04)	.15	.003**
YouTube	-0.17 (0.08)	-.10	.044*	0.08 (0.14)	.03	.558	0.01 (0.05)	.01	.800
Pinterest	0.16 (0.06)	.13	.011*	0.06 (0.11)	.03	.580	0.03 (0.04)	.05	.357

5. Discussion & Conclusion

Greater social media use was linked with increased body dissatisfaction, and both stronger endorsement of the “Superwoman” ideal and greater social media use were associated with greater disordered eating in a sample of young women. Contrary to hypotheses, however, there was no evidence of an interaction between these predictors. Furthermore, whereas greater trait perfectionism uniquely predicted greater body dissatisfaction and disordered eating, endorsement of the Superwoman ideal uniquely predicted body dissatisfaction in the opposite direction as predicted and did not uniquely predict disordered eating.

Notable limitations of the study warrant mention. First, the lack of variability in age may have resulted in a restriction of range in degree of social media use, especially given that young women report significantly greater social media use than other segments of the population [3]. Another limitation stems from the use of an eating disorder scale that was designed for diagnostic use in clinical populations, given our study’s non-clinical sample; future research using a sub-clinical assessment tool would be beneficial. Furthermore, our measure of endorsement of the Superwoman ideal was originally developed in 1994 [9] and contained a number of statements that might not tap into the self-expectations most relevant to young women today. For instance, the scale included items pertaining to parenting (e.g., “I would be disappointed if my children did not qualify for ‘gifted and talented’ programs”) and career aspirations (e.g., “I probably won’t have my own secretary”) that are likely outdated. Finally, future research with greater diversity within the sample will strengthen the generalizability of the present findings.

References

- [1] Hartman-Munick SM, Lin JA, Milliren CE, Braverman PK, Brigham KS, Fisher MM, et al. Association of the COVID-19 pandemic with adolescent and young adult eating disorder care volume. *JAMA Pediatr.* 2022;176(12):1225. doi: <https://doi.org/10.1001/jamapediatrics.2022.4346>
- [2] Padín PF, González-Rodríguez R, Verde-Diego C, Vázquez-Pérez R. Social media and eating disorder psychopathology: A systematic review. *Cyberpsychol J Psychosoc Res Cyberspace.* 2021;15(3). doi: <https://doi.org/10.5817/cp2021-3-6>
- [3] Dixon SJ. U.S. social media user distribution by gender 2024 [Internet]. Statista; 2024 [cited 2024 Jun 14]. Available from: [https://www.statista.com/statistics/1319300/us-social-media-audience-by-gender/#:~:text=As%20of%20January%202023%2C%2053.9,media%20users%20in%20the%20country](https://www.statista.com/statistics/1319300/us-social-media-audience-by-gender/#:~:text=As%20of%20January%202023%2C%2053.9,media%20users%20in%20the%20country.).
- [4] Mingoja J, Hutchinson AD, Wilson C, Gleaves DH. The relationship between social networking site use and the internalization of a thin ideal in females: A meta-analytic review. *Front Psychol.* 2017;8:1351. doi: <https://doi.org/10.3389/fpsyg.2017.01351>
- [5] Wells G, Horwitz J, Seetharaman D. Facebook knows Instagram is toxic for teen girls, company documents show - WSJ [Internet]. Wall Street Journal; 2021 [cited 2024 Jun 13]. Available from: <https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739>.

- [6] TikTok. Thatgirl [Internet]. TikTok; 2024 [cited 2024 Jun 13]. Available from: <https://www.tiktok.com/tag/thatgirl?lang=en>.
- [7] What does it mean to be that girl online?: Later glossary [Internet]. [cited 2024 Jun 13]. Available from: <https://later.com/social-media-glossary/that-girl/>.
- [8] Theran SA, Dour H. Superhero ideal, authenticity in relationships, and depressive symptoms: A multiple mediation analysis. *Acta Psychol (Amst)*. 2022;226:103559. doi: <https://doi.org/10.1016/j.actpsy.2022.103559>
- [9] Murnen SK, Smolak L, Levine MP. Development of a scale to measure adherence to the “superwoman” construct [Unpublished Manuscript]. 1994. 26 p. Located at: Kenyon College, OH.
- [10] Kurz E. I will never be good enough!! – the rise of perfectionism among young adults. *Eur J Psychother Couns*. 2021 Jan 2;23(1):85–98. doi: <https://doi.org/10.1080/13642537.2021.1881140>
- [11] Stoeber J. The dual nature of perfectionism in sports: Relationships with emotion, motivation, and performance. *Int Rev Sport Exerc Psychol*. 2011;4(2):128–45. doi: <https://doi.org/10.1080/1750984x.2011.604789>
- [12] Stackpole R, Greene D, Bills E, Egan SJ. The association between eating disorders and perfectionism in adults: A systematic review and meta-analysis. *Eat Behav*. 2023;50:101769. doi: <https://doi.org/10.1016/j.eatbeh.2023.101769>
- [13] Gunnard K, Krug I, Jiménez-Murcia S, Penelo E, Granero R, Treasure J, et al. Relevance of social and self-standards in eating disorders. *Eur Eat Disord Rev*. 2011;20(4):271–8. doi: <https://doi.org/10.1002/erv.1148>
- [14] Hock HW. Review of the worldwide epidemiology of eating disorders. *Curr Opin Psychiatry*. 2016 Nov;29(6):336–9. doi: <https://doi.org/10.1097/yco.0000000000000282>
- [15] Murnen SK, Smolak L. Negotiating gender roles to enact body appreciation and positive embodiment. In: Tylka TL, Calogero RM, Piran N, editors. *Handbook of positive body image and embodiment*. Oxford: Oxford University Press; 2019. p. 161–72. doi: <https://doi.org/10.1093/med-psych/9780190841874.003.0016>
- [16] Batista A, Hall DL. The starving “superwoman”: Personal- and context-related factors as predictors of maladaptive eating and body behaviors [Internet]. OSF; 2024 [cited 2024 Jun 13]. Available from: <https://osf.io/3cmjk>.
- [17] Batista A, Hall DL. Pursuing perfection or striving towards “superwoman”? - Investigating the roles of social media and individual differences in eating disorders [Internet]. OSF; 2024 [cited 2024 Jun 13]. Available from: https://osf.io/x2uem/?view_only=3a44ef75b03f4936be11c88b07d8ce67.
- [18] Rosen LD, Whaling K, Carrier LM, Cheever NA, Rokkum J. The media and technology usage and attitudes scale: An empirical investigation. *Comput Human Behav*. 2013;29(6):2501–11. doi: <https://doi.org/10.1016/j.chb.2013.06.006>
- [19] Flett GL, Hewitt PL, Demerjian A, Sturman ED, Sherry SB, Cheng W. Perfectionistic automatic thoughts and psychological distress in adolescents: An analysis of the perfectionism cognitions inventory. *J Ration Emot Cogn Behav Ther*. 2012;30(2):91–104. doi: <https://doi.org/10.1007/s10942-011-0131-7>
- [20] Garner DM, Olmstead MP, Polivy J. Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *Int J Eat Disord*. 1983;2(2):15–34. doi: [https://doi.org/10.1002/1098-108x\(198321\)2:2<15::aid-eat2260020203>3.0.co;2-6](https://doi.org/10.1002/1098-108x(198321)2:2<15::aid-eat2260020203>3.0.co;2-6)
- [21] Gideon N, Hawkes N, Mond J, Saunders R, Tchanturia K, Serpell L. Correction: Development and psychometric validation of the EDE-QS, a 12-item short form of the eating disorder examination questionnaire (EDE-Q). *PLoS One*. 2018;13(11). doi: <https://doi.org/10.1371/journal.pone.0207256>

Understanding What Actions and Traits Facilitate Social Media Campaigning on Transgender Issues Through Perceptions of the BBC and Its Media Coverage

Dan LLOYD^{a,1} and CHALENNOR^a

^a*Institute of Art, Design and Technology, Dún Laoghaire, Dublin, Ireland*
ORCID ID: Dan Lloyd <https://orcid.org/0009-0001-0805-3989>

Abstract. This research examined how online transgender communities perceive and react to the British Broadcasting Corporation (BBC) and its gender identity coverage. Existing research is limited due to broad focus, low sample sizes, and self-report or semi-structured interview methods [1-3]. Using reflexive thematic analysis, 354 posts on the BBC and its content from a UK transgender social media community were analyzed. The reflexive element controlled researcher bias [4]. Results showed negativity towards the BBC, a need for reporting transparency, and the community's use of social media for information-seeking and promoting collective action. The community has a complex relationship with the BBC and uses social media to mobilize against perceived threats to their identity and existence. Members confirm their identities through scrutinized media coverage but disagree with their presentation in BBC coverage. Creating an overarching anti-trans media entity helped unify the group. Limitations exist in social media data due to a lack of essential characteristics and participant anonymity.

Keywords. Transgender, social media, BBC, News, gender identity.

1. Introduction

Social media discourse can shape opinion, offer support, or inflict harm, particularly in discussions of transgender people and gender identification. The topic is divisive and causes friction between groups, with conversations developing online, on television, within governments, and in newspapers [5]. Those transitioning feels prejudiced in many areas of society, and guidance for the media when reporting on gender identity is lacking [1,2].

Studies have found that while social network services (SNS) help educate those forming gender identities, they can cause distress from anti-trans attacks and negative news [6,7]. Perceived injustice and habitual SNS use drive the spread of disinformation and engagement with news [8,9]. Emotionally charged responses are more likely to be shared based on previous beliefs [10,11]. The theoretical framework includes the Online Disinhibition Effect [12], Uses and Gratifications Theory [13], Social Identity Theory [14], the MAD Model of Moral Contagion [15] and Moral Outrage [16].

2. Theoretical Framework

2.1. The Transgender Debate on Social Media

The transgender community uses SNS to establish identity and gain support [5]. However, they also feel it is only sometimes a safe space and would welcome changes to moderation rules [17].

¹ Corresponding Author: Dan Lloyd, danjlloyd@gmail.com

2.2. Advocacy and Opposition in News

Transgender individuals perceive many media outlets as a threat, particularly in sports coverage [18]. The BBC's investigation into LGBTQ+ charity Stonewall's influence on public bodies challenged its impartiality [19,20]. Poor journalism practices can result in increased hostility towards transgender people [21].

2.3. Disinhibition Online and Individual Roles

Suler's Online Disinhibition Effect [12,22] explains how anonymity on SNS facilitates the creation of online identities that behave differently from offline. Overuse of SNS can be a poor coping strategy for anxiety [23], and emotional investment can lower well-being [3].

2.4. Understanding Reddit Use in the Transgender Debate

Uses and Gratifications Theory (UGT) [13] demonstrates how different media types deliver distinct needs to users. Reddit is a popular information resource for the transgender community, especially for health advice [24-26].

2.5. Social Identity Theory and Community

Social Identity Theory (SIT) [14] helps understand the sense of community and belonging to a group. The binary model of 'ingroup' and 'outgroup' illustrates the amplification of divisive issues during social media campaigns. Collective identity and action are linked to SIT.

2.6. Morality and Emotional Reaction on Social Media

Three theories consider these phenomena:

1. Moral Outrage [16]: An emotional response that holds perceived immoral actors accountable.
2. MAD Model of Moral Contagion [15]: Outlines motivation, attention capture, and platform design contributing to the spread of moral sentiment online.
3. Heuristic Systematic Model (HSM) [27]: Presents dual paths of information processing - systematic (careful analysis) and heuristic (minimal cognitive effort). Kohout et al. [28] found negative comments attracted the most attention when applying the heuristic route. Recent studies [29,30] found that readers overestimate moral outrage compared to post authors, and emotion plays a vital role in shaping online discourse.

This study proposes the following research questions:

RQ1. What are user perceptions of the BBC and BBC gender identity coverage on transgender Reddit forums?

RQ2. What communication strategies are used by transgender community members on Reddit when discussing the BBC and its gender identity coverage?

3. Methodology

3.1. Design

Qualitative experiential framework with a constructionist approach using reflexive thematic analysis (RTA) [4].

3.2. Data sample

Reddit, an open-source platform with diverse communities, was chosen. The UK-focused r/transgenderUK subreddit (47,000+ members) had the most comments regarding the BBC.

3.3. Data Collection and Screening

Searches are limited to the last 12 months [31]. 23 threads were identified, resulting in 698 posts and comments. Screening produced 354 posts and comments for analysis.

3.4. Data Analysis

An inductive and deductive approach identified semantic and latent codes [4]. The principal researcher, a former BBC News Editor, documented their unique subjective position through journaling.

3.5. Ethical Considerations

Approved by the Ethics Committee, Institute of Art, Design and Technology, Dún Laoghaire. APA Guidelines [32], BPS Codes of Ethics and Conduct [33], AOIR ethical guidance [34], recommendations from [35] and Reddit's content policy were followed.

4. Results

Three themes were identified from the RTA of the r/transgenderUK subreddit data.

4.1. Theme 1: Highlighting the anti-trans media agenda

An overarching theme was the concept of the BBC being influenced by transphobic, right-wing, or conservative forces, often labelled as corrupt. There was a need for information about the BBC's conduct. Sub-themes include: 1.1 Media Techniques to Frame Negatively: Detailed critiques of BBC reporting seen as biased against trans people, including comments from 'anti-trans' groups, unsuitable pictures, and headline/syntax changes to discredit the community. 1.2 The 'Transilluminati': Theories of the BBC controlled by secretive anti-trans entities like the 'Illuminati', MI5, or the UK government. Often accepted uncritically within the community. 1.3 BBC Bias: Widespread perception of anti-trans. Some disagreement and justification of reporting. 1.4 Information-Seeking and Education: Clarification sought on BBC stances, funding, and complaint procedures. Alternative news sources and a potential trans-friendly news website were discussed.

4.2. Theme 2: Threat to Identity and Existence

Highly emotional discussions conveyed anger and sadness at BBC coverage seen as dismissing or sensationalising trans identities and experiences. Sub-themes: 2.1 Transgender People Are Invisible: Desire for recognition of authentic selves; attributing societal attitudes to harmful media representation. Gender is seen as foundational to existence. 2.2 Feeling Despondent and Angry: Complex emotions, from anger to hopelessness, towards BBC content questioning trans identities. Reveals the dual nature of discourse and perceptions. 2.3 Deadnaming: Central issue provoking many responses. Importance of naming conventions and post-mortem identity. Historic examples are frequently cited. 2.4 Poor Journalism: Criticism of source credibility and fact-checking, especially on trans identities, experiences, and journeys. Insufficient representation of trans voices.

4.3. Theme 3: Transgender People Are at War

Language conveyed a constant battle against 'enemies', with comparisons to genocide. Calls for action against 'anti-trans' journalism. Sub-themes: 3.1 Us Versus Them: 'Terfs', 'cisgender' people as opposing powers. Varying perceptions of different BBC entities - news as hostile, entertainment as friendly. Dismay at 'British values'. 3.2 Genocide: BBC is seen as complicit in trans 'genocide' through negative portrayals, condoning discrimination. Evokes fear, anger, and existential vulnerability. 3.3 Culture War: Media, particularly the BBC, accused of stoking anti-trans hate through biased

coverage. Attitudes of 'enemies' are seen as entrenched and unchangeable. 3.4 Collective Action: Prevalent theme. Formal BBC complaints, TV license cancellation, and archiving webpages to deny traffic are encouraged to be heard and make a difference.

The themes reveal a complex psychological relationship between the transgender community, media, themselves, and each other on SNS, characterised by strong emotions, perceived threats, and a desire for change.

5. Discussion

5.1. The Perceptions of the BBC and Transgender Coverage (RQ1)

Perceptions of BBC media coverage were mainly negative, with headlines and portrayals of transgender people in specific articles (e.g., crime) being sources of negativity. These misgivings were vivid to the community, with users citing historical examples and rarely expressing positive feedback without scepticism. The anger towards BBC content aligns with Kohout et al.'s [28] findings on anger resulting in systematic processing, and the use of Reddit to highlight ignorance in BBC journalism supports the need for improved editorial guidance on trans-issues [1,18].

Perceptions of the BBC as an organisation were that it had an anti-transgender agenda and was controlled by entities like the 'Transilluminati,' supporting Tajfel et al.'s [14] ingroup/outgroup concept and the Us Versus Them narrative. These unfounded claims showed distrust of institutions [22] and negativity towards establishment bodies fuelling bogus information [8]. Ingroup behaviour favoured negative accounts of the BBC [10,11], but outliers showed no uniform response, aligning with Suler's [22] role of individuality in online disinhibition.

5.2. Communication Strategies (RQ2)

Subreddit members used the platform for Information-Seeking and Education, advice was often unchallenged [24]. Repeated interaction could lead to users adopting educator roles to achieve status [9,13,22].

Emotion, notably anger, contributed to increased interaction when posts had substance. Anger took forms like swearing, explicit expressions, and colloquialisms ('terfs,' 'Tories,' and 'cisgender'), following the theory of moral outrage [16] and Brady et al.'s [15] findings on positive reactions to outrage. However, outrage without relational news was not always engaged with, possibly due to the norm of seeking information.

Communication also produced despondent and fearful comments, often comparing the situation to genocide, aligning with findings on the negative impact of unfavourable content [6]. These displays may be designed to seek support [5].

Collective action, like complaining about stories or advocating for the cessation of BBC funding was another communication strategy, aligning with findings on Twitter and Social Identity Theory [24, 14]. However, this action can also drive positive change in media representation [5].

5.3. Practical Implications

The BBC recently admitted a breach of accuracy rules when omitting a murderer's trans identity [36]. Transparency in story construction and organisational structures could improve public perceptions of the BBC [37]. Additional training for journalists could enhance reporting on sensitive communities.

5.4. Theoretical Implications

Combining theories like the online disinhibition effect [22], social identity theory [14], uses and gratifications theory [13], and the MAD model of moral contagion [15] helped inform the impact of SNS on online discourse. Using multiple theories allowed for identifying more semantic and latent codes, resulting in unique themes and sub-themes.

6. Conclusion

This paper examined transgender community campaigning against the BBC, revealing that a lack of transparency creates negative perceptions. Transgender people seek education and fight for positive changes. These insights could guide media bodies and social media companies.

Despite limitations, this research contributes to an under-researched area. Future research could build upon this framework to provide guidance and counteract ill feelings towards media within the transgender community.

References

- [1] Akrivos D. Transgender reporting in the British press: editorial standards and discursive harms in the post-Leveson era. *J Media Law.* 2022;14(2):406–41. doi: <https://doi.org/10.1080/17577632.2022.2136378>
- [2] Chan ASW, Wu D, Lo IMC, Ho IMC, Yan E. Diversity and inclusion: impacts on psychological wellbeing among lesbian, gay, bisexual, transgender, and queer communities. *Front Psychol.* 2022;13:894742. doi: <https://doi.org/10.3389/fpsyg.2022.894742>
- [3] Chan RCH. Benefits and risks of LGBT social media use for sexual and gender minority individuals: an investigation of psychosocial mechanisms of LGBT social media use and well-being. *Comput Human Behav.* 2023;139:107531. doi: <https://doi.org/10.1016/j.chb.2022.107531>
- [4] Braun V, Clarke V. Thematic analysis: a practical guide. *QM&P Bull.* 2022;1(33):46–50.
- [5] Coker MC. What to withhold and when to disclose: gender transitions and privacy management on social media. *Qual Res Rep Commun.* 2021;23(1):39–45. doi: <https://doi.org/10.1080/17459435.2021.1888222>
- [6] Scheuerman MK, Branham SM, Hamidi F. Safe spaces and safe places. *Proc ACM Hum Comput Interact.* 2018;2(CSCW):1–27. doi: <https://doi.org/10.1145/3274373>
- [7] Pham A, Morgan A, Kerman HM, Albertson K, Crouch J, Inwards-Breland DJ, et al. How are transgender and gender nonconforming youth affected by the news? A qualitative study. *J Adolesc Health.* 2020;66(4):467–75.
- [8] Winterlin F, Schatto-Eckrodt T, Frischlich L, Boberg S, Reer F, Quandt T. "It's us against them up there": spreading online disinformation as populist collective action. *Comput Human Behav.* 2023;146:107784. doi: <https://doi.org/10.1016/j.chb.2023.107784>
- [9] Lee CS, Ma L. News sharing in social media: the effect of gratifications and prior experience. *Comput Human Behav.* 2012;28(2):331–9. doi: <https://doi.org/10.1016/j.chb.2011.10.002>
- [10] Rijo A, Waldzus S. That's interesting! The role of epistemic emotions and perceived credibility in the relation between prior beliefs and susceptibility to fake-news. *Comput Human Behav.* 2023;141:107619. doi: <https://doi.org/10.1016/j.chb.2023.107619>
- [11] Kim SJ. The role of social media news usage and platforms in civic and political engagement: focusing on types of usage and platforms. *Comput Human Behav.* 2023;138:107475. doi: <https://doi.org/10.1016/j.chb.2022.107475>
- [12] Suler J. The online disinhibition effect. *Cyberpsychol Behav.* 2004;7(3):321–6. doi: <https://doi.org/10.1089/1094931041291295>
- [13] Katz E, Blumler JG, Gurevitch M. Uses and gratifications research. *Public Opin Q.* 1973;37(4):509–23.
- [14] Tajfel H, Billig M, Bundy RP, Flament C. Social categorization and intergroup behaviour. *Eur J Soc Psychol.* 1971;1(2):149–78.
- [15] Brady WJ, Crockett MJ, Van Bavel JJ. The MAD model of moral contagion: the role of motivation, attention, and design in the spread of moralized content online. *Perspect Psychol Sci.* 2020;15(4):978–1010. doi: <https://doi.org/10.1177/1745691620917336>
- [16] Crockett MJ. Moral outrage in the digital age. *Nat Hum Behav.* 2017;1(11):769–71. doi: <https://doi.org/10.1038/s41562-017-0213-3>
- [17] Haimson OL, Buss J, Weinger Z, Starks DL, Gorrell D, Baron BS. Trans Time. *Proc ACM Hum Comput Interact.* 2020;4(CSCW2):1–27. doi: <https://doi.org/10.1145/3415193>
- [18] Scovel S, Nelson M, Thorpe H. Media framings of the transgender athlete as "legitimate controversy": the case of Laurel Hubbard at the Tokyo Olympics. *Commun Sport.* 2022;11(5):838–53. doi: <https://doi.org/10.1177/21674795221114891>
- [19] Thompson D. Stonewall's influence on BBC and Ofcom revealed. BBC News [Internet]. 2021 Oct 14. <https://www.bbc.com/news/uk-58917227>
- [20] BBC News (UK). Stonewall's influence on BBC and Ofcom revealed. X [Internet]. 2021 Oct 14. <https://twitter.com/BBCNews/status/1448744059590352906?s=20>
- [21] [21] Li M. Exemplifying power matters: the impact of power exemplification of transgender people in the news on issue attribution, dehumanization, and aggression tendencies. *Journal Pract.* 2021;17(3):525–53.
- [22] [22] Suler J. Psychology of the digital age: humans become electric. Cambridge: Cambridge University Press; 2016.
- [23] [23] Lerma M, Marquez C, Sandoval K, Cooper TV. Psychosocial correlates of excessive social media use in a Hispanic college sample. *Cyberpsychol Behav Soc Netw.* 2021;24(11):722–8. doi: <https://doi.org/10.1089/cyber.2020.0560>
- [24] [24] Saeed R. Twitter and development of social capital: a gender perspective. *Telemat Inform Rep.* 2023;9:100037. doi: <https://doi.org/10.1016/j.teler.2023.100037>
- [25] [25] Mohseni M. Motives of online hate speech: results from a quota sample online survey. *Cyberpsychol Behav Soc Netw.* 2023;26(7):499–506. doi: <https://doi.org/10.1089/cyber.2023.0014>

- [26] [26] Bretherton I, Thrower E, Zwickl S, Wong AFQ, Chetcuti D, Grossmann M, et al. The health and well-being of transgender Australians: a national community survey. *LGBT Health*. 2021;8(1):42–9. doi: <https://doi.org/10.1089/lgbt.2020.0179>
- [27] Chaiken S. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *J Pers Soc Psychol*. 1980;39(5):752–66. doi: <https://doi.org/10.1037/0022-3514.39.5.752>
- [28] Kohout S, Kruikemeier S, Bakker BN. May I have your attention, please? An eye tracking study on emotional social media comments. *Comput Human Behav*. 2023;139:107495. doi: <https://doi.org/10.1016/j.chb.2023.107495>
- [29] Brady WJ, McLoughlin KL, Torres MP, Luo KF, Gendron M, Crockett MJ. Overperception of moral outrage in online social networks inflates beliefs about intergroup hostility. *Nat Hum Behav*. 2023;7(6):714–23. doi: <https://doi.org/10.1038/s41562-023-01574-5>
- [30] Chen K, He Z, Chang R, May J, Lerman K. Anger breeds controversy: analyzing controversy and emotions on Reddit. In: Lecture Notes in Computer Science. 2023. p. 44–53.
- [31] Sit M, Elliott SA, Wright KS, Scott SD, Hartling L. Youth mental health help-seeking information needs and experiences: a thematic analysis of Reddit posts. *Youth Soc*. 2022;56(1):24–41. doi: <https://doi.org/10.1177/0044118X221107404>
- [32] American Psychological Association [APA]. Guidelines for psychological practice with transgender and gender nonconforming people [Internet]. 2015. <https://psycnet.apa.org/doiLanding?doi=10.1037%2Fa0039906>.
- [33] Code of Ethics and Conduct. In: British Psychological Society eBooks. 2021.
- [34] Franzke AS, Bechmann A, Ess CM, Zimmer M. Internet research: ethical guidelines 3.0 [Internet]. 2020. <https://aoir.org/reports/ethics3.pdf>.
- [35] Adams NN. 'Scraping' Reddit posts for academic research? Addressing some blurred lines of consent in growing internet-based research trend during the time of Covid-19. *Int J Soc Res Methodol*. 2022;27(1):45–58. doi: <https://doi.org/10.1080/13645579.2021.1935565>
- [36] Tobitt C. BBC says omission of trans identity of murderer Scarlet Blake was breach of accuracy rules. *Press Gazette* [Internet]. 2024 Apr 12. <https://pressgazette.co.uk/the-wire/newspaper-corrections-media-mistakes-errors-legal/bbc-trans-pronouns-scarlet-blake/>.
- [37] Prochazka F, Obermaier M. Trust through transparency? How journalistic reactions to media-critical user comments affect quality perceptions and behavior intentions. *Digit Journal*. 2022;10(3):452–72. doi: <https://doi.org/10.1080/21670811.2021.1987948>

SECTION V

CLINICAL OBSERVATIONS

Cybertherapy is a field that is growing rapidly due to today's technology and information boom.

Virtual reality and advanced technologies have been used successfully to in a variety of healthcare issues, including treatment of anxiety disorders and phobias, treatment of eating and body dysmorphic disorders, neuropsychological assessment and rehabilitation and distraction during painful or unpleasant medical procedures.

The novel applications of these technologies yield many advantages over traditional treatment modalities, and the disadvantages that accompanied the first trials of virtual reality are quickly being addressed and eliminated.

Virtual reality peripherals such as data gloves, physiological monitoring and Internet worlds are swiftly demonstrating their usefulness in cybertherapy applications.

Wiederhold & Wiederhold, 2004

Development and Testing of a Virtual Reality Intervention with Adolescents Hospitalized for Suicide-Related Crises

Marisa E. MARRACCINI^{a,1}, Robert HUBAL^b, Lauren E. DELGATY^a, Telieha J.

MIDDLETON,^a Rachel ANONICK^a, Yixin LI^c, and Jerry HENEGHAN^c

^a*School of Education, University of North Carolina at Chapel Hill*

^b*Renaissance Computing Institute, University of North Carolina at Chapel Hill*

^c*BioMojo*

ORCiD ID: Marisa Marraccini <https://orcid.org/0000-0002-7965-3703>

Abstract. The primary objective of this study was to build and test a virtual reality (VR) intervention that teaches adolescents hospitalized for suicide-related risk therapeutic skills and provides them opportunities to practice using these skills in immersive VR scenarios. The Practice Experiences for School Reintegration (PrESR) is facilitated by a clinician to help hospitalized adolescents prepare for their return to school. Our model was informed by input from adolescents with lived experience, school professionals, and hospital professionals, and the intervention was built and iteratively refined based on feedback from a community sample of adolescents (n=6), adolescents previously or currently hospitalized for suicide-related thoughts or behaviors (n=13), and hospital professionals (n=8). This paper describes the iterative development and refinement process leading to the final PrESR intervention, which is currently being tested in a pilot optimization trial.

Keywords. Suicide, virtual reality, intervention development, adolescence, cognitive behavioral therapy

1. Introduction

Increasing rates of suicide-related thoughts and behaviors [1] and hospitalization for suicide-related crises among adolescents [2,3] necessitate innovative approaches for improving treatment and recovery. Following hospital discharge, adolescents are at heightened risk of additional attempts or re-hospitalization [4]. Limited research has addressed how to prepare adolescents during this transition for handling stressors and reacclimating to school [5,6].

Although inpatient programs use diverse treatment approaches, treatment typically focuses on stabilization and psychopharmacotherapy [7,8]. One well-suited treatment, Cognitive Behavioral Therapy (CBT) [8], includes teaching, modeling, and practicing therapeutic skills. Yet, within the confines of hospital treatment, meaningful practice opportunities that could translate to daily settings are difficult to execute because of time constraints and engagement difficulties. Consequently, we developed a Virtual Reality (VR) intervention that allows hospitalized adolescents to learn and practice CBT skills. Specifically, Practice Experiences for School Reintegration (PrESR) aims to increase skill knowledge and use, to improve school reintegration and prevent suicide [9]. Adolescent patients using PrESR can learn and practice three skills (affect regulation, cognitive restructuring, problem-solving) in an immersive school scenario. Development and testing of PrESR occurred over three phases. First, we recruited adolescents previously hospitalized for suicide-related crises and school and hospital professionals to provide input on the PrESR model. We conducted qualitative analyses on transcripts of interviews [9] to refine the model. Second, we recruited participants representative of a range of roles to review scenario scripts and/or playtest a PrESR prototype. Third, we initiated a pilot optimization trial in two psychiatric inpatient hospitals with adolescents

¹ Corresponding Author: Marisa E. Marraccini, mmarracc@unc.edu.

hospitalized for suicide-related risk. Here we present findings from the second phase of this research, which involved feedback from a sample community of adolescents, adolescents with lived experiences, and hospital professionals.

2. Methods

Methods for the development of PrESR were informed by the Person-Based Approach to Intervention Development [10]. This approach includes four stages of research towards developing and testing digital health-related behavior change interventions, from planning to implementation and trialing [10]. We have previously shared early results from the design stage addressing model development [9] and here expand on results from this stage addressing intervention development. All procedures were reviewed and approved by an Institutional Review Board prior to participant interaction.

2.1. Participants

Participants included a community sample of adolescents (n=6), adolescents currently or previously hospitalized for suicide-related crises (n=13), and hospital professionals (n=8). Participants provided permission, assent, and/or consent (depending on age) prior to participation.

Community sample. Adolescents from the community were recruited through local social media and listservs using a snowball sampling approach. Eligibility criteria included: (1) ages 13-18; (2) attended school recently; (3) ability to speak, read, and understand English; (4) consent of a parent/legal guardian; and (5) adolescent assent. Participants identified their race as White (n=4), Mexican American (n=1), or Black or African American (n=1), and ethnicity as Hispanic (n=1) or non-Hispanic (n=5). Two identified as woman/girl and reported their sex to be female, two identified as man/boy and reported their sex to be male, one identified in some other way and reported their sex to be female, and one identified as man/boy and reported their sex to be female.

Hospitalized sample. Adolescents with current or previous suicide-related risk were identified from medical records from one child and adolescent psychiatric inpatient hospital located in the southeast. For previously hospitalized adolescents, we mailed an opt out letter to parents/guardians or adult adolescents and followed up with phone calls. For adolescents currently hospitalized, we called parents or legal guardians to seek permission before approaching hospitalized adolescents for assent. Eligibility criteria included: (1) hospitalization for suicidal thoughts or behaviors in past six months; (2) ages 13-18; (3) return (or expected return) to school following discharge; (4) ability to speak, read, and understand English; (5) consent of a parent/legal guardian; and (5) adolescent assent. Minimal risk for motion sickness was required for testing the VR specifically, and if of concern participants were able to provide feedback on components of the intervention without engaging in VR. Participants identified their race as Asian and White (n=1), White (n=8), Black or African American (n=2), and White and Hispanic (n=1), and ethnicity as Hispanic (n=3) or non-Hispanic (n=10).¹ Four identified as woman/girl and reported their sex to be female, seven identified as man/boy and reported their sex to be male, one identified in some other way and reported their sex to be female, and one identified as man/boy and reported their sex to be female.

Professional sample. Clinicians (n=8) were recruited from a child and adolescent psychiatric inpatient hospital and identified their role as occupational therapist (n=2), recreational therapist (n=1), or teacher (n=5). They identified as Black or African American (n=2) and White (n=5), and ethnicity as non-Hispanic (n=7). Five identified as female and reported their sex to be female, one identified as transgender and reported their sex to be female, and one identified as male and reported their sex to be male.²

¹ One adolescent selected “prefer not to answer” when asked about their race.

² Demographics are missing from one hospital professional.

2.2. Instrumentation

We collected demographic information and structured feedback from participants using a self-report questionnaire. We audio recorded qualitative, semi-structured interviews using a think aloud protocol, and completed a debrief following each session to identify disruptions and key takeaways from interviews for rapid analysis. An observation protocol from a previously validated rubric of decision making and social competency skills [11] was used for all sessions; it involved eight ratings of engagement, verbalizations, comprehension, emotional control, provocativeness, reaction time, affect, and reflection. Two independent raters provided a rating to each of the eight categories on a scale from 1 (very low) to 5 (very high), meeting to come to consensus.

We gave the Motion Sickness Questionnaire, which measures susceptibility to motion sickness via different modes of transportation [12], to adolescents previously or currently hospitalized and to hospital professionals as a screening method for determining appropriateness of VR. We used established norms [12] and excluded participants above the 50th percentile. We provided the Simulator Sickness Questionnaire, which measures cybersickness levels after VR, to assess nausea and oculomotor symptoms [13]. Adolescents using VR additionally completed the PROMIS Anxiety and Depression instrument, which measures levels of anxiety and depression [14].

3. Intervention Development

3.1. Procedures

We developed the intervention sequentially based on iterative feedback from participants. First, we developed scripts for difficult school scenarios and lessons about CBT based on data collected during the first phase of the study [9]. Specifically, we developed (1) scripts that taught and reinforced introduction to CBT skills (introduction to CBT, affect regulation, cognitive restructuring, and problem solving); and (2) practice scenarios involving conversations between peers and adults in the school reflecting the most important and commonly recommended situations adolescents may face when returning to school. We designed three scenarios to integrate multiple stressors and interactions, including: (1) a classroom setting in which the adolescent engages in a lesson, interacts with peers in a group activity, and discusses missed work and absences with a teacher; (2) a cafeteria setting in which the adolescent interacts with different peers about their recent absence due to their hospitalization; and (3) a meeting setting in which adults in the school discuss re-entry planning with the adolescent's parent or legal guardian. Findings from the first phase of the study informed several other ideas for the intervention (e.g., virtual incentives in the intervention, additional scenarios) [9], but, due to resource constraints, we focused on core components of the intervention.

Second, we recruited participants to provide feedback across different stages of intervention development (see Figure 1). We introduced participants to PrESR with a brief overview describing its goals and aims and then invited them to playtest (from scripts to actual VR). While playtesting, participants were invited to describe their experience with whatever came to mind. We monitored participant safety during VR testing by having participants describe their level of distress using the Subjective Units of Distress Scale (SUDS) [15]. We thoroughly cleaned all equipment before and after each use and maintained an observer to monitor physical safety.

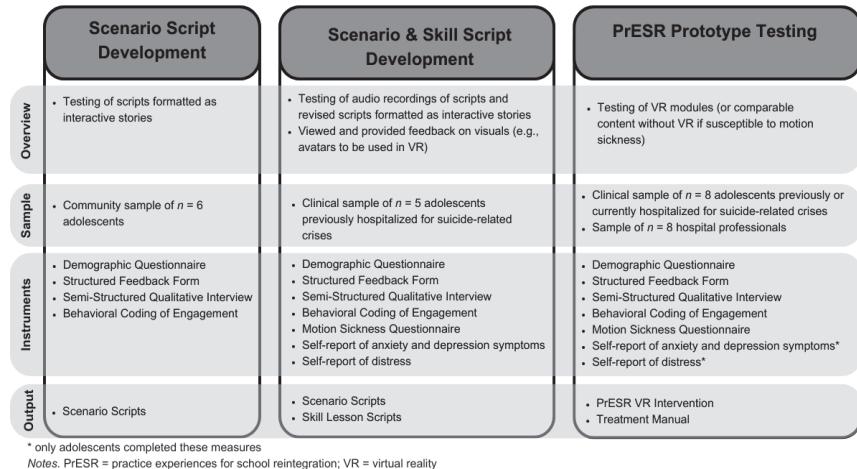


Figure 1. Stages of Intervention Development

During the first stage of intervention development, a community sample of $n=6$ adolescents tested the three scenarios by reading and responding to interactive stories developed in Inky [16]. We integrated findings from this feedback/playtesting and then presented to a clinical sample of $n=5$ previously hospitalized adolescents. These adolescents read and responded to the revised interactive stories; tested interactive audio recordings of lessons teaching affect regulation, cognitive restructuring, and problem solving; and provided feedback about their experience, as well as about some of the visuals developed for the intervention. We further refined scripts and content for inclusion in an initial computer (non-VR) prototype of the intervention. Due to resource constraints, we selected two of the scenarios and three of the skill lessons to include in the VR intervention. We prioritized the cafeteria and classroom scenes because they included a combination of supervised (classroom) and unsupervised (cafeteria) settings and reflected day-to-day experiences. (In contrast, a re-entry meeting would only occur once, if at all.) VR skill lessons focused on affect regulation, problem solving, and cognitive restructuring; the clinician would introduce CBT in-person instead of including it as a VR experience.

We internally tested and refined the initial computer build and then converted it to a VR prototype. We next recruited another clinical sample of adolescents, including adolescents previously ($n=4$) and currently ($n=4$) hospitalized for a suicide-related crises, as well as a sample of hospital professionals ($n=8$), to test and provide feedback about the VR scenarios and skill lessons (delivered in VR or using similar methods employed in earlier development sessions in cases of susceptibility for motion sickness).¹

3.2. Analyses and Results

A trained researcher documented participant response choices, as well as their verbal responses to each scenario prompt. We recorded major points and impressions in debrief summaries following each session that we integrated into subsequent presentations and/or prototypes. We documented recommendations for changes to scenes and skills at all stages, as well as solutions to glitches or errors embedded in the syntax or application and also integrated these into subsequent sessions. Examples of recommendations for scenarios from community participants included toning down the way bullying was presented and adding more positive response options. Examples of recommendations from adolescents with lived experiences included identifying the best time to practice the skill in a given scenario and making the scenarios and skills more interactive.

We evaluated participant ratings from behavioral coding to gauge adolescents' willingness to engage within the scenarios and their variation. Findings supported variability in behavioral responses in our clinical samples, with scores for engagement,

¹ Technical glitches disrupted testing for one hospital professional, so a combination of VR and components of the intervention was presented in this case.

emotional control, reflection, and comprehension ratings ranging from 3 (moderate) to 5 (very high); scores for reaction time ranging from 2 (low) to 4 (high); scores for provocativeness ranging from 2 (low) to 5 (very high); and scores for verbalizations and affect ranging from the 1 (very low or positive affect) to 5 (very high or negative affect).

3.3. Final Intervention

We developed the final intervention based on the iterative feedback from participants, internal testing, and recommendations from the VR development team. PrESR is designed to be delivered by trained clinicians adhering to fidelity checklists and the PrESR treatment manual. Sessions include rapport building, goal setting, VR experience(s), and a clinical debrief. The VR experiences include up to three skill lessons and two practice scenarios. Cybersickness following the first VR experience is assessed and clinicians teach adolescents to use SUDS to gauge distress throughout the intervention. Following completion of the VR sessions, the clinician and adolescent collaborate to develop a school-focused safety planning intervention to share with schools (with permission from patients and families).

4. Conclusion and Future Directions

Findings from this study informed the development of PrESR, a novel intervention for teaching therapeutic skills to adolescents hospitalized for suicide-related crises. Our next step is to evaluate the intervention's feasibility and potential for teaching skills in an inpatient setting. A pilot optimization trial (NCT05934396) is currently underway, in which we are recruiting adolescents (ages 13–18) hospitalized for a suicide-related crisis at two child/adolescent inpatient psychiatric hospitals. Although outcomes of PrESR are still to be determined, the model aims to help adolescents increase their tolerance to difficult social experiences and change their thinking patterns, ultimately supporting a smoother transition back to school and contributing to suicide prevention among previously hospitalized adolescents [9].

References

- [1] Youth risk behavior survey data [Internet]. 2022 [cited 2024 Jun 14]. Available from: <https://www.cdc.gov/yrbs>
- [2] Plemmons G, Hall M, Doupnik S, Gay J, Brown C, Browning W, et al. Hospitalization for suicide ideation or attempt: 2008–2015. *Pediatrics*. 2018;141(6):e20172426. doi: <https://doi.org/10.1542/peds.2017-2426>
- [3] Yard E, Radhakrishnan L, Ballesteros MF, Sheppard M, Gates A, Stein Z, et al. Emergency department visits for suspected suicide attempts among persons aged 12–25 years before and during the COVID-19 pandemic—United States, January 2019–May 2021. *MMWR Morb Mortal Wkly Rep*. 2021;70(24):888–94. doi: <https://doi.org/10.15585/mmwr.mm7024e1>
- [4] Chung DT, Ryan CJ, Hadzi-Pavlović D, Singh SP, Stanton C, Large MM. Suicide rates after discharge from psychiatric facilities: a systematic review and meta-analysis. *JAMA Psychiatry*. 2017;74(7):694–702. doi: <https://doi.org/10.1001/jamapsychiatry.2017.1044>
- [5] Marraccini ME, Pittelman C, Toole EN, Griffard MR. School supports for reintegration following a suicide-related crisis: a mixed methods study informing hospital recommendations for schools during discharge. *Psychiatr Q*. 2022;93(1):1–37. doi: <https://doi.org/10.1007/s11126-021-09942-7>
- [6] Tougas A, Houle A, Leduc K, Frenette-Bergeron É, Marcil K. Framework for successful school reintegration after psychiatric hospitalization: a systematic synthesis of expert recommendations. *Psychol Sch*. 2023;60(3):793–813. doi: <https://doi.org/10.1002/pits.22714>
- [7] Hayes C, Palmer V, Hamilton B, Simons C, Hopwood M. What nonpharmacological therapeutic interventions are provided to adolescents admitted to general mental health inpatient units? A descriptive review. *Int J Men Health Nurs*. 2019;28(3):671–86. doi: <https://doi.org/10.1111/ijn.12575>
- [8] Frazier E, Thompson A, Wolff J, Hunt J. Implementing evidence-based psychotherapy in an adolescent inpatient setting. *Brown Univ Child Adolesc Behav Lett*. 2016;32(5):1 – 6. doi: <https://doi.org/10.1002/cbl.30121>
- [9] Marraccini M, Anonick R, Delgaty L, Middleton T, Toole E, Ying J, et al. Practice experiences for school reintegration (PrESR): iterative development of a virtual reality intervention to supplement cognitive behavioral therapy for hospitalized youth. Forthcoming 2024.
- [10] Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. *J Med Internet Res*. 2015;17(1):e4055. doi: <https://doi.org/10.2196/jmir.4055>

- [11] Hubal RC, Fishbein DH, Sheppard MS, Paschall MJ, Eldreth DL, Hyde CT. How do varied populations interact with embodied conversational agents? Findings from inner-city adolescents and prisoners. *Comput Hum Behav*. 2008;24(3):1104–38. doi: <https://doi.org/10.1016/j.chb.2007.03.010>
- [12] Golding JF. Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. *Brain Res Bull*. 1998;47(5):507–16. doi: [https://doi.org/10.1016/S0361-9230\(98\)00091-4](https://doi.org/10.1016/S0361-9230(98)00091-4)
- [13] Bouchard S, Berthiaume M, Robillard G, Forget H, Daudelin-Peltier C, Renaud P, et al. Arguing in favor of revising the simulator sickness questionnaire factor structure when assessing side effects induced by immersion in virtual reality. *Front Psychiatry*. 2021;12:739742. doi: <https://doi.org/10.3389/fpsyg.2021.739742>
- [14] Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care*. 2007;45(5 Suppl 1):S22–31. doi: <https://doi.org/10.1097/01.mlr.0000250483.85507.04>
- [15] Wolpe J. Subjective units of distress scale (SUDS). In: APA PsycTests. 1969. doi: <https://doi.org/10.1037/t05183-000>
- [16] Inky [Internet]. Inkle Ltd; [cited 2024 Jun 14]. Available from: <https://www.inklestudios.com/inklewriter/>

The Analysis of Digital Biomarkers for Identifying and Predicting Frailty and Healthy Elderly through Machine Learning

Stefano DE GASPARI ^{a,b,1}, Irene Alice CHICCHI GIGLIOLI ^{a,c,1}, Chiara PUPILLO^{a,b,},
Elena SAJNO ^{a,b}, Daniele DI LERNIA ^{a,c}, Alessandro CAPRIOTTI ^d, Paolo
OLIVETTI^d and Giuseppe RIVA ^{a,c,e}

^a*Humane Technology Lab, Università Cattolica del Sacro Cuore, Largo Gemelli, 1,
20121 Milan, Italy*

^b*Department of Computer Science, University of Pisa, Largo B. Pontecorvo, 3 56127
Pisa*

^c*Department of Psychology, Università Cattolica del Sacro Cuore, Largo Gemelli 1,
20100 Milan, Italy*

^d*Research and Development Unit, Tech4care srl, Via Primo Maggio 20, 60131 Ancona*

^e*Applied Technology for Neuro-Psychology Laboratory, IRCCS Istituto Auxologico
Italiano, Via Magnasco 2, 20149 Milan, Italy*

ORCiD ID: Stefano De Gaspari <https://orcid.org/0009-0006-6083-2134>

Abstract. This study explores using an AI model to identify and predict frailty in elderly individuals by analyzing digital biomarkers of head and hand movements during virtual reality (VR) tasks. A sample of 20 participants, divided equally into healthy and frail groups, was assessed using traditional questionnaires and VR-based cognitive activities. The study utilized machine learning (ML) algorithms, including Decision Tree, Random Forest, and Logistic Regression, to analyze the digital biomarkers and predict the health status of the subjects. The results demonstrated that the Logistic Regression model achieved an accuracy of 0.83 and a ROC-AUC of 0.83, indicating its reliability in classifying frailty and healthy elderly. The research highlights the potential of combining digital biomarkers and VR with ML techniques to detect frailty conditions, suggesting a novel assessment modality that could enhance early interventions and improve the quality of life for the elderly.

Keywords. Frailty, Virtual reality, Digital biomarker, Artificial intelligence, Machine learning, Elderly

1. Introduction

Frailty in elderly people is a clinical condition characterized by a decline in physiological reserves and function across multiple organ systems, resulting in increased vulnerability to stressors [1]. This condition leads to a heightened risk of adverse health outcomes, including falls, hospitalization, and other clinical conditions such as Mild Cognitive Impairment (MCI) and Alzheimer's disease (AD) [1-3]. Frailty is not synonymous with aging itself but is a distinct and critical aspect of health in older adults [4]. Its traditional assessment includes questionnaires (yes/no), such as the Tilburg Frailty Indicator (TFI) [5] that considers three main components: a) physicals, including items on walking difficulty, maintaining balance difficulty, poor hearing and eyesight, lack of strength in the hands, and physical tiredness, b) psychological, including items on lack of memory, feeling down or nervous or anxious in the last month and c) socials, including items on whether the subject lives alone, on the feeling of lack of social support and whether the subject feels to receive on sufficient support from other people [1].

Despite their validity, these questionnaires have limitations in terms of: a) lack of standardization because there isn't a universal standard for frailty assessment, b) social

¹ These authors contributed equally
Corresponding Author: Stefano De Gaspari, s.degasperi@phd.unipi.it

desirability bias, c) applicability in clinical practice; some instruments indeed, despite being valid, require specific time and expertise to be applied, thus making complex their use in daily practice [6, 7], and d) fail to detect behavioral aspects in daily life activities. These limitations can slow down or hinder the diagnostic process, when timely detection and intervention at this early stage can significantly alter the trajectory of cognitive and physical decline, enhancing the quality of life for affected individuals [8].

In this context, virtual reality (VR) provides a unique platform for simulating real-world environments and assessing cognitive function in a controlled setting [9]. VR-based cognitive assessments can evaluate spatial navigation, memory recall, and executive functions, offering insights into specific cognitive domains for the evaluation of different health conditions such as: Parkinson's and Alzheimer's disease, vascular dementia, frailty, and others [10-12]. VR not only allows collecting behavioral data such as correct/incorrect answers and execution times, but can also capture digital biomarkers (DB): these can be defined as indicators of physiological and behavioral data, collected by digital devices, that can offer insights into health, disease progression, and treatment responses in a real-world setting [13]. In VR, DB can be identified by head, hand, and gait movements [14, 8]. In addition, the vast amount of data generated by digital biomarkers can be analyzed through Machine Learning (ML) algorithms providing simulations and predictions of behavioral performance trends and patterns [18]. Previous studies on DB in aging populations showed the feasibility of tracking head, hands, and gait movements in VR to support the assessment of clinical conditions, such as MCI [8], and predicting future risks of cognitive impairments [14]. Indeed, decreasing speed, longer stride time, and increasing stride-to-stride variability represent some key features in predicting MCI cognitive impairments, related to executive functions [14]. In frailty's condition, studies using ML techniques showed good accuracy in the assessment of frailty [15], however, the combination of DB in VR and ML hasn't been addressed yet. According to these premises, this study explores the development and potential of an ML model including and integrating head and hands digital biomarkers and VR for identifying and predicting frail and healthy elderly.

2. Methods

10 healthy subjects (F= 1, M=9; Mean Age= 71,3, SD=5,21) and 10 frailty subjects (F= 10; Mean Age= 75,2, SD=7,66) over 65 years of age have been screened using Tilburg Frailty Indicator (TFI) [5] and the Mini-Mental State Examination (MMSE) [16].

The inclusion criteria for the subjects were: (a) age 65 years or older; (b) absence of cognitive impairment as measured by the MMSE (Score >24) [16]; (c) absence or presence of frailty (according to the group), with a score of 5 or less on the TFI [5]; and (d) autonomy or mild dependence in activities of daily living measured by Index of Independence in Activities of Daily Living (ADL) and Instrumental Activities of Daily Living Scale (IADL). The exclusion criteria for both subject groups were: (a) history of stroke or acute TIA; (b) presence of MCI; (c) presence of dementia; (d) presence of physical and/or functional deficits preventing the use of VR; (e) comorbidity with severe neurological diseases or with psychiatric conditions not under pharmacological treatment or in the acute phase; and (f) history of head trauma with loss of consciousness.

2.1 Virtual Reality

After the screening, each subject was introduced to the virtual reality visor (PICO 4; <https://www.picoxr.com/it>) and detailed explanations of the system's components and interactions were provided. The subject then wore the visor and entered the virtual environment, starting with familiarization activities, such as grasping and dragging an object and selecting objects with open and closed hands. After the familiarization phase, the subject completed four cognitive activities in serious game environments: a) an activity on visual-spatial abilities by dragging a mole from point A to point B avoiding obstacles; b) an activity on working and short-term memory by matching pairs of cards; c) an activity on attention and inhibitory control by selecting moles and avoiding penguins, and d) an activity on memory and visual-spatial skills by recalling a shopping

list and selecting items in a virtual supermarket. The total duration of the four activities in VR was 20 minutes.

2.2 Data and ML models

Head and hand movements have been recorded during the VR activities with a frequency of 20 Hz. The development of the ML model, integrating head and hands digital biomarkers for identifying and predicting frailty and healthy elderly people, included the following phases:

Dataset preparation phase:

- Creation of a starter dataset containing all participants' head and hand movements data and TFI total score.
- Size reduction of the dataset calculating means and standard deviations for each movement data. This made the data more manageable for the next steps [17].
- Data pre-processing, checking missing values (NaN) to avoid bias in the ML models, and data standardization, ensuring that the features present all had the same scale of values.

ML algorithms selection and model's phase:

- The following algorithms have been selected: Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), Support Vector Machine (SVM), Logistic Regression (LR) and K-Nearest Neighbors (KNN) [17].
- The dataset has been divided into two parts: one part of the dataset, the train set, was used to train all ML models to identify the 2 classes we were interested in, frailty and healthy. With the remaining part of the dataset, the test set, we evaluated the ability of the algorithms to predict the two conditions on which they were previously trained, healthy and frailty.
- Next, a Repeated Stratified K-Fold Cross-Validation with 3 folds and 3 repetitions has been applied to further evaluate the performance of the ML models and obtain a more reliable and robust evaluation. After this evaluation, the algorithms with the best performance have been selected and hyperparameters have been applied to re-train the models and improve their performance. For the hyperparameters, Grid Search and Randomized Search have been used. After this final training and testing phase the performance of the selected models using the same metrics as before have been evaluated.
- To evaluate the performance of the models, the following metrics have been used: Accuracy, AUC-ROC, Precision, Recall and F1 score. Accuracy measures the percentage of correct predictions, while AUC-ROC assesses the model's ability to distinguish between classes. Precision, recall and F1 score provide further information on the balance between false positives and false negatives.

3. Results

The analysis results of the three machine learning models (Random Forest, Decision Tree, and Logistic Regression) are shown below.

The Random Forest model (**Table 1**) showed moderate performance with all metrics (accuracy, precision, recall, F1 score, specificity and ROC-AUC) being 0.67. This indicates that the model has a balanced but not excellent ability to correctly classify both healthy (class 0) and frail elderly subjects (class 1).

Table 1. Random Forest (*Class 0 = Healthy elderly subjects; Class 1=Frail elderly subjects)

	Precision	Recall	F1 score	Support
Class 0*	0.67	0.67	0.67	3
Class 1*	0.67	0.67	0.67	3
Accuracy			0.67	6
Macro avg	0.67	0.67	0.67	6
Weighted avg	0.67	0.67	0.67	6

The Decision Tree model (**Table 2**) showed perfect performance with all metrics equal to 1.00. This indicates that the model correctly classified all healthy elderly subjects (class 0) and frail elderly subjects (class 1) without errors.

Table 2. Decision Tree (*Class 0 = Healthy elderly subjects; Class 1=Frail elderly subjects)

	Precision	Recall	F1 score	Support
Class 0*	1.00	1.00	1.00	3
Class 1*	1.00	1.00	1.00	3
Accuracy			1.00	6
Macro avg	1.00	1.00	1.00	6
Weighted avg	1.00	1.00	1.00	6

The Logistic Regression model (**Table 3**) showed good overall performance with an accuracy of 0.83 and a ROC-AUC of 0.83. The precision for class 1 (frail subjects) is 0.75, indicating that 75% of the positive predictions were correct, while the recall is 1.00, indicating that all frail subjects were correctly identified. The specificity of 0.67 indicates that the model correctly classified 67% of healthy subjects.

Table 3. Logistic Regression (*Class 0 = Healthy elderly subjects; Class 1=Frail elderly subjects)

	Precision	Recall	F1 score	Support
Class 0*	1.00	0.67	0.80	3
Class 1*	0.75	1.00	0.86	3
Accuracy			0.83	6
Macro avg	0.88	0.83	0.83	6
Weighted avg	0.88	0.83	0.83	6

4. Discussion and Conclusion

This study aimed to explore the development and potential of ML models including and integrating head and hands digital biomarkers and VR for the identification and prediction of frail and healthy elderly. According to the results, the Logistic Regression model performed well (accuracy of 0.83, ROC-AUC of 0.83). The precision for the class of frail subjects was 0.75, while the recall reached 100%, indicating that the model was able to correctly identify all frail subjects. This balance between precision and recall, reflected in an F1 score of 0.86, makes Logistic Regression a reliable model for this type of classification. In contrast, the Decision Tree model showed perfect performance with all metrics equal to 1.00. However, such perfect performance can often be an indicator of overfitting. This makes the model less reliable for practical applications, where it is essential that the model can generalize well to new data. The Random Forest model showed modest performance (accuracy of 0.67, ROC-AUC of 0.67). However, the precision for the class of frail subjects was 0.67, lower than that of the Logistic Regression, and the specificity of 67% suggests that the model has a fair ability to correctly identify healthy subjects.

In conclusion, the analysis showed that Logistic Regression is the most reliable model for the classification of healthy and frail elderly subjects. This model offers a good balance between accuracy, precision, recall, and specificity while providing greater interpretability and stability than other models. This preliminary study aimed to evaluate the possibility of creating a novel assessment modality for the health status of the elderly. The aim was to use behavioral data (digital biomarkers) collected during some VR tasks, (connected to cognitive functions such as attentional capacity, visuospatial, etc.), and the results of a paper and pencil questionnaire (Tilburg Frailty Indicator - TFI) to train ML models in predicting two health conditions associated with the elderly, namely the healthy condition and the frailty condition.

Despite the good performance in predicting the two classes by one of the trained algorithms (Linear Regression), this preliminary study has several limitations: firstly, the study involved a total of 20 elderly subjects, which corresponds to a small sample size for ML. Consequently, future studies could consider recruiting a significantly larger sample to obtain more generalizable results and to further test our models. Furthermore, the VR behavioral data (digital biomarkers) in their raw form corresponded to time series, which were then transformed into means and standard deviations to make the dataset more manageable. However, future studies might consider the idea of working directly with this type of data, which, being very extensive and detailed, could provide relevant

information for trained ML models by identifying patterns that are fundamental for the prediction of the elderly's health condition.

Funding

This work has been co-funded by the European Union—Next Generation EU under Project: “A novel public-private alliance to generate socioeconomic, biomedical and technological solutions for an inclusive Italian ageing society” (Age-IT), CUP: B83C22004800006.

References

- [1] Fried LP, Tangen CM, Walston J, et al.; Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci.* 2001;56(3):M146–M156. doi: <https://doi.org/10.1093/gerona/56.3.M146>
- [2] Ávila-Funes JA, Amieva H, Barberger-Gateau P, et al. Cognitive impairment improves the predictive validity of the phenotype of frailty for adverse health outcomes: the Three-City Study. *J Am Geriatr Soc.* 2009;57(3):453–461. doi: <https://doi.org/10.1111/j.1532-5415.2008.02136.x>
- [3] Robertson DA, Savva GM, Kenny RA. Frailty and cognitive impairment: a review of the evidence and causal mechanisms. *Ageing Res Rev.* 2013;12(4):840–851. doi: <https://doi.org/10.1016/j.arr.2013.06.004>
- [4] Ahmed N, Mandel R, Fain MJ. Frailty: an emerging geriatric syndrome. *Am J Med.* 2007;120(9):748–753. doi: <https://doi.org/10.1016/j.amjmed.2006.10.018>
- [5] Mulasso A, Roppolo M, Gobbens RJ, Rabaglietti E. The Italian version of the Tilburg Frailty Indicator: analysis of psychometric properties. *Res Aging.* 2016;38(8):842–863. doi: <https://doi.org/10.1177/0164027515606192>
- [6] Dent E, Kowal P, Hoogendoijk EO. Frailty measurement in research and clinical practice: a review. *Eur J Intern Med.* 2016;31:3–10. doi: <https://doi.org/10.1016/j.ejim.2016.03.007>
- [7] Dolenc E, Rotar-Pavlič D. Frailty assessment scales for the elderly and their application in primary care: a systematic literature review. *Slov J Public Health.* 2019;58(2):91–98. doi: <https://doi.org/10.2478/sjph-2019-0012>
- [8] Seo K, Kim J, Oh DHK, Ryu H, Choi H. Virtual daily living test to screen for mild cognitive impairment using kinematic movement analysis. *PLoS One.* 2017;12(7):e0181883. doi: <https://doi.org/10.1371/journal.pone.0181883>
- [9] Riva G, Wiederhold BK, Chirico A, Di Lernia D, Mantovani F, Gaggioli A. Brain and virtual reality: what do they have in common and how to exploit their potential. *Annu Rev CyberTherapy Telemed.* 2018;16:3–7. doi: https://doi.org/10.3233/978-1-61499-852-5_3
- [10] [10] Kubicki A. Functional assessment in older adults: should we use timed up and go or gait speed test? *Neurosci Lett.* 2014;577:89–94. doi: <https://doi.org/10.1016/j.neulet.2014.06.055>
- [11] Negut A. Cognitive assessment and rehabilitation in virtual reality: theoretical review and practical implications. *Rom J Appl Psychol.* 2014;16(1):1–7. doi: <https://doi.org/10.2478/rjap-2014-0001>
- [12] Paliokas I, Kalamaras E, Votis K, Doumpoulakis S, Lakka E, Kotsani M, et al. Using a virtual reality serious game to assess the performance of older adults with frailty. In: GeNeDis 2018: Geriatrics. Cham: Springer International Publishing; 2020. p. 127–139. doi: https://doi.org/10.1007/978-3-030-24913-9_10
- [13] Cavedoni S, Chirico A, Pedroli E, Cipresso P, Riva G. Digital biomarkers for the early detection of mild cognitive impairment: artificial intelligence meets virtual reality. *Front Hum Neurosci.* 2020;14:245. doi: <https://doi.org/10.3389/fnhum.2020.00245>
- [14] Martín-Gonzalo JA, Pulido-Valdeolivas I, Wang Y, Wang T, Chiclana-Actis G, Algarra-Lucas MDC, et al. Permutation entropy and irreversibility in gait kinematic time series from patients with mild cognitive decline and early Alzheimer's dementia. *Entropy.* 2019;21(9):868. doi: <https://doi.org/10.3390/e21090868>
- [15] Oliosi E, Guede-Fernández F, Londral A. Machine learning approaches for the frailty screening: a narrative review. *Int J Environ Res Public Health.* 2022;19(14):8825. doi: <https://doi.org/10.3390/ijerph19148825>
- [16] Magni E, Binetti G, Padovani A, Cappa SF, Bianchetti A, Trabucchi M. The Mini-Mental State Examination in Alzheimer's disease and multi-infarct dementia. *Int Psychogeriatr.* 1996;8(1):127–134. doi: <https://doi.org/10.1017/S1041610296004922>
- [17] Tsai CF, Chen CC, Wu EHK, Chung CR, Huang CY, Tsai PY, Yeh S. A machine-learning-based assessment method for early-stage neurocognitive impairment by an immersive virtual supermarket. *IEEE Trans Neural Syst Rehabil Eng.* 2021;29:2124–2132. doi: <https://doi.org/10.1109/TNSRE.2021.3111654>
- [18] Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. *Science.* 2015;349(6245):255–260. doi: <https://doi.org/10.1126/science.aaa8415>

Virtual Reality Training for Spatial Perspective-Taking in Children with ASD

Luna MADDALON^{a,1}, Maria Eleonora MINISSI^a, Thomas D. PARSONS^{b,c}, Amaia HERVÁS-ZÚÑIGA^c and Mariano ALCAÑIZ^a

^a *Human-Centered Technology Research Institute (HUMAN-Tech). Polytechnic University of Valencia, Valencia, Spain*

^b *Grace Center, Edson College, Arizona State University, Tempe, AZ 85281, USA*

^c *Computational Neuropsychology and Simulation, (CNS) Arizona State University, Tempe, AZ 85281, USA*

^d *Fundació de Docencia Y Recerca Mútua Terrassa (MTA) - Grup Salut Mental Infanto Juvenil, Terrassa, Barcelona, Spain*

ORCID ID: Luna MADDALON <https://orcid.org/0000-0002-2789-6703>

Abstract. Autism Spectrum Disorder (ASD) has profound deficits in social-cognitive skills, between them spatial perspective-taking ability, requiring alternative treatment methods in the context of increasing diagnoses. Emerging technologies like virtual reality (VR) offer promising avenues for intervention due to their immersive nature and ability to dynamically adapt to individual needs, potentially addressing the heterogeneity of ASD symptom severity. This study introduces a novel approach by integrating an adaptive VR serious game into traditional interventions, aiming to complement existing methods and improve treatment efficacy. The study involved 15 participants diagnosed with ASD severity level 1. A VR task was designed to immerse participants in a virtual playground environment, where they interacted with an avatar and responded to spatial perspective-taking questions. Over the course of four sessions, participants underwent adaptive training, with task difficulty adjusted incrementally based on performance. In addition to the difficulty level, the system was governed by applied behavioral analysis intervention principles tailored to each participant's performance. Data analysis revealed a significant improvement in performance metrics between sessions and decreased use of aids as participants demonstrated greater competence, suggesting a successful learning process. Data should be interpreted with caution, given the small sample size. However, findings underscore the potential of VR-based adaptive interventions to address social-cognitive deficits in ASD. Indeed, the innovative methodology, based on traditional psychological principles and frameworks, could provide a solid framework for future studies on adaptive training.

Keywords. Autism spectrum disorder; virtual reality; serious game; perspective taking; children

1. Introduction

Children diagnosed with Autism Spectrum Disorder (ASD) often exhibit significant challenges in social-cognitive skills, essential for navigating daily interactions and fostering effective peer relationships. Guided by the essential role of social-cognitive development in societal integration, interventions targeting ASD children should prioritize nuanced aspects of the Theory of Mind, including understanding the perceptions of others. A key component is the spatial perspective-taking ability, which involves occupying the other's place and understanding the relative position of objects [1]. Egocentric errors occur when someone chooses his or her own perspective instead of the other person's, which is frequently observed in ASD children.

With the escalating global ASD prevalence rates, the demand for alternative treatment methods is required. Recent studies have endeavored to bridge this gap by integrating novel technologies into traditional interventions [2]. Indeed, given the limitations of conventional methods, technology systems may hold promise to complement training based on objective evidence [3]. Notably, the immersive nature of

¹ Corresponding Author: Luna Maddalon, lmaddal@htech.upv.es

VR fosters learning transfer to real-life scenarios by engaging users through a sense of physical presence in a virtual realm. As an embodied skill, spatial perspective-taking finds VR technologies a good tool for training and skills consolidation. Another benefit of using such technologies lies in their ability to dynamically adapt the training content according to the needs of each individual [4]. An adaptive intervention consists of a sequence of clinical decision rules that guide how and when to change the dose and type of treatment based on given measures [3]. This approach fits with the heterogeneity of symptom severity observed in ASD, as it requires the implementation of individualized treatment approaches that target specific symptom configurations for each individual.

Recognizing the advantages of early intervention, we developed an adaptive VR task to improve spatial perspective-taking skills in ASD children. The VR task was based on the traditional Theory of Mind Task Battery (ToMTB) [5]. The purpose of the current study is twofold: (a) to delineate a novel methodological paradigm in spatial perspective-taking training, and (b) to conduct a preliminary validation study to explore the evolution of spatial perspective-taking skills throughout four sessions, laying the foundation for future interventions. To investigate this latter point, analyses were conducted concerning the accuracy achieved by participants in the adaptive VR task. Thus, it was expected that throughout the sessions, children would achieve better performance as the need for support decreased.

2. Methods

2.1. Participants

Fifteen children diagnosed with ASD severity level 1, according to the Autism Diagnostic Observation Schedule 2 (ADOS-2), were enrolled in the study. The sample, aged 6 to 8 years ($M = 91.73$ months, $sd = 8.71$), comprised predominantly males ($N = 14$) and one female participant. Participants were recruited voluntarily from Fundació Assistencial Mútua Terrassa, a specialized clinical center for ASD in Spain. Before the study, caregivers were provided with detailed information regarding the experimental protocol and requested to provide written consent for their children's participation. The study received approval from the Ethical Committee of the Fundació Assistencial Mútua Terrassa (ID: P/23-111/).

2.2. Spatial Perspective-Taking VR Serious Game

The experimental study was conducted within an ad hoc setting equipped with a portable device setup, facilitating a semi-immersive experience [6]. The system incorporated an 86" screen monitor, a computer, and the Azure Kinect DK, enabling interaction by projecting the participant's body as a non-filled virtual shape within the virtual environment. The user's mobility area, delineated with marks on the floor, was set as two square meters at 1.5 m in front of the monitor.

The VR spatial perspective-taking task goal was to enhance social-cognitive skills in ASD children, focusing on two key components: (a) understanding that perceptions, such as visual or auditory cues, provide access to knowledge and influence beliefs and (b) recognizing that individuals may have different perspectives based on their spatial orientation. Clinical experts agreed that effective training should last at least one month, encompassing approximately a weekly session with incremental adjustments in-game difficulty adjusted by performance while keeping the number of sessions constant to maintain control over this variable. Finally, if positive linear performance was observed, the total number of sessions was adjusted to the three difficulty levels and a learning consolidation session.

Employing Unity® software, the VR task featured an interaction between participants and an avatar character resembling a child in a virtual playground environment. The task commenced with a static scene featuring the avatar and an object situated within the playground environment for a duration of 10 seconds to allow the participant to process the visual stimuli. Following, participants were exposed to an introductory scene wherein the avatar expressed a search for a specific object within the playground. An example of a target object for spatial positioning is a ball located in front

of the fountain (see **Figure 1a**). From this moment onward, following each interaction and VR animation, a five-second static interval was designated to allow for information processing and retention. Afterward, participants were prompted by the avatar to answer two sequential questions about the spatial positioning of the target object from the participant's and avatar's perspectives, respectively. After each question, participants' virtual bodies were displayed alongside response options, allowing them to indicate their selection by positioning their hands within the collision area of the chosen option for a duration of 1.5 seconds. Answer options visually described the object's location relative to various park features. Participants were allowed to visually explore the virtual environment to answer questions. By rule, the participant underwent level 1 in the first session, which featured two response alternatives; level 2 instead comprised three response options (see **Figure 1b**); and level 3 presented four options.

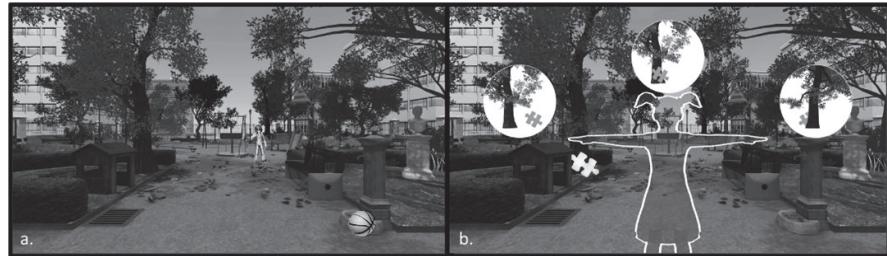


Figure 1: Screen captions of the task. a. avatar in the playground; b. projected user shape and answer options.

While the task structure remained consistent throughout, the system automatically adapted the training method. The system was ruled by principles of applied behavioral analysis intervention tailored to each participant's performance through the provision of aids, feedback mechanisms, correct answer elucidation, reinforcement strategies, and adjustments in difficulty levels. Refer to **Figure 2** to visualize the serious game flow explained here. Aids were administered upon the expiration of the 10-second response window after each question, entailing the repetition of the question up to two additional times. Therefore, participants had 30 seconds in total to answer questions within a maximum of two aids. Immediate feedback was provided following each response, with correct answers indicated by color-coded answer option borders (green) and corresponding auditory cues, while incorrect responses were marked in red. Whenever participants did not respond within the allocated 30-second total timeframe after each question, the correct answer was revealed only visually. After answering both questions, if an incorrect answer was selected, the avatar verbally and visually provided a correct answer elucidation at the end of the task. Afterward, a “completed game” visual feedback was presented to indicate the game was finished. Visual and auditory reinforcement “Good Job!” was administered after the task if participants achieved a 50% accuracy rate (see **Figure 2**). Following this criterion, participants progressed to higher difficulty levels in subsequent sessions, while those failing to attain the 50% accuracy threshold repeated the same level. The 50% rule was adopted due to the small number of responses, which prevented the establishment of a majority rule. Additionally, even if a participant progressed to a higher level, the same questions were repeated iteratively, allowing continuous performance evaluation over time. Thus, the 50% rule served as a symbolic threshold for level advancement. In fact, accuracy analysis was based on overall performance metrics per session, weighted by difficulty level. Finally, two versions of each difficulty level were created to mitigate potential learning effects, featuring distinct stimuli and locations. The alternate version was given whenever participants repeated the same difficulty level in subsequent sessions. Each version of the difficulty level encompassed unique stimuli chosen a priori, with levels stratified according to the number of response options presented.

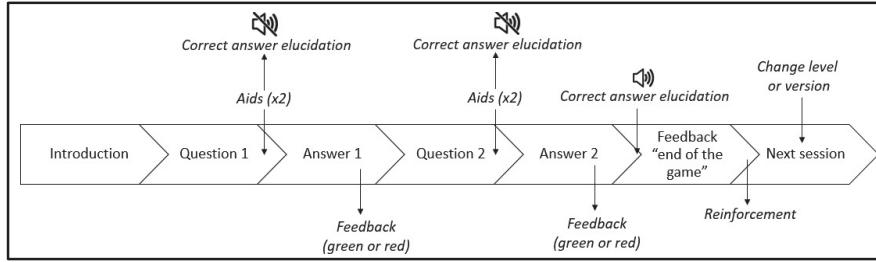


Figure 2: Spatial perspective-taking VR serious game flow

2.3. Procedure

Relevant data were collected within the framework of a broader experimental methodology. The following disclosures are pertinent to the parameters of this study. At the onset of the first session, participants underwent a familiarization phase with the semi-immersive VR system to verify their ability to interact with the system. Participants were asked to choose between male or female virtual body shapes to facilitate their meta-self-recognition in the virtual body, which provided interaction with the system. Subsequently, participants engaged in the VR spatial perspective-taking serious game. This procedure was repeated four times in a one-month period.

2.4. Data analysis

Data analysis was performed utilizing the JASP 0.18.3 (University of Amsterdam, 2013). Accuracy, quantified as the number of correct responses, was aggregated for each participant for each session. To weight accuracy by session difficulty level, accuracy scores were multiplied according to the level of difficulty participants were subjected to for each session. Therefore, aggregate accuracy scores encompassed the cumulative sum of accuracy within a session (maximum 2 points) multiplied by the difficulty level (1-3). The data distribution's normality assumption was assessed using Shapiro-Wilk's test.

Whether assumptions were violated, Kendall's tau-b correlation non-parametric test was employed to determine the relationship between aggregated accuracy and session progression and between aggregated accuracy and received aids. All analyses were subjected to two-tailed tests, with a significance level set at $\alpha = 0.05$.

3. Results

The analysis revealed a strong, positive correlation between aggregated accuracy and sessions, signifying statistical significance ($\tau b(58) = .616$, $p < .001$). Furthermore, a moderately negative correlation between aggregated accuracy and aids was observed to be statistically significant ($\tau b(58) = -.368$, $p = .001$).

4. Discussion

This study aimed to investigate the effectiveness of an adaptive VR task in enhancing spatial perspective-taking skills in children diagnosed with ASD. Building upon the traditional ToMTB, a novel VR task specifically tailored to target spatial perspective-taking abilities has been developed. The task was designed to engage participants in a semi-immersive VR environment, leveraging the portable device system, which had been previously validated for such applications [6]. Objectives encompassed delineating a new methodological approach to spatial perspective-taking training and conducting a preliminary validation study to assess the task's efficacy over four sessions. The findings seem to provide valuable insights into the potential of VR-based interventions to address social-cognitive challenges in ASD, laying the groundwork for future research and intervention strategies in this area.

The findings suggested a gradual improvement in performance accuracy over sessions. This trend reflected the increase in correct task answers, consistent with difficulty levels. Such improvements, observed longitudinally, could underscore the effectiveness of adaptive training to promote spatial perspective-taking skills. Furthermore, the results revealed a remarkable tendency to diminish reliance on aids consistently with improved performance. This observation suggests that although aids were initially beneficial during the learning phase, their need gradually decreased as participants possibly demonstrated greater competence in task completion. This phenomenon aligns with the concept of educational scaffolding, wherein supportive structures are gradually eliminated as learners gain competence and autonomy in their learning endeavors [7]. In summary, the designed task showed a notable enhancement in performance throughout the training regimen, concomitant with a reduction in aid administration, corresponding to the observed performance improvements.

Although our study has provided valuable insights, these findings should be interpreted cautiously, as limitations exist. The primary limitation of this study is its small sample size, as the data analysis and interpretation were conducted in a preliminary study. Indeed, such did not allow the establishment of the relationships among the three variables analyzed. Furthermore, children's ability to take a spatial perspective was not previously assessed to establish a baseline measure of their initial ability. We anticipate that future research with a larger sample and clinically standardized measures will yield more reliable explanations. Additionally, future studies are necessary to explore and develop a broader range of perspective-taking tasks, building upon the foundational findings of this study.

5. Conclusion

This study presents a pioneering methodological approach utilizing VR to tackle deficits in social-cognitive skills prevalent in children with ASD. Specifically, the aim was to devise an immersive and engaging platform tailored to augment spatial perspective-taking capabilities. Indeed, the innovative methodology, grounded on traditional psychological principles and structure, could provide a robust framework for future adaptive training studies. Noteworthy performance enhancements were observed across successive sessions, coupled with a concurrent decline in aid administration, underscoring the possible promising efficacy of this serious game in refining spatial perspective-taking skills.

References

- [19] Parsons S. Learning to work together: designing a multi-user virtual reality game for social collaboration and perspective-taking for children with autism. *Int J Child Comput Interact.* 2015;6:28–38. doi: <https://doi.org/10.1016/j.ijcci.2015.12.002>
- [20] Barletta VS, Caruso F, Di Mascio T, Piccinno A. Serious games for autism based on immersive virtual reality: a lens on methodological and technological challenges. In: Proceedings of the 12th International Conference in Methodologies and Intelligent Systems for Technology Enhanced Learning; 2022; Salamanca, Spain. Cham: Springer; 2023. p. 181–195. doi: https://doi.org/10.1007/978-3-031-20617-7_23
- [21] Alcañiz M, Maddalon L, Minissi ME, Sirera M, Abad L, Chicchi Giglioli IA. Intervenciones tecnológicas adaptativas para el trastorno del espectro autista: una revisión bibliográfica. Medicina (B Aires). 2022;82:54–58. ISSN 1669-9106.
- [22] Maddalon L, Minissi ME, Parsons TD, Hervás Zúñiga A, Alcañiz Raya M. Adaptive virtual reality systems employed for children with autism spectrum disorder interventions: a systematic review. *J Med Internet Res.* Forthcoming. doi: <https://doi.org/10.2196/57093>
- [23] Hutchins TL, Prelock PA, Chace W. Test-retest reliability of a theory of mind task battery for children with autism spectrum disorders. *Focus Autism Other Dev Disabil.* 2008;23(4):195–206. doi: <https://doi.org/10.1177/1088357608322995>
- [24] Maddalon L, Minissi ME, Cervera-Torres S, Hervás A, Gómez-García S, Alcañiz M. Multimodal interaction in ASD children: a usability study of a portable hybrid VR system. In: Proceedings of the International Conference on Human-Computer Interaction; 2023; Copenhagen, Denmark. Cham: Springer; 2023. p. 614–624. doi: https://doi.org/10.1007/978-3-031-35681-0_40
- [25] Munshi A, Biswas G, Baker R, Ocumpaugh J, Hutt S, Paquette L. Analysing adaptive scaffolds that help students develop self-regulated learning behaviours. *J Comput Assist Learn.* 2023;39(2):351–368. doi: <https://doi.org/10.1111/jcal.12761>

Physiological Response to a Virtual Reality Simulation for Preoperative Stress Inoculation

Catherine PROULX^{a,1}, Michael S. D. SMITH^a, Vincent GAGNON SHAIGETZ^a,
Gabrielle S. LOGAN^b, Jordana L. SOMMER^{b,c}, Pamela HEBBARD^{d,e},
Kristin REYNOLDS^{c,f}, W. Alan MUTCHE^b, Natalie MOTA^{c,f,g},
Jessica L. MAPLES-KELLER^h, Rakesh C. ARORA^e, and Renée
ELGABALAWY^{b,c,d,f,g}

^a National Research Council of Canada

^b Department of Anesthesiology, Perioperative and Pain Medicine, University of Manitoba, Winnipeg, MB, Canada

^c Department of Psychology, University of Manitoba, Winnipeg, MB, Canada

^d CancerCare Manitoba, Winnipeg, MB, Canada

^e Harrington Heart and Vascular Institute, University Hospitals, Cleveland, OH, United States

^f Department of Psychiatry, University of Manitoba, Winnipeg, MB, Canada

^g Department of Clinical Health Psychology, University of Manitoba, Winnipeg, MB, Canada

^h Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, GA, United States

ORCID ID: Catherine Proulx <https://orcid.org/0000-0002-1776-2082>

Abstract. This paper describes the development of a novel immersive virtual reality (VR) simulation designed to reduce preoperative state anxiety in patients undergoing breast cancer surgery. A custom interactive VR simulation allows participants to experience the setting of an operating room and the key preoperative stages, all the way through the administration of general anesthesia. Interactivity is provided through a self-avatar with which the various simulated medical personnel interact directly. We evaluate the capacity of the simulation to induce an emotional response as measured by the participants' galvanic skin response (GSR). To our knowledge, this is the first fully interactive simulation of an oncology surgery induction procedure for stress inoculation, and the first preoperative VR study to measure emotional impact using GSR. Out of a larger trial, we analyzed 6 participants who had been randomized to the simulation group and for whom baseline and intra-simulation GSR data had been successfully acquired. Three-minute samples were compared for statistical difference with a 95% confidence interval on the mean. 5 out of 6 showed a statistically significant and visually noticeable increase in GSR, and participants reported a high sense of spatial presence. Early results are encouraging, showing that the described simulation can induce a physiological response consistent with the participants' subjective evaluation of presence. While this was a limited experiment, it provides a basis for a larger trial to be conducted in the future.

Keywords. Interactive simulation, virtual reality, preoperative anxiety, physiological monitoring

1. Introduction and Background

This paper describes the technical development of a novel immersive virtual reality (VR) simulation designed to reduce preoperative state anxiety (PSA) – anxiety and distress prior to a surgical intervention – in patients undergoing breast cancer surgery. We evaluate the capacity of an interactive simulation to induce an emotional response as measured by the participants' galvanic skin response (GSR). A previous paper on this initiative studied the feasibility and utility of the system with data from a case series [1] derived from a larger feasibility trial [2]. This complementary paper describes in detail

¹ Corresponding Author: Catherine Proulx, catherine.proulx@nrc.ca

the design of the simulation, especially as it pertains to factors influencing presence and immersion and measures the impact of the intervention on quantifiable physiological responses of a subset of the participants (who were randomized to the treatment VR group) in the feasibility trial.

Preoperative state anxiety has been associated with a range of poor postoperative health outcomes [3, 4]. A few studies have measured the impact of VR-based preoperative stress inoculation (i.e., pre-emptive exposure to stressful environments) with mixed results. Most systems use 360° videos [5-7] to simulate the surgery process, even though interactivity increases presence [8] and is considered more beneficial in the treatment of phobias [9]. To our knowledge, only one previous investigation utilized a patient-perspective interactive VR for surgery simulation in a pediatric setting [10]. The researchers reported no significant decrease in pre-operative anxiety but reported less use of post-operative pain management drugs in the VR treatment group.

2. Design of the Simulation

A custom VR simulation was designed to allow participants to experience the setting of an operating room and the key preoperative steps designed in accordance with operating room appearance and procedures at a large provincial tertiary care center in central Canada. The steps included: lying down on the operating table, having medical sensors attached to the self-avatar, a safety briefing, placing the oxygen mask, and ending with the administration of general anesthesia. While engaged in the simulation, the participant sits and reclines on a real-world surface whose position coincides with that of the virtual operating table. Further elements regarding the simulation are detailed in the feasibility protocol [2].

An iterative and collaborative development process incorporated the input of various subject matter experts. The simulation was built using Unity 2020.1.8 and delivered through an Oculus Rift S off-the-shelf headset. Using a VR headset tethered to a PC enables clinicians or research staff to monitor the user's activities (**Figure 1**). A few design elements were hypothesized as key for increasing the user's emotional response and sense of embodiment in the simulation.

2.1. Interactivity of the Self-avatar and User Agency

The user is represented by a female self-avatar miming their real-life upper-body movement. The position of the self-avatar's lower half is kept fixed with the legs hidden under a virtual blanket. On top of increasing the feeling of presence [11], using a self-avatar enables virtual characters in the simulation to interact with the user. Natural transitions triggered by user actions give the user a greater sense of control during the simulation. For instance, after having a chance to look around the operating room, the next stage of the simulation does not proceed until the participant physically lies down on the operating table. In cases where a patient would not have agency in the surgery process, transitions are triggered by the operator.



Figure 1. Clinician view of the surgery, including first person POV and additional top and side views

2.2. Multi-Sensory Scene Representative of Operating Room Conditions

Development team observations of pre-operative procedures, reference photos and videos, and input from medical experts guided the creation of a representative virtual operating room. To enhance immersion [12], attention was paid to multi-sensory aspects of the simulation including realistic sounds like monitor beeps and clink of surgical instruments as well as bright lighting.

2.3. Realism of the Virtual Medical Staff and Devices

Efforts were put on the appearance and behavior of the medical personnel avatars since unnatural motions or expressions can easily break the user's immersion [13]. Characters located closer to the user and character interactions with the user were prioritized as these would be most obvious. For example, it was important to optimize the animations of the nurse placing sensors on the user. To give primary characters a more natural appearance, eyes were animated to include gaze direction, blinking, and eye saccades. Multiple devices were attached to the self-avatar, including an intravenous port, a wristband, a blood pressure cuff, an oxygen mask, wires, and a photoplethysmography (PPG) sensor since these devices can induce claustrophobia and fear of pain [14].

3. Methods

A post-experiment analysis of biometric data acquired as part of a larger feasibility trial in which 12 participants were exposed to the VR simulation was undertaken [2]. Participants received the VR simulation 1-2 weeks prior to their scheduled breast cancer surgery. Control participants ($n=11$) were not analyzed for the purposes of the current study as they did not undergo VR exposure. GSR was selected as the physiological measure of choice, based on earlier research on phobia exposure therapy linking GSR with a sense of presence and with the desensitization process' effectiveness [15, 16]. GSR was measured (using Mindfield® eSense Skin Response) before the VR experiment to provide a baseline and then during the simulation itself.

Of the 12 participants in the intervention group, 6 yielded usable GSR data. For the other 6 participants (age range 46-64 (mean = 56)), GSR data was not acquired due to communication issues between the sensor and the recording app, or poor sensor contact with the skin; these participants were excluded from the analysis. All participants in the analyzed subset self-reported as female and had experienced a previous surgical procedure. Three-minute samples for both the baseline and the intra-simulation measurements were compared with a 95% confidence interval on the mean. In addition, a t-test statistical test was used to compare both samples. Data was also smoothed with a median filter with a 60 s window for visualization purposes.

The iGroup Presence Questionnaire [17] assessed the presence associated with the VR intervention, defined as "the sense of being in the virtual environment" along three subscales: involvement (range -12 to 12), spatial presence (range -15 to 15) and realism (range -12 to 12).

4. Results

Results of the analysis are shown in Table 1, and time series for each of the participants are presented in **Figure 2**. 5 out of 6 participants showed a statistically significant increase in GSR ($p < 1e-10$) with a large enough difference to be considered physiologically significant ($>1 \mu\text{S}$). Participant D showed only a small ($<0.1 \mu\text{S}$) increase in mean GSR: there is a statistically significant difference in the mean but given the noise on the measurements and the absence of a visually detectable step in the smoothed data, we estimated that this was most likely not clinically relevant.

Table 1. Confidence intervals on the mean, and p-values for individual t-tests between the baseline and the intra-simulation samples. Non-overlapping of the intervals indicates statistical significance.

Participant id	95% C.I. on the mean, baseline (μ S)	95% C.I. on the mean, intra-simulation (μ S)	Δ mean (μ S)	p
A	[2.82 2.85]	[5.39 5.53]	2.63	< 1e -10
B	[2.95 2.97]	[4.50 4.56]	1.57	< 1e -10
C	[0.87 0.90]	[1.87 1.92]	1.01	< 1e -10
D	[1.78 1.80]	[1.85 1.87]	0.07	< 1e -10
E	[4.21 4.33]	[5.40 5.56]	1.21	< 1e -10
F	[3.24 3.32]	[4.46 4.55]	1.22	< 1e -10

Responses to the Presence Questionnaire indicate moderate involvement (mean = 1) and moderate realism (mean = 1) scores, but high spatial presence (mean = 8).

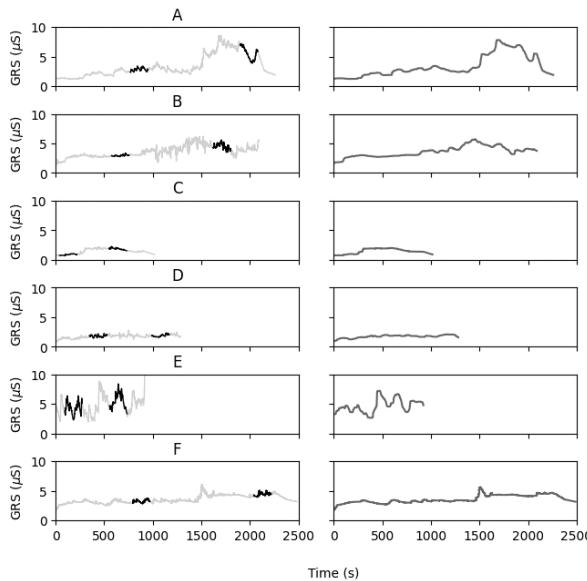


Figure 2 Time series for the 6 participants studied. Samples used for analysis are shown in black. The right column shows the smoothed data.

5. Discussion and Conclusion

This study describes key design elements of a VR simulation created to reduce preoperative state anxiety in patients undergoing breast cancer surgery. It also presents preliminary evidence that immersion in this simulation induces a measurable physiological response for most participants.

Three key design elements that could contribute to engaging and embodying virtual experiences are highlighted: interactivity and user agency, multi-sensory stimulation, and well-designed virtual characters and medical devices. While the impact of these factors has already been documented for a variety of virtual conditions, this is one of the first fully interactive simulations of an oncology surgery induction procedure for stress inoculation, and the first preoperative VR study to measure the emotional impact using GSR. Subjective evaluation of involvement, realism, and spatial presence indicate areas of improvement for the next version of the simulation.

Analysis of the physiological data reveals that the simulation appears to increase the mean GSR level for most participants. A limitation of the paper is the manual GSR acquisition process which led to data loss, and the lack of automatic synchronization with the VR data. Manual recording of the simulation start time was imprecise, so we were limited to only analyzing samples of the signals to account for that imprecision. The lack of an active control group is another significant limitation; an active control group was initially planned but had to be eliminated due to recruitment challenges [2].

This study allowed us to observe some indicators of the simulation's effectiveness in eliciting genuine emotional reactions. While not definitive proof of efficacy, it facilitated the identification of potential improvements in the experiment design, the simulation, and the GSR acquisition processes to improve the next arm of the study.

References

- [1] Sommer K, Reynolds P, Hebbard M, Smith MS, Mota N, Mutch WAC, et al. Preoperative virtual reality to expose patients with breast cancer to the operating room environment: feasibility and pilot case series study. *JMIR Form Res.* 2024;8(1):e46367. doi: <https://doi.org/10.2196/46367>.
- [2] El-Gabalawy JL, Sommer P, Hebbard K, Reynolds GS, Logan MS, Smith MS, et al. An immersive virtual reality intervention for preoperative anxiety and distress among adults undergoing oncological surgery: protocol for a 3-phase development and feasibility trial. *JMIR Res Protoc.* 2024;13(1):e55692. doi: <https://doi.org/10.2196/55692>.
- [3] Skeppholm M, Fransson R, Hammar M, Olerud C. The association between preoperative mental distress and patient-reported outcome measures in patients treated surgically for cervical radiculopathy. *Spine J.* 2017;17(6):790–798. doi: <https://doi.org/10.1016/j.spinee.2016.02.037>.
- [4] Utrillas-Compaied A, Basilio J, Tebar-Martínez AJ, Asúnsolo-Del Barco Á. Does preoperative psychologic distress influence pain, function, and quality of life after TKA? *Clin Orthop Relat Res.* 2014;472(8):2457–2465. doi: <https://doi.org/10.1007/s11999-014-3570-5>.
- [5] Bekelis D, Calnan N, Simmons N, MacKenzie TA, Kakoulides G. Effect of an immersive preoperative virtual reality experience on patient-reported outcomes: a randomized controlled trial. *Ann Surg.* 2017. doi: <https://doi.org/10.1097/SLA.00000000000002094>.
- [6] Chiu PL, Li H, Yap KYL, Lam KMC, Yip PLR, Wong CL. Virtual reality-based intervention to reduce preoperative anxiety in adults undergoing elective surgery: a randomized clinical trial. *JAMA Netw Open.* 2023;6(10):e2340588. doi: <https://doi.org/10.1001/jamanetworkopen.2023.40588>.
- [7] Dehghan F, Jalali R, Bashiri H. The effect of virtual reality technology on preoperative anxiety in children: a Solomon four-group randomized clinical trial. *Perioper Med.* 2019;8:1. doi: <https://doi.org/10.1186/s13741-019-0116-0>.
- [8] Slater M. Immersion and the illusion of presence in virtual reality. *Br J Psychol.* 2018;109(3):431–433. doi: <https://doi.org/10.1111/bjop.12305>.
- [9] Guillén-Sanz H, Checa D, Miguel-Alonso I, Bustillo A. A systematic review of wearable biosensor usage in immersive virtual reality experiences. *Virtual Reality.* 2024;28(2):74. doi: <https://doi.org/10.1007/s10055-024-00970-9>.
- [10] Eijlers R, Dierckx B, Staals LM, Berghmans JM, van der Schroeff MP, Strabbing EM, et al. Virtual reality exposure before elective day care surgery to reduce anxiety and pain in children: a randomized controlled trial. *Eur J Anaesthesiol.* 2019;36(10):728–737. doi: <https://doi.org/10.1097/EJA.0000000000001059>.
- [11] Steed A, Pan Y, Zisch F, Steptoe W. The impact of a self-avatar on cognitive load in immersive virtual reality. In: Proc IEEE Virtual Reality (VR). 2016:191–198. doi: <https://doi.org/10.1109/VR.2016.7504689>.
- [12] Diemer J, Alpers GW, Peperkorn HM, Shiban Y, Mühlberger A. The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Front Psychol.* 2015;6:111605. doi: <https://doi.org/10.3389/fpsyg.2015.00026>.
- [13] Zibrek K, Kokkinara E, McDonnell R. The effect of realistic appearance of virtual characters in immersive environments—Does the character's personality play a role? *IEEE Trans Vis Comput Graph.* 2018;24(4):1681–1690. doi: <https://doi.org/10.1109/TVCG.2018.2794638>.
- [14] Arakelian E, Laurssen E, Öster C. Older patients' worries in connection with general anesthesia and surgery: a qualitative study. *J PeriAnesth Nurs.* 2018;33(6):822–833. doi: <https://doi.org/10.1016/j.jopan.2018.01.005>.
- [15] Notzon S, Deppeermann S, Fallgatter A, Diemer J, Kroczeck A, Domschke K, et al. Psychophysiological effects of an iTBS modulated virtual reality challenge including participants with spider phobia. *Biol Psychol.* 2015;112:66–76. doi: <https://doi.org/10.1016/j.biopsych.2015.10.003>.
- [16] Wiederhold BK, Jang DP, Kim SI, Wiederhold MD. Physiological monitoring as an objective tool in virtual reality therapy. *Cyberpsychol Behav.* 2002;5(1):77–82. doi: <https://doi.org/10.1089/109493102753685908>.
- [17] Berkman MI, Çatak G. I-group presence questionnaire: psychometrically revised English version. *Mugla J Sci Technol.* 2021;7:1–10. doi: <https://doi.org/10.22531/muglajsci.882271>.

Inclusive Cognitive Commercial Training for Children: A Preliminary Evaluation of the Kokoro Kids App in Enhancing Visuospatial Abilities in Autistic and Neurotypical Children

Maria Eleonora MINISSI^{a,1}, Luna MADDALON^a, Andrea LOPEZ-CASTILLO^b, Rocio SAEZ^b, Carles PONS^b, Elena PARRA-VARGAS^a, Mariano ALCAÑIZ^a

^a*Instituto Universitario de Investigación en Tecnología Centrada en el Ser Humano (HUMAN-tech), Universitat Politècnica de Valencia, Valencia, Spain*

^b*Lernin Games, Valencia, Spain*

ORCiD ID: Maria Eleonora Minissi <https://orcid.org/0000-0001-6326-0609>

Abstract. This paper presents a preliminary evaluation of an inclusive cognitive training program using the Kokoro Kids app. The study aimed to assess the effectiveness of specific games within the Kokoro Kids app in improving visuospatial abilities in both neurotypical and autistic children. The methodology involved 10 children aged 5 to 8 years, with five children diagnosed with autism spectrum disorder and five neurotypical children. Participants engaged with 11 Kokoro Kids games targeting visuospatial skills over a period of 15 consecutive days. Visuospatial abilities were assessed using 5 standardized subtests from the NEPSY-II and DTVP-3 batteries, administered before and after the training period. The overall findings indicated significant improvements in one NEPSY-II subtest, irrespective of the groups, suggesting positive visuospatial enhancements. Additionally, there were non-significant tendencies toward improvement in two of the other subtests in both groups. In terms of game performance, both groups showed varying degrees of accuracy improvement across game levels in 8 games, with progressive and constant difficulty. Future research should focus on expanding the sample size, and refining the methodological procedure, such as controlling for additional variables to better understand the impact of such digital cognitive training. Further investigation involving diverse populations is also recommended to enhance the inclusivity of the training.

Keywords. Autism spectrum disorder, typical development, visuospatial abilities, cognitive training, commercial video game

1. Introduction

The increasing accessibility of technologies to children, such as mobile phones and tablets, offers a chance to transform cognitive training and education. These technologies facilitate remote access to playful training materials at any time. Despite the growing interest in the effectiveness of digital games for boosting cognitive skills in both neurotypical and neurodivergent children [e.g., 1,2], few studies have explored whether the cognitive improvement offered by specific digital games in the neurotypical population can be replicated in the neurodivergent population [e.g., 3], whilst, to our knowledge, there is a lack of studies applying the same game-based cognitive training in the two populations to assess inclusive training of cognitive abilities.

In this preliminary study, we evaluated several games from the Kokoro Kids commercial app to assess their effectiveness in enhancing visuospatial abilities in a small group of neurotypical children, that is, children with typical development (TD), and in a small group of children with autism spectrum disorder (ASD). Kokoro Kids is a commercial app for tablets encompassing more than 100 games. Specific games targeting

¹ Corresponding Author: Maria Eleonora Minissi, meminissi@i3b.upv.es.

visuospatial skills were chosen and proposed to the two groups of children to study whether they could enhance visuospatial skills through playful training. To assess their effectiveness in enhancing visuospatial abilities, standardized visuospatial subtests were administered before and after training to both groups. Additionally, we checked whether game accuracy improved over game levels with both constant and progressive difficulty. In games with constant difficulty, higher accuracy would indicate that game exposure may improve visuospatial abilities. In games with progressive difficulty, improvements in visuospatial skills, as reflected by increased accuracy, would be driven by the increasing difficulty of the games.

2. Methods

2.1. Participants

In Ten Spanish children aged between 5 and 8 years who have never played Kokoro Kids participated in this preliminary study. Five of these children with no diagnosis or risk of psychiatric, affective, learning, or neurodevelopmental disorders were classified as children with TD (age in months, $M = 82.00$; $sd = 14.30$; females = 3). The other five children formed the ASD group (age in months, $M = 87.80$; $sd = 14.60$; females = 2). They had a previous diagnosis of level 1 ASD with no intellectual disability. Three children of the ASD group had a language delay; therefore, Spanish comprehension was checked in the entire group. Participants were recruited through Lernin Games, which provided families with online coupons upon completion of the study. Before the study began, caregivers received detailed information about the experimental protocol and were asked to provide written consent for their children's participation. The study has been approved by the Ethical Committee of the Polytechnic University of Valencia (P06 24-04-2024).

2.2. Kokoro Kids Games for Visuospatial Training

Eleven digital cognitive games within the Kokoro Kids commercial app, identified by independent expert evaluators as targeting different narrow skills of visuospatial ability according to the Cattell-Horn-Carroll theory [4], were used in the study. The games varied in complexity, with 8 featuring progressively increasing difficulty and 3 remaining constant. They also have a varied number of levels (see **Table 1**).

Table 1. Game taxonomy: type of difficulty, number of levels, and narrow visuospatial skill involved.

Type	Game	Levels	Narrow visuospatial skill
Constant difficult	Shapes Tangram	28	visual association
	Body Tangram	12	visual association
	Spot the Differences	9	visual search and discrimination ability
	Candy Colors	39	visual association
	Half Shapes	9	visual association
	Shapes Classify	40	visual association
Progressive difficulty	Hide & Seek	20	visual search and discrimination ability
	Car Jam Puzzle	29	sequential spatial ability
	The Treasure Map	25	visual association
	Mice & Coding	19	sequential spatial ability
	Draw the Path	19	sequential spatial ability

The eleven games are briefly described below (see **Figure 1**).



Figure 1. Screen captures of the eleven digital cognitive games: (a) Shapes Tangram; (b) Body Tangram; (c) Spot the Differences; (d) Candy Colors; (e) Half Shapes; (f) Shape Classify; (g) Hide & Seek; (h) Car Jam Puzzle; (i)The Treasure Map; (j) Mice &Coding; (k) Draw the Path.

The Shapes and Body Tangram games aimed at creating shapes by combining different geometric pieces. The Spot the Difference game aimed to identify differences between similar stimuli. The Candy Colors game aimed to drag candies to the corresponding tubes. The Half Shapes game aimed at matching pieces to complete different shapes. Similarly, the Shape Classify game required recognizing and classifying simple geometric shapes. The Hide & Seek game required actively searching for hidden characters in a scenario. The Car Jam Puzzle game aimed at moving elements to allow the car to leave the parking lot. The Treasure Map game consisted of selecting and manipulating the correct path piece and placing it strategically to reach the treasure. The Mice and Coding game aimed at interacting with arrows to form a sequence of movements to guide the mouse to the cheese. Finally, the Draw the Path game aimed at drawing lines to connect islands and create a visual path to reach the goal.

2.3. Visuospatial Measures Before and After Training

Five visuospatial subtests from the NEPSY-II [5] and DTVP-3 [6] were implemented online to be administered remotely and to check potential visuospatial enhancements in the two groups after the training. Both batteries have also been validated with children with ASD. Participants of both groups completed the 5 subtests in a randomized order before and after the training period.

Both NEPSY-II subtests (Route Finding and Arrow) measure sequential spatial ability. Specifically, implementing the Route-Finding subtest was intended to investigate the understanding of the directionality of visuospatial relationships. With the Arrow subtest, the ability to judge the orientation of a line was tested.

Two of the three DTVP-3 subtests measure visual search and discrimination ability. Namely, the Figure-Ground subtest investigated the ability to extract hidden figures from a complex and confusing background. The Visual Closure subtest tested the ability to mentally provide the missing parts of a given figure. Finally, the Form Constancy subtest measured the visual association ability by recognizing a given figure hidden in a set of different distractor stimuli.

2.4. Experimental Procedure

During the recruitment phase, an online survey completed by the children's caregivers gathered information on the participants' sociodemographic characteristics, and the Child Behaviour Checklist (CBCL) was administered to check comorbidities.

Pre-tests on visuospatial abilities were administered online for both groups upon starting the training period. During the experimental training, all participants engaged with the 11 commercial games of the Kokoro Kids app for 15-20 minutes daily over a

maximum of 15 consecutive days. To play games, participants needed a tablet with at least iOS 13 or Android 18. The experimentation was carried out remotely. Children were allowed to choose the order of the games during training sessions to encourage self-engagement. However, to standardize the training experience, each game level could only be played once; after completing all levels, a game was no longer available. Additionally, at the end of the daily training, caregivers answered online control questions to identify any factor that might have influenced the cognitive training.

Finally, visuospatial ability post-tests were administered online for both groups at the end of the training period. During game training, the system tracked participants' accuracy for each level, which was computed using the number of errors and prompts.

2.5. Data Analysis

Data was analyzed using IBM SPSS Statistics 2 (IBM, 2018). The normality assumption of the data distribution was evaluated using Shapiro-Wilk's test. Likewise, Levene's test checked the homogeneity assumption. Differences between participants' ages and differences in the duration of the training period in the two groups were analyzed using independent samples t-tests. Potential enhancements in visuospatial abilities after training were assessed through repeated measures ANOVA on raw scores of each visuospatial measure using testing (pre-test/post-test) as within subject factor, and group (ASD and TD) as between subject factor. Post hoc analyses were run using Bonferroni correction. Finally, the potential improvement in game accuracy over levels has been checked through simple linear regression analysis for each game per group.

3. Results

First, the two groups did not differ in age ($t(8) = -.634; p = .544$), and participants engaged in gaming for a training period that ranged from 9 to 15 days, with no significant differences in duration between the groups ($t(8) = .893; p = .398$). Control questions posed to caregivers showed that most of the time children of both groups played games at home and did not play similar games that could affect intervention.

With respect to visuospatial abilities assessed pre- and post-training, a significant main effect of digital cognitive training was noted in the Arrows subtest irrespective of the group ($F(1,8) = 7.578; p = .025$). In the post-tests, participants achieved higher scores ($M = 26.900; SE = .975$) than they did prior to the digital cognitive training ($M = 20.500; SE = 2.454$).

Furthermore, despite the lack of statistical significance, performance in both groups showed a tendency to improve in the Visual Closure and Figure-Ground subtests. Visual improvement was also observed in the Route Finding and Form Constancy subtests for the neurotypical group only, although these were not statistically significant (see Figure 2).

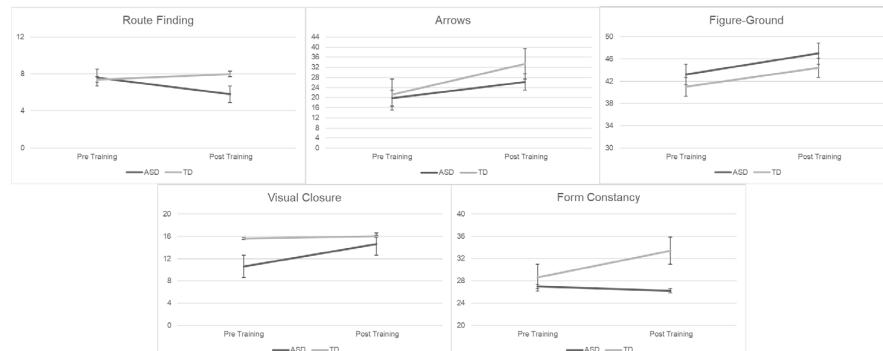


Figure 2. Visual representation of mean scores and typical error differences between pre- and post-testing in the two groups in the visuospatial subtests.

In terms of gaming accuracy over levels, statistically significant results are shown in **Table 2**.

Table 2. Games in which the regression analysis was statistically significant.

Game	Group	Regression coefficient β	Adj-R squared	F Statistics
Shapes Tangram	TD	.019	.090	$F(1,48) = 5.927; p = .019^*$
	ASD	.038	.421	$F(1,98) = 72.822; p < .001^{***}$
Spot the Differences	TD	.060	.029	$F(1,48) = 2.533; p = .119$
	ASD	.081	.247	$F(1,98) = 33.465; p < .001^{***}$
Candy Colors	TD	.049	.137	$F(1,48) = 4.399; p = .041^*$
	ASD	.167	.772	$F(1,98) = 337.1; p < .001^{***}$
Half Shapes	TD	.103	.349	$F(1,48) = 20.049; p < .001^{***}$
	ASD	.130	.669	$F(1,98) = 76.355; p < .001^{***}$
Shapes Classify	TD	.007	-.010	$F(1,48) = .508; p = .477$
	ASD	.167	.772	$F(1,98) = 337.1; p < .001^{***}$
Car Jam Puzzle	TD	.060	.029	$F(1,48) = 2.533; p = .119$
	ASD	.081	.247	$F(1,98) = 33.465; p < .001^{***}$
Treasure Map	TD	-.006	-.006	$F(1,48) = .471; p = .496$
	ASD	.167	.772	$F(1,98) = 337.1; p < .001^{***}$
Draw the Path	TD	-.006	-.006	$F(1,48) = 0.471; p = .496$
	ASD	.029	.083	$F(1,98) = 9.956; p = .002^{**}$

* $p < .05$, ** $p < .01$, *** $p < .001$.

4. Discussion

This preliminary study wants to assess the potential of digital games from the Kokoro Kids app to train specific cognitive skills, such as visuospatial abilities, in an inclusive cohort of neurotypical and neurodivergent children.

Despite the small sample size, positive cognitive enhancements have been found in the Arrow subtest after training, regardless of the group. In addition, there was a visual tendency to improve in post-tests of both groups in Visual Closure and Figure-Ground, although a statistically significant difference was not found. These findings may suggest that increasing the sample size could aid in achieving visuospatial cognitive enhancements between pre- and post-training in the three tests mirrored by statistically significant differences. Moreover, visual and statistical improvements in the three tests should be related to visuospatial cognitive training provided by the Kokoro Kids app due to the control that caregivers carried out over training on influencing factors, and the good test-retest reliability coefficient of chosen tests after 15 days [Brooks; Hammill]. Nevertheless, future studies should keep control over more variables besides involving more participants.

Regarding accuracy improvement over game levels, level switching explained different portions of accuracy variance in the games, ranging from .83 to .772. In the games with constant difficulty, both ASD and TD children showed a limited improvement in accuracy on Shapes Tangram, while only ASD children improved similarly on Spot the Difference. No accuracy improvement was found in the Body Tangram for either group. For games with progressive difficulty, both groups of children improved their accuracy over game levels on Candy Colors and Half Shapes, particularly in the latter. Finally, only ASD children reported improved accuracy over game levels on Shapes Classify, Car Jam Puzzle, Treasure Map, and Draw the Path. No accuracy improvement over levels was found in Hide & Seek and Mice & Coding. Among these, Shape Classify and Treasure Map were the games in which accuracy variance was most explained by level switching in ASD. Overall, the effect of level switching on accuracy was reduced for both groups in most games, suggesting that factors besides visuospatial

training play a role in accuracy modulation (e.g., IQ, engagement, anxiety, affective states, sociodemographic). Furthermore, there is evidence that children with ASD may prefer negative feedback stimulation over positive feedback, sometimes choosing the wrong answer on purpose [7]. In addition, some games may have provided either floor effects in performance at their first levels or ceiling effects, particularly in children with TD. Nevertheless, cognitive training may still occur even if it is not reflected in accuracy improvement, and visuospatial testing represents the most precise option to assess improvements after training.

The present findings suggest that the same cognitive training may improve visuospatial skills in an inclusive manner although at different extents in a cohort of neurotypical and neurodivergent children. However, further research with a larger sample size, more precise evaluations, and control variables needs to be performed.

5. Conclusion

These preliminary findings highlight the potential of specific digital games within the Kokoro Kids app to enhance targeted aspects of visuospatial cognition in a diverse group of children. Future research should expand the sample size and experimental methodology, as well as include a broader range of neurodiversity to thoroughly investigate how selected games can improve visuospatial abilities, building on the results and directions reported here.

References

- [1] Rezayi S, Tehrani-Doost M, Shahmoradi L. Features and effects of computer-based games on cognitive impairments in children with autism spectrum disorder: an evidence-based systematic literature review. *BMC Psychiatry*. 2023;23(1):2. doi: <https://doi.org/10.1186/s12888-022-04550-6>
- [2] Gashaj V, Dapp LC, Trninic D, Roebers CM. The effect of video games, exergames and board games on executive functions in kindergarten and 2nd grade: an explorative longitudinal study. *Trends Neurosci Educ*. 2021;25:100162. doi: <https://doi.org/10.1016/j.tine.2021.100162>
- [3] Fantasia V, Markant DB, Valeri G, Perri N, Ruggeri A. Memory enhancements from active control of learning in children with autism spectrum disorder. *Autism*. 2020;24(8):1995–2007. doi: <https://doi.org/10.1177/1362361320935951>
- [4] McGrew KS. CHC theory and the human cognitive abilities project: standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*. 2009;37(1):1–10. doi: <https://doi.org/10.1016/j.intell.2008.08.004>
- [5] Brooks BL, Sherman EM, Strauss E. NEPSY-II: A developmental neuropsychological assessment. *Child Neuropsychol*. 2009;16(1):80–101. doi: <https://doi.org/10.1080/09297040903146966>
- [6] Hammill DD, Pearson NA, Voress JK. DTV-P-3: Developmental test of visual perception. Austin, TX: Pro-Ed; 2014.
- [7] Zakari HM, Ma M, Simmons D. A review of serious games for children with autism spectrum disorders (ASD). In: Proceedings of the Serious Games Development and Applications: 5th International Conference, SGDA 2014; 2014 Oct 9–10; Berlin, Germany. Springer International Publishing; 2014. p. 93–106. doi: https://doi.org/10.1007/978-3-319-11623-5_9

Virtual Standardized LLM-AI Patients for Clinical Practice

Patrick G. KENNY^{a,1} and Thomas D. PARSONS^{b,a}

^aGeer AI

^bComputational Neuropsychology & Simulation (CNS) Laboratory, Arizona State University (ASU)

ORCID ID: Thomas D. Parsons <https://orcid.org/0000-0003-0331-5019>

Abstract. The integration of artificially intelligent (AI) conversational agents acting as virtual standardized patients into the practice for therapists represents a significant advancement in the field of medical training and simulation. Over the past four decades, the development of cognitive architectures, chat bots, behavior trees, and state machines has aimed to create believable, human-like behavior in embodied agents. These virtual humans have been utilized in many applications, from non-player characters in video games to virtual patients for clinical medical diagnosis, to conversational therapists in training and entertainment. Studies have shown that these systems can be effective in training and intervention. However, numerous challenges remain in designing and implementing these systems effectively. One of the most promising developments in this area is the advent of Large Language Models (LLMs) and associated AI technologies such as Natural Language Processing (NLP), Natural Language understanding (NLU) and Generation (NLG), voice and behavior recognition and generation and integration of multi-modal inputs for emotion and gesture tracking to enhance behavior generation. LLMs have demonstrated remarkable capabilities in generating human-like text and engaging in meaningful conversational discourse, but have limited application in creating realistic interactive 3D virtual human characters.

Keywords. Virtual standardized patients, artificial intelligence, large language models, mental health, psychology

1. Introduction

The goal of this paper is to discuss the potential of using advanced AI technologies such as LLMs to enhance the design and implementation of embodied conversational virtual standardized patients (VSPs). LLMs can address the challenges of knowledge acquisition, reasoning and decision making, realistic behavior generation, and the integration of diverse cultural and psychological attributes to develop virtual standardized patients. These characters are agents that use text or voice input, perform question answering and carry out interactive conversations while expressing the output in dialog, gestures and facial expressions and can simulate a psychological condition with high degree of fidelity to enhance training.

In our previous research we developed VSPs for clinical therapist training [1] we developed characters that exhibited clinical characteristics (e.g., ADHD; PTSD; depression) [3,4,5, 6]. These characters were tested with many clinical trainees. The results showed that virtual humans can be used as virtual patients to elicit the appropriate questions for a therapist to make a proper diagnosis [7]. The development of these systems at the time used early research versions of technology from statically question answering [8] and voice recognition systems to enable a 3D character to generate appropriate animations. Technology has come a long way since then and VSP characters today can greatly benefit from using the state of the art LLMs.

¹ Corresponding Author: Patrick Kenny, geerkenny@gmail.com

2. LLMs and AI Agents for virtual standardized patients design and development

The power in using LLMs as part of the development and implementation process for realistic characters has changed the way they will be created for now and the future. LLM's with their vast knowledge and ability to generate relevant data, text, pictures and voices in real time has modified the underlying architecture for realistic virtual humans. However, it is not without some challenges and new research areas to integrate LLM's and more advanced AI into these architectures. Although there is no standard virtual human architecture, it can be divided into three main areas, 1) the input from a user or the world, 2) the processing and decision making of the data and the output 3) in dialog, voice and animations [9,10].

2.1. AI Generated Challenges

The 3 main layers in a virtual human architecture are broken down into modules that communicate with each other. The way these modules work and are integrated for realtime conversational agents is an ongoing topic of research. The game industry has made great strides in making Non-player characters (NPC), with various techniques for decision making, but these are mostly scripted to follow the story or player. The film industry has the luxury of being able to process the data in non-real time allowing refinement of the result. For real-time interactive embodied conversational characters the modules can be defined as: **Input:** Speech recognition, text input, vision, EEG signals. **Processing:** Interpretation and classification of the input data, decision making systems such as finite state machines (FSM), behavior trees (BT), rule based systems or statistical question answering systems. Short term and long term memory, integration of emotion and personalities, reasoning based on character, history and persona. More advanced virtual humans can use cognitive architectures such as SOAR [11][12]. **Output:** Voice Generation, face animations, lip sync, autonomic behaviors, gesture animation and postures, gaze and rapport. **External modules:** Connections to Language models, databases, tutoring systems, saving and loading states and changing behavior over time and even a director agent that can give the other agents tasks, goals or behaviors.

LLM's can be integrated into or replace many modules by setting up the system to generate appropriate responses from specific LLM prompts. Ideally there would be a tighter coupling between the LLM and the architecture, but generating and modifying general LLM's in real-time is unfeasible at this time. The goal is to make the LLM be a part of the system, not just use it as a chatbot to generate text output to be displayed through a 3D character.

2.2. Knowledge Acquisition and Input

The data used in the system can be either input data from sensors and modules, such as voice recognition or text input and even vision and EEG data or LLM knowledge used as rules. The input data would be processed and used for the decision modules. One of the main issues with any autonomous system is acquiring accurate and relevant knowledge from subject matter experts or encoding the tasks by hand or in code. LLM's can play a significant role in helping with this process by asking the LLM to generate the characters knowledge. Experts can verify the output if it's a critical data or just use it as it from specific or verified prompts sent to the LLM. Virtual patients need detailed information about various medical conditions to simulate realistic symptoms and responses. More specifically in the case of therapist training the conditions could be hidden and only revealed through dialog, behavior and deep question answering to the virtual patient. This process involves not only gathering medical facts but also understanding the nuanced ways in which symptoms manifest and progress in patients. LLM's like GPT have an enormous corpus of human knowledge and understanding in dialog and discourse, and can talk in simple language.

In the design process one can ask the LLM to generate a Persona of a human character based on some simple input prompt (See examples below). Others have used LLMs to fabricate a history, life stories, talking style, emotional state and many other

attributes that can be used to create a virtual human patient[13,14,15]. This enhances the design process as it can be performed instantly and iterated upon to get the right character data for the situation. Additionally, the character could be tailored and personalized for a teaching objective verified by experts or fall within some pre-defined set of rules or constraints for the learning objective. Sometimes it's also nice to just get a set of human like responses that would give a student unexpected challenges that push them into deeper decision making.).

2.3. Decision Making

In embodied conversational agents, like a chatbot, selecting the right answer is the main goal. There have been several types of systems developed for virtual human characters. 1) Scripted interactions, where responses are selected based on the structure of the interview, if a question is not asked or answered, then the student moves to the next question. 2) Question /answering systems select a best answer generated statically based on what is asked. 3) Finite State Machines and Behavior Trees select a response or action based on a set of pre-defined conditions to move from one state to another.

BT's have a hierarchy of states and are more adaptable and manageable. 4) Rule Based Systems, like Expert systems that mimic human expert decision-making from a selected set of rules encoded by humans. 5) Cognitive architectures system, like SOAR or ACT-R [16] model more psychology aspects of how the brain works and offer more promise, but the complexity of getting all the parts correct requires experts in computers and psychology to develop full models. 6) Neural Network systems and Large Language Models (all knowledge is trained and embedded into the system) can perform many tasks. LLM's have been shown to person well in conversations [17,18].

Each of these approaches has their issues, but it has become clear that LLM's have almost replaced the other methods or are used to generate the knowledge and rules, and can also perform planning and reasoning. LLM's can provide almost all of the expert knowledge from the core neural network, and has an understanding of human dialog and discourse and even medical conditions from the DSM. The task now is how to properly format the query prompts to get the relevant responses out of the LLM, there will be limitations that arise, and additional modules might need to be added to the agent. Several techniques exist to get the model to perform better, from chain of thought prompts to one-shot prompts that would get the LLM to reason more about the question through intermediate steps [19,20]. Another limitation of LLM's is limited or non-existent short and long term memory and tracking state over several sessions. Generative Agent work has been shown to address these limitations by enhancing agents with LLM's to interact with each other, make decisions, record short term and long term memory and generate human like behavior in simulated worlds [21] and companies [22].

2.4. Realistic Behavior Generation

Once the LLM generates the dialog for the virtual patient conversation, it can augment it with character state annotations (See examples below), the AI Agent will then need to generate the corresponding believable behaviors. This includes generating or playing the appropriate facial expressions, gestures animations, and vocal intonations that align with the simulated medical conditions and dialog. The virtual patient can also have autonomic behaviors, like eye gaze, twitches, scratching and idle movements. It is also possible that the system can get interrupted before all of the actions are generated and displayed by a user with questions, clarifications, or an event like spilling water. The system needs to be reactive and constantly monitoring and planning based on inputs, which is not an easy task to perform

3. Examples of using LLMs in the design and implementation process

3.1. Creating a character Persona

Creating Personas for the virtual patients using the LLM require describing the attributes for the patient. An example of a LLM generated Persona Profile Template for a VSP

includes: Demographics; Medical History; Mental Health Information; Behavioral Characteristics; Social and Cultural Context; Interaction History; and Additional Interaction Styles. This template can be filled out for each virtual patient to provide a detailed profile to help in generating realistic and contextually appropriate interactions using an LLM. The data is cached so that it can be used when priming the input prompts to the system to be used later when performing the conversations. Additionally the system can be told what values would change over time and how they might change between sessions so the generated questions and answers would be more relevant.

3.2. Using the LLM to generate a range of answers to questions in a session

This is an example of a multi layered prompt and responses to the LLM. The prompt would be: "I am going to do an example here where you will act as a patient with PTSD from a recent war. I want you to answer the question I give you, but also generate 5 responses and rate them." **The question is: Why is the sky blue?**

Response 1: "The sky is blue because of the way the Earth's atmosphere scatters sunlight. Shorter blue wavelengths are scattered more than other colors, making the sky appear blue." Rating: 5/5 Rationale: This response is scientifically accurate, concise, and straightforward, providing a clear explanation of Rayleigh scattering. **Response 2:** "The blue color of the sky is due to the scattering of sunlight by the gases and particles in the atmosphere. Blue light is scattered in all directions, which is why we see a blue sky most of the time." Rating: 4/5 Rationale: This response is also accurate but slightly less precise in explaining the scattering process compared to the first response. **Response 3:** "I used to look at the blue sky and feel a sense of peace, but now it just reminds me of the open spaces in the desert where everything went wrong." Rating: 5/5 Rationale: This response integrates the PTSD context, showing the patient's emotional state and how their perception of the sky has changed due to their experiences. **Response 4:** "Honestly, I don't really think about why the sky is blue anymore. It's just there, you know? Like a constant reminder of everything that's happened." Rating: 4/5 Rationale: This response is relevant to the PTSD context, reflecting the patient's disinterest in mundane details due to their trauma, though it doesn't address the scientific aspect. **Response 5:** "The sky is blue because of some science stuff, but right now, it feels like the sky could be any color, and it wouldn't change how I feel." Rating: 3/5 Rationale: This response touches on the scientific reason but is vague. It effectively conveys the patient's emotional disconnect but could be clearer on the scientific explanation. The responses range from scientifically accurate to emotionally expressive, reflecting the complexity of answering questions from the perspective of a virtual standardized patient.

3.3. Simulated Conversation: Therapist and a generated character

The LLM, programmed with John's persona profile (an example person), generates realistic responses based on the trainee's questions and prompts. There are several ways to use the LLM in generating conversations. The LLM can generate both the therapists' questions and the patient's responses. The transcripts could be refined and checked by experts. The LLM can also be handed a question to ask the virtual patient that will respond based on its Persona and other provided data. Current methods that would not use an LLM and require experts to create a complete script that actors would play out for that specific session is how current OSCE [23] are performed. But it could also be possible to take the dialog students and the actors say in a session and input it into the LLM, as training data if there was enough, or just to evaluate it. Here is an example summary of the John character generated by the LLM: John is a VSP who has been experiencing severe PTSD symptoms, particularly frequent nightmares, stemming from a traumatic ambush in Afghanistan. He often wakes up in a state of extreme distress, drenched in sweat, and finds it challenging to calm down, even when employing coping strategies like deep breathing. His emotional state is characterized by sadness, anxiety, and frustration, which are reflected in his facial expressions, body language, and tone of voice. John struggles to maintain eye contact and displays physical signs of tension, indicating the impact of his trauma on his daily life. .

3.4. Example of Generating Behavior based on the output

Here is an example dialog for John and a therapist generated by the LLM based on the Persona, it is also augmented with additional output to generate realistic behavior, voice, intonation for the various virtual patient modules in a 3D character. The output can be annotated into a JSON string processed by the system. Here the LLM acts as the therapist, generates a question and then acts as the virtual patient to generate a response.

In a training environment all the dialog could be tracked and scored.

Annotated Conversation: Therapist: Good morning, John. How have you been feeling since our last session? **Client John:** Text: Morning. It's been a tough week.

I've been having more nightmares than usual (Emotions: Sadness, Fatigue; Facial Gestures: Downcast eyes, slight frown; Behavior Animations: Slow head nod, shoulders slightly hunched; Intonations: Low, tired tone; Gaze and Rapport: Avoids direct eye contact, looks down). **Therapist:** I'm sorry to hear that. Can you tell me more about the nightmares? Are they related to your experiences in Afghanistan? **Client John:** Text: Yeah, they're mostly about the ambush we went through. It feels so real every time, and I wake up drenched in sweat (Emotions: Anxiety, Distress; Facial Gestures: Furrowed brow, widened eyes; Behavior Animations: Quick, shallow breathing, wiping sweat from forehead; Intonations: Strained, shaky voice; and Gaze and Rapport: Glances up briefly, then looks away). **Therapist:** That sounds really difficult. Have you been able to use any of the coping strategies we discussed, like deep breathing or mindfulness exercises, when you wake up from these nightmares? **Client John:** Text: I've tried the deep breathing, but it's hard to calm down when I wake up so shaken. It takes a while for me to feel normal again (Emotions: Frustration, Hopelessness; Facial Gestures: Tightened jaw, pursed lips; Behavior Animations: Slight head shake, hand clenched; Intonations: Slightly raised voice, sighs at the end; and Gaze and Rapport: Looks directly at the therapist briefly, then looks away again). These annotations aim to capture the emotional and physical state of the virtual patient enhancing the realism and depth of the interaction for training or therapeutic purposes.

4. Discussion

By integrating an LLM with a well-defined virtual patient persona, training sessions can provide immersive, realistic, and interactive learning experiences for healthcare professionals, enhancing their skills and preparedness for real-world clinical situations. There are not enough actors and qualified clinicians to create realistic dialogs for effective clinical therapist training. Addressing these challenges is essential for the development of virtual human AI agents that can effectively serve as training tools in medical practice. By leveraging advanced AI technologies such as Large Language Models (LLMs), there is potential to overcome these obstacles and create more realistic, dynamic, and educationally valuable virtual patients. The LLM AI Agents can be personalized with structured dialogs and sessions for the virtual patient and can be generated relatively quickly with only high level English like descriptions to the system. The underlying 3D virtual patient characters can use a wide range of modules that can gain increasingly complex functionality as the LLM and agent architectures develop. Once more trust is developed with how the LLM's can generate the Persona's, dialog and conversations, and take into account biases, this technology will be more fully adopted.

References

- [8] Kenny P, Parsons T. Conversational Agents and Natural Language Interaction: Techniques and Effective Practices. 2011.
- [9] Kenny P, Parsons T. Embodied Conversational Virtual Human Patients. 2011. doi: <https://doi.org/10.4018/978-160960-617-6.ch011>.
- [10] Kenny P, Parsons T, Garrity P. Virtual Patients for Virtual Sick Call Medical Training. 2010.
- [11] Kenny PG, Parsons TD, Reger G, Pataki C, Pato M, Sugar J, et al. Virtual Patients for Future Leaders. In: Interservice/Industry Training, Simulation and Education Conference (I/ITSEC); 2008.

- [12] Kenny P, Parsons T, Rizzo A. A Comparative Analysis between Experts and Novices Interacting with a Virtual Patient with PTSD. *Stud Health Technol Inform.* 2009;144:122-4. doi: <https://doi.org/10.3233/978-1-60750017-9-122>.
- [13] Kenny P, Parsons T, Gratch J, Rizzo A. Evaluation of Justina: A Virtual Patient with PTSD. *Lect Notes Comput Sci.* 2008;394-408. doi: https://doi.org/10.1007/978-3-540-85483-8_40.
- [14] Kenny P, Parsons T, Gratch J, Rizzo A. Virtual humans for assisted health care. In: Proceedings of the 1st international conference on Pervasive Technologies related to assistive environments. 2008. doi: <https://doi.org/10.1145/1389586.1389594>.
- [15] Leuski A, Traum D. A Statistical Approach for Text Processing in Virtual Humans. In: Proceedings of the 26th Army Science Conference. 2008.
- [16] Kenny P, Hartholt A, Gratch J, Swartout W, Traum D, Marsella S, et al. Building interactive virtual humans for training environments. *Proc IEEE.* 2007.
- [17] Swartout W, Gratch J, Hill R, Hovy E, Marsella S, Rickel J, et al. Toward virtual humans. *AI Mag.* 2006;27(1).
- [18] Jones R, Laird J, Nielsen P, Coulter K, Kenny P, Koss F. Automated Intelligent Pilots for Combat Flight Simulation. *AI Mag.* 1999;20. doi: <https://doi.org/10.1609/aimag.v20i1.1438>.
- [19] Laird J, VanLent M. Human-Level AI's Killer Application: Interactive Computer Games. *AI Mag.* 2001;22(2):15. doi: <https://doi.org/10.1609/aimag.v22i2.1558>.
- [20] Darcy A, Daniels J, Salinger D, Wicks P, Robinson A. Digital Therapeutic Alliance: Is Bond with a Relational Conversational Agent Possible? Symposium presented at the annual meeting of the Association for Behavioral and Cognitive Therapies; 2021; New Orleans, LA.
- [21] Xie Q, et al. Human Simulacra: Benchmarking the Personification of Large Language Models. arXiv preprint arXiv:2402.18180v5 [cs.CY]. 2024.
- [22] Li J, et al. Agent hospital: A simulacrum of hospital with evolvable medical agents. arXiv preprint arXiv:2405.02957. 2024.
- [23] Newell A. Unified Theories of Cognition. Cambridge, MA: Harvard University Press; 1990.
- [24] Jones CR, Bergen BK. People cannot distinguish GPT-4 from a human in a Turing test. 2024. doi: <https://doi.org/10.48550/arXiv.2405.08007>.
- [25] Darcy A, Daniels J, Salinger D, Wicks P, Robinson A. Evidence of Human level Bond Established with a Digital Conversational Agent: An Observational Study. *JMIR Form Res.* doi: <https://doi.org/10.2196/27868>.
- [26] Gao T, Fisch A, Chen D. Making Pre-trained Language Models Better Few-shot Learners. *CoRR.* 2020;abs/2012.15723. doi: <https://doi.org/10.48550/arXiv.2012.15723>.
- [27] Wu T, Terry M, Cai CJ. AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems; 2022.
- [28] Park JS, O'Brien J, Cai CJ, Morris MR, Liang P, Bernstein MS. Generative agents: Interactive simulacra of human behavior. In: Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology; 2023. p. 1-22.
- [29] Qian C, Cong X, Yang C, Chen W, Su Y, Xu J, et al. Communicative agents for software development. arXiv preprint arXiv:2307.07924. 2023.
- [30] Parsons T, Kenny P, Ntuen C, Pataki C, Pato M, Rizzo A, et al. Objective structured clinical interview training using a virtual human patient. *Stud Health Technol Inform.* 2008;132:357-62.

Novel Memory Assessment Methods in Minimally Invasive Epilepsy Surgery

Daniel DRANE^{a,b,c,1}, Noah OKADA^{a,d}, Yousef RAJEH^c, Ellie PAEK^c, Keith ADAMS^f,
Emily WALL^e

^a*Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA*

^b*Department of Pediatrics, Emory University School of Medicine, Atlanta, GA, USA*

^c*Department of Neurology, University of Washington School of Medicine, Seattle, WA, USA*

^d*California Institute of Technology, Pasadena, CA, USA*

^e*Department of Computer Science, Emory University, Atlanta, GA, USA*

^f*SpinVFX, Atlanta, GA, USA*

ORCiD ID: Daniel Drane <https://orcid.org/0000-0001-8747-672X>

Abstract. The rise of minimally invasive neurosurgical procedures (e.g., stereotactic laser ablation) coupled with technological advances has revealed gaping holes in cognitive theory and our ability to thoroughly assess such constructs. Being able to create focal surgical destruction zones has revealed a mismatch between the existing structure-function theory of the brain and post-surgical results. For example, extant research literature often focuses on the involvement of the medial temporal lobes in memory or the fusiform gyrus in semantic memory/language, yet these highly precise lesional studies are showing theory to often be incomplete or incorrect. In the setting of SLA in epilepsy surgery, some of our worst post-surgical memory outcomes occur when extra-medial TL regions are destroyed rather than medial TL structures. This is likely because cognitive theory has been based on indirect, correlative measures of brain function (e.g., fMRI) or large lesions in the brain resulting from disease or surgery. Additionally, most clinical measures of cognitive and emotional functioning are kept simplistic in nature to allow for the most straightforward interpretation. For example, memory testing is usually done in a sensory domain-specific manner (e.g., visual vs. auditory) rather than allowing for the integration of memory features (e.g., visual, auditory, semantic, autobiographical, historical being integrated and assessed holistically). We highlight emerging weaknesses in theory as well as shortcomings in cognitive assessment and present data to demonstrate how novel tests can be developed using videography, gamification, internet delivery to allow for longer windows of delayed recall, and updated theory to better assess neural network interactions.

Keywords. Memory assessment, telemedicine/virtual assessment, surgical outcome, unity game engine

5. Introduction

The perfect storm is brewing between the advent of cutting-edge advances in both technology and neuroscientific theory and promises to unlock new opportunities to study the brain and to advance our options for diagnostic inquiry and treatment alike. Nowhere is this clearer than in the neurosurgical setting, where minimally invasive techniques have opened new windows into the structure-function relationships of key cognitive and socioemotional functions while challenging many long-held conceptual and neuroanatomical theories of such constructs [1,2,3] We will explore some examples of these occurrences from our own work in the epilepsy surgery setting, demonstrating that the field of epilepsy surgery has been severely under-evaluating the construct of memory since its inception 70 years ago, and putting forth our initial efforts to solve some of these problems through novel assessment paradigms and the application of virtual assessment and machine learning algorithms for administration and scoring.

¹ Corresponding Author: Daniel Drane, ddrane@emory.edu

6. Minimally Invasive Surgical Procedures Reveal that Cognitive Theories are Often Wrong

The advent of minimally invasive surgical procedures, such as stereotactic laser amygdalohippocampotomy (SLAH) for carrying out a selected temporal surgery restricted to the amygdalar-hippocampal complex (AHC), has demonstrated that we can do a better job of preserving cognitive function while still conducting surgery for the control of seizures [2,4]. Rather than removing a large portion of the temporal lobe, as is the case in a standard anterior temporal lobectomy (ATL), the neurosurgeon can now restrict surgery using a laser device to just a portion of the AHC. We have demonstrated over the past decade that we can virtually spare all language functions and even lessen episodic memory deficits by minimizing the anatomical scope of surgery [5,6,7,8].

Taking the example of sparing language is a perfect example of the mismatch between scientific theory and the novel revelations resulting from the interface of technology and neuroscience, multiple research articles have suggested a role for the hippocampus in naming ability, and many researchers have suggested that sparing the hippocampus could lead to better language outcomes [9] There have been at least four lines of research pointing to this hypothesis: (a) associations between presurgical naming scores and hippocampal volumes [10], (b) hippocampal activation observed during fMRI tasks involving object and face naming [11], (c) worse naming scores observed in TLE patients with MTS than in those without,[12,13] and (d) observations that naming deficits always occurred when the hippocampus was included in ATLs despite patients undergoing cortical stimulation mapping [9] All of this supporting research is primarily indirect and correlative rather than a “causal” model. In contrast, our work has shown that actually ablating the language-dominant AHC virtually never leads to any form of language dysfunction [5,6] and that deactivating the hippocampus through a selective posterior cerebral artery Wada procedure likewise leads to no language disruption [14].

Similarly, with episodic memory, the conventional clinical wisdom has been to spare the hippocampus whenever possible with the goal of preserving memory functioning, with the prevailing thought being that these structures are highly critical to forming new memories. Nevertheless, surgeries sparing the hippocampus have not seemed to preserve memory [15], and our own work with laser ablation has generally shown a favorable memory outcome when the hippocampus is destroyed. While episodic verbal memory can be affected by a language-dominant SLAH, fewer patients tend to decline and the extent of decline appears less in magnitude [7,8,16]. In contrast, some of our worst outcomes on conventional verbal memory measures have occurred following focal surgical lesions created in the fusiform gyrus or the temporal pole [4] Overall, these findings and much of the animal data on memory suggest that memory is likely dependent on a wide range of structures that exceed the medial TL region [17,18,19].

7. Technology and Videography Can Improve Our Ability to Assess Cognition

While innovative surgical technologies have led to many challenges of long-held structure-function models of cognition, they have also highlighted where there are deficiencies in our current test paradigms. Despite sparing of function on conventional tests following epilepsy surgery, some patients and their families continue to report new onset cognitive problems. We have been carefully interviewing them as part of a long-term outcome research project and have found that there are several areas where our standard tasks are falling woefully short of a complete evaluation. These include an absence of measures designed to explore long-term consolidation of information (i.e., we typically only test recall following a half-hour delay), the emotional contribution to memory, and the contextual information of when and where the information was learned, among other constructs. We would like to spend the remainder of this brief review, highlighting how technology may overcome these glaring absences in our assessment armamentarium.

Learning new information prospectively has rarely ever been investigated systematically in any age group, though the consequences of deficits in prospective learning are likely to be much more severe than losing access to stored knowledge and information. Our group has demonstrated that language-dominant open resections lead

to a greatly diminished ability to recall names of famous persons and that right TL open resections diminish one's ability to recognize such individuals or to recognize familiar landmarks and buildings [20, 21] We have speculated that patients with these deficits, both left and right TL resection patients, will have difficulty putting information with faces or locations, albeit for different mechanistic reasons. Such deficits have been reported in the research literature by patients and are particularly troubling [22,23] as one would assume that the effects of these deficits would compound with each passing year (i.e., more and more information comes along that such individuals cannot learn), and this leads to greater mental health issues and life limitations.

Furthermore, assessing the multimodal integration of information, such as combining sensory and motor input with semantic/linguistic concepts and memories, is completely absent in the clinical assessment of patients. Aspects of neuroscience research have focused on determining how multisensory input is bound in the brain, which has been driven in recent decades by discoveries of multimodal cells in single-unit recordings [24,25] and evidence that primary sensory areas are interconnected more than once realized [26]. Current work has suggested that multisensory integration is critical for many complex behavioral routines to be successfully performed [27], and it should be intuitively obvious that we are always making use of multisensory input to navigate the world, and this processing has to be tightly interwoven with our memories and ability to recognize patterns in the world. Understanding the neural substrates of such integration remains in the early stages of development, and clinical assessment of these fundamental yet highly complex interactions must be integrated into clinical assessments.

To advance towards these objectives, we have created the Emory Multimodal Learning Task (EMLT) and Emory Pediatric Multimodal Learning Task (EPMLT), a set of novel tools specifically designed for concurrent assessment of various memory processes. This task aims to evaluate both traditional memory functions, such as face/object processing and semantic/episodic learning and memory, as well as novel, intricate learning processes involving the integration of sensory information with linguistic, semantic, and episodic elements. This comprehensive assessment takes advantage of modern technological tools to present stimuli (video clips, images) in a multiplatform (desktop, tablet) gamified environment, which allows for virtual testing to allow for longer periods of follow-up (e.g., one week, one month, etc. rather than just 30-minute recall). The video clips feature professional actors each in a unique setting and scene, and these were produced under the supervision of SpinVFX, inc.

Using the Unity Game Engine, we developed a stand-alone application to administer this multimodal memory assessment. Each of our tasks consists of two blocks: a learning block and a delayed recall block. During the learning block, participants are immersed in a gamified town environment where they can view a series of 11 videos that display actors conducting daily living activities. In each video, the actors introduce novel information (faces, names, locations, objects), which can be used to assess memory across multiple modalities. The delayed recall block consists of an interactive user interface that allows participants to respond to several questions designed to evaluate participants' capacity to integrate this information effectively and accurately identify contextual features embedded within the videos. To ensure robustness in data collection, the framework is designed to be compatible with peripheral monitoring devices such as eye-tracking devices and EEG (See **Figure 1**).



Figure 1. A new multimodal tool that is under development for the assessment of multiple domains of cognition and their integration, along with simultaneous recorded eye position and pupil diameter data.

We have successfully developed and validated this gamified framework for use in patients with epilepsy. Preliminary results indicate that this assessment can effectively measure participants' ability to successfully recall and integrate visual perceptual, auditory, semantic, and episodic information. Additionally, we are developing normative data for healthy subjects and patients for delayed recall intervals of 1 week and 1 month and have found delayed recall of these more meaningful stimuli to be more recallable than traditional measures (which tend to be rote in nature). Preliminary data suggests that both memory deficits and strengths were being routinely missed by classic standard measures in our epilepsy surgery patients and that these novel measures will allow for more thorough, ecologically valid assessments of surgical outcome, the potential for more effective rehabilitative intervention, and a heightened knowledge of structure-function brain relationships and new models of memory processes.

References

- [1] Sperling MR, Barshinger B, Miocinovic S, et al. Stereotactic laser ablation for mesial temporal lobe epilepsy: a prospective, multicenter, single-arm study. *Epilepsia*. 2020;61:1183–1189. doi: <https://doi.org/10.1111/epi.16541>
- [2] Willie JT, Laxpati NG, Drane DL, et al. Real-time magnetic resonance-guided stereotactic laser amygdalohippocampotomy (SLAH) for mesial temporal lobe epilepsy. *Neurosurgery*. 2014;74:569–585. doi: <https://doi.org/10.1227/NEU.0000000000000262>
- [3] Prada F, Kalani MYS, Yagmurlu K, et al. Applications of focused ultrasound in cerebrovascular disease and brain tumors. *Neurotherapeutics*. 2019;16:67–87. doi: <https://doi.org/10.1007/s13311-018-00689-4>
- [4] Drane DL. MRI-guided stereotactic laser ablation for epilepsy surgery: promising preliminary results for cognitive outcome. *Epilepsy Res.* 2018;142:170–175. doi: <https://doi.org/10.1016/j.eplepsyres.2018.03.009>
- [5] Drane DL, Loring DW, Voets NL, et al. Better object recognition and naming outcome with MRI-guided stereotactic laser amygdalohippocampotomy for temporal lobe epilepsy. *Epilepsia*. 2015;56:101–113. doi: <https://doi.org/10.1111/epi.12854>
- [6] Hageboutros K, Drane DL, Loring DW, et al. Comparison of minimally invasive to standard temporal lobectomy approaches to epilepsy surgery: seizure relief and language outcomes. *Epilepsy Behav.* (in press).
- [7] Gross RE, Stern MA, Willie JT, et al. One-year outcomes following stereotactic laser amygdalohippocampotomy for the treatment of mesial temporal lobe epilepsy: 5-year experience in 58 patients. *Ann Neurol*. 2018;83:575–587. doi: <https://doi.org/10.1002/ana.25187>
- [8] Drane DL, Pedersen NP, Qiu D, et al. Superior verbal memory outcome after stereotactic laser amygdalohippocampotomy. *Front Neurol*. 2021;12:779495. doi: <https://doi.org/10.3389/fneur.2021.779495>
- [9] Hamberger MJ, Seidel WT, McKhann GM, Goodman RR. Hippocampal removal affects visual but not auditory naming. *Neurology*. 2010;74:1488–1493. doi: <https://doi.org/10.1212/WNL.0b013e3181dd4e46>
- [10] Seidenberg M, Geary E, Hermann B. Investigating temporal lobe contribution to confrontational naming using MRI quantitative volumetrics. *J Int Neuropsychol Soc*. 2005;11:358–366. doi: <https://doi.org/10.1017/S1355617705050401>
- [11] Bonelli SB, Powell R, Yogarajah M, et al. Hippocampal activation correlates with visual confrontation naming: fMRI findings in controls and patients with temporal lobe epilepsy. *Epilepsy Res*. 2011;95:246–254. doi: <https://doi.org/10.1016/j.eplepsyres.2011.03.006>
- [12] Baxendale SA, Thompson PJ, Harkness WF, et al. The relationship between quantitative MRI and neuropsychological functioning in temporal lobe epilepsy. *Epilepsia*. 1998;39:158–166. doi: <https://doi.org/10.1111/j.1528-1157.1998.tb01364.x>

- [13] Davies KG, Bell BD, Bush AJ, et al. Naming decline after left anterior temporal lobectomy correlates with pathological status of resected hippocampus. *Epilepsia*. 1998;39:407–419. doi: <https://doi.org/10.1111/j.1528-1157.1998.tb01404.x>
- [14] Drane DL, Pedersen NP, Loring DW, et al. Selective posterior cerebral artery Wada better predicts good memory and naming outcomes following selective stereotactic thermal ablation for medial temporal lobe epilepsy than internal carotid artery Wada. *medRxiv*. 2024.
- [15] Wagner K, Willmes K, Urbach H, et al. Memory outcome after hippocampus sparing resections in the temporal lobe. *J Neurol Neurosurg Psychiatry*. 2013;84:630–636. doi: <https://doi.org/10.1136/jnnp-2012-304644>
- [16] Donos C, Breier J, Tandon N. Laser ablation for mesial temporal lobe epilepsy: surgical and cognitive outcomes with and without mesial temporal sclerosis. *Epilepsia*. 2018;59:1421–1432. doi: <https://doi.org/10.1111/epi.14439>
- [17] Saling MM. Verbal memory in mesial temporal lobe epilepsy: beyond material specificity. *Brain*. 2009;132:570–582. doi: <https://doi.org/10.1093/brain/awn339>
- [18] Basile BM, Hampton RR. Nonnavigational spatial memory performance is unaffected by hippocampal damage in monkeys. *Hippocampus*. 2018;28:1–9. doi: <https://doi.org/10.1002/hipo.22966>
- [19] Zola-Morgan S, Squire LR, Amaral DG, Suzuki WA. Lesions of the perirhinal and parahippocampal cortex that spare the amygdala and hippocampal formation produce severe memory impairment. *J Neurosci*. 1989;9:4355–4370. doi: <https://doi.org/10.1523/JNEUROSCI.09-12-04355.1989>

Development and Application of a Three-Screen Enveloping Projection Cyberspace that Is Easy to Disassemble and Assemble, with the Aim of Contributing to People's Wellbeing

Mieko OHSUGA^{a,1}

^a*Faculty of Robotics and Design, Osaka Institute of Technology*
ORCID ID: Mieko Ohsuga <https://orcid.org/0009-0007-0006-7853>

Abstract. We have developed a three-sided housing type projection system that is easy to disassemble and reassemble. As for the application of this system, we proposed a cyber system that allows hospitalized patients to walk through the outside world and feel nature, and a cyber system in which people at a distance meet via telepresence robots. We also developed a prototype system in which two of these systems can be connected via a network to create a virtual third place where people at a distance can enjoy conversation and share experiences. These trials were favorably received at exhibitions and in preliminary experiments. Solving various technical issues and conducting evaluation experiments are future tasks.

Keywords. Cyberspace, projection system, human interaction, third space

1. Introduction

Large-scale fixed multi-surface projection systems have been introduced not only in research facilities but also in many exhibition and amusement facilities, providing multi-user interactive contents. As a departure from these facilities, we have constructed a system that is relatively inexpensive, easy to disassemble and reassemble, and portable. The reason for this is that we would like to have people experience the system by taking it to places where it could be introduced, and repeatedly improve the contents based on feedback from the people there.

We have developed a three-sided enclosure-type cyberspace that is easy to disassemble and reassemble in order to support human interaction, improve wellbeing, and alleviate stress for a variety of people. In this presentation, the features and advantages of the developed system, some case studies and future issues will be discussed. The system is an encircling projection system capable of life-size projection of just over two meters on each side. The space consists of three screens and a floor, with three projectors and a distortion

¹ Corresponding Author: Mieko Ohsuga, mieko.ohsuga@oit.ac.jp

correction device. The walls are foldable and the floor is divided into four sections, each of which can be carried by one person.

Compared to a head-mounted display (HMD), the system gives less burden and has a lower risk of VR sickness, making it suitable for a wide range of users and applications, including the older people, children, and people with physical disabilities. Two sets of the same configuration were prepared. By connecting the two locations, it is possible to give people at a distance the experience of being in a connected space or in the same space.

2. System Configuration

The three-panel projection system consists of a foldable freestanding panel, three projectors, a video processor, a floor that can be disassembled into a portable size, and a disassembled frame to secure their positioning (**Figure 1**). The video processor provides edge blending and geometric correction in hardware to achieve seamless multi-projection. **Figure 2** shows the projection onto a single surface. An ultra-short focus projector placed on the floor and projected onto a screen through a hole in the raised floor. The panel is covered with an elastic fabric with good reflectivity.

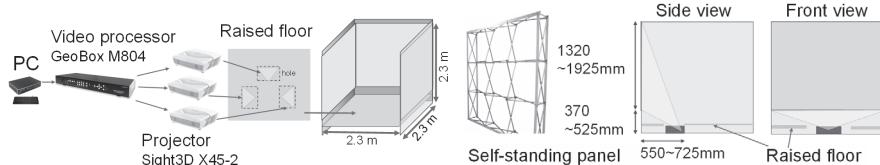


Figure 1. System configuration of the developed system

Figure 2. Projection on the wall

Two sets of such systems were produced and networked (**Figure 3**) to realize various virtual arrangements of two cyberspaces, as shown in **Figure 4**.

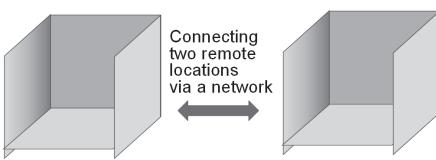


Figure 3. Two sets connect

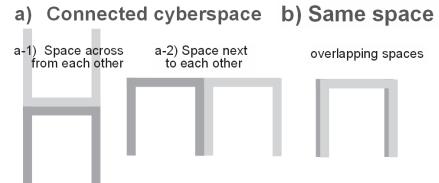


Figure 4. Virtual arrangement of cyberspaces

3. Application Examples

3.1. Walking System in Virtual Space as an Avatar

More than 20 years ago we attempted to implement a VR system for palliative care of cancer patients [1, 2]. The concept was well received by both medical staff and patients. We evaluated the effectiveness of the system on a small number of patients and found that it provided relief from pain and fatigue. However, the system was not put into practical use due to the space required, the time needed for the medical staff to operate the system, and the high cost of implementation. Several years ago, we took up this challenge again and developed a system that allows patients in sterile rooms to feel the outside world through projection mapping [3, 4]. After preliminary studies and the initiation of a case study with patients, the study was interrupted due to COVID-19.

We started the third attempt with this simplified projection system, taking advantage of the flexible configuration and portability of the system. This is because the system is suited for having inpatients experience the prototype, exploring expectations of acceptability and effectiveness, and extracting points for improvement. The cyberspace was created using Unity assets with the first content being a walk through a forest. The first step was to introduce a third-person avatar, although the viewpoint could be switched, and Hand gestures, a keypad, and a game controller were prepared for the avatar's movement within the space for the user to choose from.

In addition, we have added 3D acoustics using Sound-XR (Yamaha corp.). Not only does 3D acoustics give a sense of spatial spaciousness, it also facilitates to build a mental map of the space and can be used for navigation. An example of a spatial map is shown on the right side of **Figure 5**. Visitors can stroll through the virtual forest based on the direction and distance from where they hear invisible birds or animals, and enjoy the sounds of murmuring streams and wind.



Figure 5. Virtual forest walk

3.2. Cyberspace Connecting People at a Distance

The next example connects people at a distance. However, only one set of systems is used. Neither user comes to the real location of this system. They meet in cyberspace via two telepresence robots (iPresence, Inc.), each equipped with an iPad and capable of remotely controlling motors and operating two axes of rotation via a network. While enjoying the scenery and contents of cyberspace, the user can see the other person's face, creating a natural conversation. The cyberspace provided can be created not only in a three-dimensional space created by computer graphics, but also by attaching a 360-degree moving image to the inside of a spherical dome installed in the Unity Space. It is envisioned to be used for conversing in a relaxed atmosphere at home or in a familiar café, or in a place you have never been to before, while looking out the window at the scenery and events.



Figure 6. Cyberspace connecting people at a distance using 2 telepresence robots

3.3. Cyber Third Place

The third example is the provision of a third place that is neither work, school, nor home. The need for a third place as a real space that functions to reduce loneliness and improve well-being by allowing people to interact with each other has been pointed out. Based on this, we considered that a cyber third place would also be useful for people who cannot meet in person, allowing them to spend important time together and enjoy conversation in a place that is not physically accessible.

It takes advantage of the strengths of the virtual world, such as the ability to choose one's avatar depending on the person one meets, and the ability to easily change the scenery and interior design according to the mood of the users at the time. **Figure 7** shows an example of a conversation with a friend in a wooden log house, the arrangement of a-

1) in **Figure 4**. It uses a first-person perspective in which the other person's avatar is visible but his or her own avatar is not.

Figure 8 is another example using the arrangement in **Figure 4 b)**, where virtual spaces such as a café and a living room were also provided for users to choose from. The objects that make up the virtual space were created using Unity and free assets, with the addition of functions to change the music and lighting (**Figure 9**). The avatars were created using VRoid for appearance and mixamo for motion. Two 3-sided projection systems are connected to the cloud using PUN2. A prototype was created and preliminary experiments were conducted to evaluate system functionality and avatar operability, and improvements were made based on feedback from participants.



Figure 7. Cyber third space: a wooden log house



Figure 8. Cyber third space: a café and a living room



Figure 9. Examples of changing the lighting

4. Preliminary Evaluation Experiments

4.1. Virtual Forest Walk and Warp to Natural Space (ethics approval number 2023-81)

In addition to the forest walk shown in 3.1, we created content that warps participants into a natural space with a 360-degree video projection, and held an experience session with patients and medical staff to collect their comments. In content using 360-degree video, the viewpoint can only be changed from the left, right, up, and down; walk-throughs were not possible. The scenes included a seaside, a forest on a plateau, a park with cherry blossoms, and a sunflower field. The hands-on session was held in the family briefing room of the orthopedic surgery ward at the Osaka International Cancer Center. Since two sides of the room were empty white walls, three projectors on the floor were used to project images on these two walls and one freestanding wall of the system. The floor was not raised to allow easy access for wheelchair users and people on intravenous drips. The opinions of patients, physicians, nurses and physiotherapists were collected during the trial session. The trial was held for two days. 6 patients, 5 doctors, 5 nurses, and 11 staff members participated in the trial through posters and calls from the doctors. The ages of the participants ranged from 19 to 60, 11 males and 11 females.

As a result, except for one negative respondent who said that the image quality was poor and the immersive feeling was not enough, the other told that it was good to experience the outside world and refreshing when they could not go outside, that they would like to experience it again, and that they would be happy to have it in the hospital ward. More than half of the participants liked both CG and real images, and several patients preferred real, especially the ocean. The sound was highly evaluated, perhaps because 3D sound experiences were rare, even though similar video experiences were available. There were also requests for unrealistic worlds such as outer space and the deep sea. Participants requested a wider vertical angle of view, walk-throughs in real image contents, and a sense of wind.

4.2. Cyber Third Place (ethics approval number 2023-81)

Two pairs of healthy university students (four people) who gave informed consent experienced the prototypes before and after improvement, and compared the functions and operability of the prototypes and the occurrence of VIMS (visually induced motion sickness), by subjective evaluation, and collected opinions about the future prospects of utilizing the proposed system.

As a result, most of the participants commented on the functional improvements, especially the avatar and spatial setting improvements such as the addition of avatar gestures, adjustment of the speed of viewpoint movement, and changes in the display of lighting. The results were divided into two groups, with two people reporting reduced VIMS and two reporting increased VIMS, with the comment that the increased VIMS was caused by the intense viewpoint movement and motion during the interaction using the avatars.

5. Considerations and Future Work

A three-plane projection system characterized by ease of disassembly and assembly, and several applications were introduced. There are many research issues to be addressed,

including the realization of communication between two points without delay, measures against VIMS, the introduction of tactile and force sensations and the measurement of the human state mainly physiological measurements with less burden, and real-time adaptation of content using these measurements. In parallel with technological development, we believe it is important to conduct demonstration research combining technologies that can be stably realized at that point in time to confirm the acceptability and needs of the field. In actual operation, the system is envisioned to be installed in third places that have already been realized as actual spaces, such as family visiting rooms in hospitals, facilities for the elderly, communal facilities in local communities, fitness clubs, and karaoke boxes. Furthermore, for those who cannot visit the installation location, we would like to allow them to choose the style of participation (**Figure 11**).

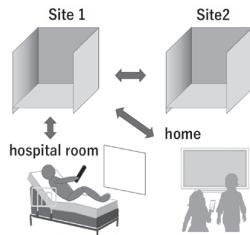


Figure 11. Various participation styles

Acknowledgment

A part of this work (chapter 3.2 and 4.2) was supported by JSPS KAKENHI Grant Number JP 18K10728. In implementing the system, I received a lot of support from the faculty and staff, members of the Cyberspace Working Group, Solidlay co. ltd., Yamaha

Corporation, forestdigital, Inc. I thank to Dr. Yoshiyuki Kamakura who provided gesture identification technology. I am grateful to the students of the Wellness Lab, especially Mr. Masashi Koizumi, Mr. Keito Shoda, and Dr. Hidenori Boutani, for their efforts in content development.

References

- [1] Ohsuga M, Tatsuno Y, Shimono F, Hirasawa K, Oyama H, Okamura H. Development of a bedside wellness system. *CyberPsychology & Behavior*. 1998;1(2):105–112. doi: <https://doi.org/10.1089/cpb.1998.1.105>.
- [2] Oyama H, Ohsuga M, Tatsuno Y, Katsumata N. Evaluation of the psycho-oncological effectiveness of the bedside wellness system. *CyberPsychology & Behavior*. 1999;2(1):81–84. doi: <https://doi.org/10.1089/cpb.1999.2.81>
- [3] Ohsuga M, Tada Y, Ishikawa J. Interactive environment for hematopoietic stem-cell transplant patients. In: 2017 International Conference on Virtual Rehabilitation (ICVR). IEEE; 2017. p. 1–2. doi: <https://doi.org/10.1109/ICVR.2017.8007469>
- [4] Joko S, Ohsuga M, Tada Y, Ishikawa J. Evaluation of a mental care system for patients recuperating in a sterile room after hematopoietic cell transplantation. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2019. p. 1314–1317. doi: <https://doi.org/10.1109/EMBC.2019.8857157>

SECTION VI

WORK IN PROGRESS

It is important to emphasize the importance of developing technological strategies (such as artificial intelligence or augmented reality) that can provide either new enhanced experiences or technological systems also nurtured by artificial intelligence techniques developed by humans.

These new mixed ICT tools might evolve into experts in “helping others,” with the objective of making our net-shared experience increasingly more competitive, creative, and capable in the task of helping others. Of course, this has significant ethical implications, which will also need to be explored at greater depth.

*Botella, Riva, Gaggioli,
Wiederhold, Alcaniz,
and Banos, 2012*

New Considerations for Memory Assessment

Anne M. CLEARY^{a,1}

^a*Department of Psychology, Colorado State University, 1876 Campus Delivery, Fort Collins, CO 80523, USA*

ORCiD ID: Anne Cleary <https://orcid.org/0000-0003-1856-9070>

Abstract. Standard memory assessments typically do not measure all, or even most, facets of human memory. Although some of the more nuanced aspects of human memory, such as the ability to detect familiarity with something despite failing to recall specifics, might seem inconsequential to overall memory function, a growing body of research suggests that these more nuanced aspects of memory may be among the most important components to overall memory function. The present work covers some of the more nuanced aspects of memory function that are largely overlooked in standard approaches to assessing memory function, including the unique role that one's surroundings may play in autobiographical and event memory, the important role that general familiarity-detection ability might play in directing the mind toward memory search effort, why memory oddities like *déjà vu* and tip-of-the-tongue experiences might be useful and important aspects of memory, and why the multimodal nature of memory is important to consider.

Keywords. Memory assessment, familiarity, familiarity-detection, scenes.

1. Introduction

Human memory is multi-faceted, involving different types of information and different ways of exhibiting memory for them. Standard memory assessments have been developed, tested and applied with this understanding in mind [1-5]. For example, the Memory Assessment Scales measure Verbal, Visual, and Global memory [6] and the Wechsler Memory Scale [7] contains subtests for remembering personal and current information, mental control, digit span, logical memory, visual reproduction, and association learning. However, emerging research suggests that memory is even more nuanced than in the ways captured through standard neuropsychological testing, and in ways that are important to consider [8]. An example of a memory ability not captured by standard memory assessments is the phenomenon of Highly Superior Autobiographical Memory (HSAM), a rare condition in which a person can remember nearly every experience. When given a date and a year, someone with HSAM can state what day of the week that date fell on and what significant newsworthy events happened on that day, as well as other verifiable details such as what the weather was like [9, 10]. Despite this exceptional ability, people with HSAM tend to show unremarkable memory performance on standard neuropsychological assessments [11] and standard memory laboratory tasks [10]. Thus, whatever is at the heart of this exceptional memory ability is not captured in standard memory assessments. HSAM may result from superior consolidation of memories over time [12], which itself is not a part of typical memory assessment procedures. Below, I describe several more nuanced facets of human memory that are likely getting missed in standard memory assessments and for which there is reason to believe they are important.

2. Nuanced facets of human memory that are likely important

2.1. Memory for one's surroundings

One of the more nuanced aspects of human memory concerns the role that one's surroundings play in it [13]. The demonstrated role of scene information in

¹ Corresponding Author: Anne M. Cleary, Anne.Cleary@colostate.edu

autobiographical and episodic memory [14,15] and the shared neural mechanisms between spatial navigation and episodic memory ability [16] has led to the recent suggestion that episodic memory may be memory for the scene from one's surroundings at the time of the encoded experience [17]. However, as reviewed in [13, 15], nearly all standard memory tasks hold the person's surroundings constant throughout the task. The person is typically sitting in the same room throughout the encoding and retrieval portions of the memory task, either using a computer in front of them or with cards or pencil and paper. Their surroundings are unchanging throughout the memory task while memory for a series of things within that constant set of surroundings is examined. Even episodic memory is typically tested this way, with the most common procedure involving studying a list of words and then later being given a test a short time later on memory for those words. Yet, in real-life, different memories of different events take place in changing sets of distinct surroundings, and evidence suggests that the surroundings in which something occurred are a critical component of human memory [13, 14, 15].

It may not have been feasible in decades past to try to incorporate changing sets of surroundings into memory assessment tasks (as it would entail having to physically move across different settings as part of the memory assessment task); however, the advent of virtual reality (VR) makes this feasible today. The VR task used by [13] does so. In it, participants are immersed within randomly ordered successively changing scenes. Participants are later tested on their memory for scenes that earlier served as their VR surroundings using a combination of spatially similar and dissimilar scenes as cues and counterbalancing test scene cues across participants regarding whether they corresponded to studied or unstudied scenes. All facets of scene memory that were examined (e.g., cued recall of scenes, familiarity-detection with scenes spatially resembling encoded scenes, etc.) were enhanced when the scenes had been encoded as VR surroundings compared to when the scenes were encoded as images from a 2D screen in a room that was unchanging throughout the task. This suggests that it is not just scene *stimuli* that play a uniquely important role in memory, but scenes as *surroundings*, and this aspect of memory can be captured using VR.

2.2. Familiarity-detection

Another of the more nuanced aspects of human memory that is not usually considered in assessment is familiarity-detection, which refers to a general sense of recognition that something was encountered in the past, as opposed to recalling a specific instance in which that situation was previously encountered. As it is thought to result from global resemblance to information stored in memory [18], familiarity-detection can occur when a situation resembles a previously encountered situation, and it can occur even without recalling any specific prior relevant experience [13, 19-23].

Although the ability to detect familiarity with something despite failing to recall specifics might seem inconsequential to overall memory function, increasing evidence suggests familiarity-detection is likely an important component to overall memory function. Not only might intact familiarity-detection ability during other memory impairment be useful for rehabilitative purposes, but feelings of familiarity during recall failure are associated with likely proxies of memory search effort, such as an increased tendency to generate candidate information (commission errors), to devote increased time trying to retrieve the critical piece of information, and to exhibit elevated curiosity [13, 24]. An initial familiarity-detection mechanism may indicate the presence of something in memory before it gets retrieved, and such familiarity-detection may even determine whether or not an *attempt* to retrieve occurs [25, 26]. As such, familiarity-detection may be a significant modulator of attention, “flipping” attention inward toward retrieval search effort when familiarity is detected [27]. Whereas factors like novelty and salience draw attention outward toward the novel or salient stimulus for encoding, familiarity-detection instead likely sends attention inward toward memory retrieval search effort. This assertion is consistent with recent evidence suggesting that attention switches from being externally oriented to being oriented internally toward memory [28] as well as with neural evidence suggesting that the same neural circuit is involved in mutually exclusive hippocampal encoding vs. retrieval modes [29].

Thus, an overlooked possibility in memory assessment is that familiarity-detection is a critical neural mechanism that modulates the direction of attention in ways potentially important to the rest of cognitive function. Methods of measuring familiarity-detection when recall fails [13, 18-24] hold promise for tapping into this mechanism and its potential to drive attention inward toward memory search effort. A failure to detect familiarity in the environment may underlie some memory and attentional deficiencies.

2.3. Tip-of-the-tongue experiences

Another nuanced aspect of human memory not typically considered in assessment is the subjective sensation of a tip-of-the-tongue (TOT) experience, which occurs when a person feels on the verge of retrieving a word [30]. In the cognitive psychology literature, TOTs are measured by directly asking people if they feel that they are experiencing a TOT state for a given target word [30, 31]. TOTs are particularly relevant to cases of word anomia (clinical cases of extreme word retrieval failure). However, in assessments used in the clinical literature, rather than directly asking people for their subjective reports regarding whether or not a TOT state is present for the unretrieved word, TOTs are instead inferred in naming tasks on the basis of other potential proxies for TOT states, such as an ability to articulate partial target information or to gesture relevant information about the target [32, 33]. This type of indirect inference-based measuring of TOTs is problematic because research has demonstrated differences between partial attribute generation and the subjective sensation of a TOT state, and that TOT states are associated with an increased tendency to simply guess at candidate partial attributes without being correct [31]. As past clinical work has suggested that people with left Temporal Lobe Epilepsy (TLE) are more prone to word anomia than people with right TLE [33], it is important to distinguish between word anomia that is mere word retrieval failure unaccompanied by a subjective TOT state and word anomia that is word retrieval failure accompanied by a subjective TOT state. A reason why this is important is because the subjective sensation of a TOT state has useful qualities, such as prompting motivation, curiosity, and persistence at trying to find the answer [34, 35], which in turn could be capitalized on in rehabilitative strategies for patients with anomia. A way to address this gap in the future would be to incorporate a simple question into standard naming tasks used in memory assessment so that every time a person fails to retrieve a word in the naming task, they are asked if they are experiencing a TOT sensation for the word.

2.4. Déjà vu

Déjà vu—the striking feeling of having experienced something before when that seems impossible—is a facet of human memory that is rarely formally considered in memory assessment. Although seemingly just an interesting memory quirk or oddity, it actually has great clinical relevance. One of the most well-documented ways that déjà vu is relevant to assessment is in regard to medial temporal lobe (MTL) seizure disorders. Frequent déjà vu is often an indicator of seizure activity originating in the MTL [36]. Thus, when used in conjunction with other assessments, the frequency with which a person experiences déjà vu can be a useful indicator of the status of the neural circuitry underlying their memory system and of whether or not the person might be experiencing subclinical or focal seizures that might otherwise go undetected. A measure of a person's déjà vu frequency commonly used in research is the Inventory for Déjà vu Experiences Assessment (IDEA) [37]. This assessment tool could be used alongside other assessment measures to gain a fuller picture of a person's memory function and neural status.

2.5. The Multi-modal nature of memory

Finally, another facet of human memory that should be considered for the future of memory assessment is the multi-modal nature of memory. Memory in day-to-day life does not occur for individual pieces of information that occur in isolation. Instead, memory is multi-faceted, involving encoding one's surroundings and one's place within those surroundings at the same time as encoding surrounding people's identities, what people are saying, people's voices, the soundscape accompanying the visual surroundings, the emotions accompanying the experience, and many other facets. Thus,

memory assessments that include a multi-modal informational format could potentially detect more nuanced memory difficulties while also presenting more potential avenues for rehabilitative strategies. A step in this direction has been developed by [8]. They created an immersive gaming environment where the participant experiences a real-world scene in which someone is talking to them and events are happening (such as someone in the background shooting basketball hoops). This method allows for multiple aspects of memory to later be tested, such as face recognition, name recall ability, and object location memory, among others, and thus holds promise as a future memory assessment method.

3. Conclusion

The discovery through minimally invasive surgical interventions for the treatment of epilepsy that affected aspects of memory can go undetected by standard memory assessments [8], along with a number of new findings in cognitive psychology [27, 13, 34] suggest a need for new approaches to memory assessment that examine more nuanced facets of memory as well as that examine memory in more life-like contexts. A number of promising approaches are laid out here that have the potential to expand the current repertoire of available memory assessment tools to help provide a more comprehensive picture of a person's overall memory function and likely neural status. In addition to the above-mentioned suggested avenues, future memory assessment should also incorporate longer-term memory testing to include memory consolidation mechanisms over time as part of the assessment process as well.

References

- [1] Donders J. The incremental value of neuropsychological assessment: A critical review. *Clin Neuropsychol*. 2020;34(1):56–87. doi: <https://doi.org/10.1080/13854046.2019.1575471>
- [2] Donders J, Maher T. Clinical utility of the Neuropsychological Assessment Battery Judgment Subtest in the evaluation of older adults. *Arch Clin Neuropsychol*. 2024 May 6:acac035. doi: <https://doi.org/10.1093/arclin/acae035>
- [3] Drane DL, Willie JT, Pedersen NP, Qiu D, Voets NL, Millis SR, et al. Superior verbal memory outcome after stereotactic laser amygdalohippocamotomy. *Front Neurol*. 2021;12:779495. doi: <https://doi.org/10.3389/fneur.2021.779495>
- [4] Kostic B, Cleary AM. Memory assessment. In: Leong FTL, editor. *Encyclopedia of Counseling Psychology*. Sage Publications; 2008.
- [5] Loring DW, Papanicolaou AC. Memory assessment in neuropsychology: theoretical considerations and practical utility. *J Clin Exp Neuropsychol*. 1987;9(4):340–58. doi: <https://doi.org/10.1080/01688638708405055>
- [6] Loring DW, Hermann BP, Lee GP, Drane DL, Meador KJ. The Memory Assessment Scales and lateralized temporal lobe epilepsy. *J Clin Psychol*. 2000;56(4):563–70. doi: [https://doi.org/10.1002/\(sici\)1097-4679\(200004\)56:4<563::aid-jclp9>3.0.co;2-k](https://doi.org/10.1002/(sici)1097-4679(200004)56:4<563::aid-jclp9>3.0.co;2-k)
- [7] Ryan JJ, Kreiner DS, Gontkovsky ST, Teichner G. Wechsler memory scale-fourth edition (WMS-IV) in the neuropsychological evaluation of patients diagnosed with probable Alzheimer's disease. *Appl Neuropsychol Adult*. 2023;1–8. doi: <https://doi.org/10.1080/23279095.2023.2164858>
- [8] Okada N, Rajeh Y, Skodje T, Brady E, Shade T, Ravikanti K, et al. A gamified framework for multimodal memory assessment in patients undergoing epilepsy surgery. *Am Epil Soc*. 2023;1012. Available from: <https://aesnet.org/abstractslisting/a-gamified-framework-for-multimodal-memory-assessment-in-patients-undergoing-epilepsy-surgery>
- [9] Parker ES, Cahill L, McGaugh JL. A case of unusual autobiographical remembering. *Neurocase*. 2006;12(1):35–49. doi: <https://doi.org/10.1080/13554790500473680>
- [10] LePort AK, Mattfeld AT, Dickinson-Anson H, Fallon JH, Stark CE, Kruggel F, et al. Behavioral and neuroanatomical investigation of highly superior autobiographical memory (HSAM). *Neurobiol Learn Mem*. 2012;98(1):78–92. doi: <https://doi.org/10.1016/j.nlm.2012.05.002>
- [11] LePort AK, Stark SM, McGaugh JL, Stark CE. A cognitive assessment of highly superior autobiographical memory. *Memory*. 2017;25(2):276–88. doi: <https://doi.org/10.1080/09658211.2016.1160126>
- [12] LePort AK, Stark SM, McGaugh JL, Stark CE. Highly superior autobiographical memory: quality and quantity of retention over time. *Front Psychol*. 2016;6:2017. doi: <https://doi.org/10.3389/fpsyg.2015.02017>
- [13] Okada NS, McNeely-White KL, Cleary AM, Carlaw BN, Drane DL, Parsons TD, et al. A virtual reality paradigm with dynamic scene stimuli for use in memory research. *Behav Res Methods*. 2023. doi: <https://doi.org/10.3758/s13428-023-02243-w>
- [14] Rubin DC, Deffler SA, Umanath S. Scenes enable a sense of reliving: implications for autobiographical memory. *Cognition*. 2019;183:44–56. doi: <https://doi.org/10.1016/j.cognition.2018.10.024>
- [15] Rubin DC, Umanath S. Event memory: a theory of memory for laboratory, autobiographical, and

- fictional events. *Psychol Rev.* 2015;122(1):1–23. doi: <https://doi.org/10.1037/a0037907>
- [16] Maguire EA, Mullally SL. The hippocampus: a manifesto for change. *J Exp Psychol Gen.* 2013;142(4):1180–9. doi: <https://doi.org/10.1037/a0033650>
- [17] Rubin DC. A conceptual space for episodic and semantic memory. *Mem Cognit.* 2022;50(3):464–77. doi: <https://doi.org/10.3758/s13421-021-01148-3>
- [18] McNeely-White KL, McNeely-White DG, Cleary AM. Global matching in music familiarity: how musical features combine across memory traces to increase familiarity with the whole in which they are embedded. *J Mem Lang.* 2021;118:104217.
- [19] Cleary AM, Claxton AB. Familiarity from the configuration of objects in 3-dimensional space and its relation to déjà vu: a virtual reality investigation. *Conscious Cogn.* 2012;21(2):969–75. doi: <https://doi.org/10.1016/j.concog.2012.01.008>
- [20] Cleary AM, Claxton AB. Déjà vu: an illusion of prediction. *Psych Sci.* 2018;29(4):635–44. doi: <https://doi.org/10.1177/0956797618756191>
- [21] Huebert AM, McNeely-White KL, Cleary AM. Can cue familiarity during recall failure prompt illusory recollective experience? *Mem Cognit.* 2022;50:1–15. doi: <https://doi.org/10.3758/s13421-022-01337-4>
- [22] Ryals AJ, Cleary AM, Seger CA. Recall versus familiarity when recall fails for words and scenes: the differential roles of the hippocampus, perirhinal cortex, and category-specific cortical regions. *Brain Res.* 2013;1492:72–91. doi: <https://doi.org/10.1016/j.brainres.2012.11.001>
- [23] Carlaw BN, Huebert AM, McNeely-White KL, Rhodes MG, Cleary AM. Detecting a familiar person behind the surgical mask: recognition without identification among masked versus sunglasses-covered faces. *Cogn Res Princ Implic.* 2022;7(1):90. doi: <https://doi.org/10.1186/s41235-022-00440-3>
- [24] McNeely-White KL, Cleary AM. Piquing curiosity: déjà vu-like states are associated with feelings of curiosity and information-seeking behaviors. *J Intell.* 2023;11(6):112. doi: <https://doi.org/10.3390/intelligence11060112>
- [25] Nhousyvanisvong A, Reder LM. Rapid feeling-of-knowing: a strategy selection mechanism. In: *Metacognition: Cognitive and Social Dimensions.* 1998;35:52.
- [26] Paynter CA, Reder LM, Kieffaber PD. Knowing we know before we know: ERP correlates of initial feeling-of-knowing. *Neuropsychologia.* 2009;47(3):796–803. doi: <https://doi.org/10.1016/j.neuropsychologia.2008.12.009>
- [27] Cleary AM, Irving ZC, Mills C. What flips attention? *Cogn Sci.* 2023;47(4):e13274. doi: <https://doi.org/10.1111/cogs.13274>
- [28] Servais A, Hurter C, Barbeau EJ. Attentional switch to memory: an early and critical phase of the cognitive cascade allowing autobiographical memory retrieval. *Psychon Bull Rev.* 2023;30(5):1707–21. doi: <https://doi.org/10.3758/s13423-023-02270-w>
- [29] Barbeau EJ, Chauvel P, Moulin CJ, Régis J, Liégeois-Chauvel C. Hippocampus duality: memory and novelty detection are subserved by distinct mechanisms. *Hippocampus.* 2017;27(4):405–16. doi: <https://doi.org/10.1002/hipo.22699>
- [30] Schwartz BL. The relation of tip-of-the-tongue states and retrieval time. *Mem Cognit.* 2001;29:117–26.
- [31] Huebert AM, McNeely-White KL, Cleary AM. On the relationship between tip-of-the-tongue states and partial recollective experience: illusory partial recollective access during tip-of-the-tongue states. *J Exp Psychol Gen.* 2023;152(2):542–70. doi: <https://doi.org/10.1037/xge0001292>
- [32] Hamberger MJ, Seidel WT. Auditory and visual naming tests: normative and patient data for accuracy, response time, and tip-of-the-tongue. *J Int Neuropsychol Soc.* 2003;9(3):479–89. doi: <https://doi.org/10.1017/S135561770393013X>
- [33] Trebuchon-Da Fonseca A, Guedj E, Alario FX, Laguitton V, Mundler O, Chauvel P, et al. Brain regions underlying word finding difficulties in temporal lobe epilepsy. *Brain.* 2009;132(10):2772–84. doi: <https://doi.org/10.1093/brain/awp083>
- [34] Cleary AM, McNeely-White KL, Russell SA, Huebert AM, Hausman H. The tip-of-the-tongue state as a form of access to information: use of tip-of-the-tongue states for strategic adaptive test-taking. *J Appl Res Mem Cogn.* 2021;10(1):131–42. doi: <https://doi.org/10.1016/j.jarmac.2020.08.013>
- [35] Metcalfe J, Schwartz BL, Bloom PA. The tip-of-the-tongue state and curiosity. *Cogn Res Princ Implic.* 2017;2(1):31. doi: <https://doi.org/10.1186/s41235-017-0065-4>
- [36] Cleary AM, Brown AS. The Déjà Vu Experience. 2nd ed. Routledge; 2022.
- [37] Sno HN, Schalken HF, de Jonghe F, Koeter MW. The inventory for déjà vu experiences assessment: development, utility, reliability, and validity. *J Nerv Ment Dis.* 1994;182(1):27–33. doi: <https://doi.org/10.1097/00005053-199401000-00006>

Found Value in the Artifice of Semblance: Engaging With Abstraction Within Generative AI and Integration Practice in PAT

Xavier NOKES^{a,1} and Pavan TURAGA^b

^a*Arizona State University*

Abstract. Abstract goes here. Artificial intelligence (AI) and psychedelic-assisted therapy (PAT) exist in dynamic interstices (real/unreal, professional/recreational, critiqued/lauded), and shared amongst them is the rapidity of development and implementation. We find potential for exploration in the ebb and flow of affordances and criticisms that these spaces present one another. In PAT, integration is defined as the processes by which an individual assimilates emotions, sensations, and insights experienced in non-ordinary states of consciousness (NOSC). Within the context of this work, we will specifically detail image-making as an example. Integration is distinct in that its value is particular to the individual, intrinsic to ones' perspective, existing without precedent due to the often ineffable nature of NOSCs. Within AI-aided creation, there is consensus that successful results demonstrate high fidelity and accuracy to the prompt. Many results incentivize re-engagement, taking the form of prompt refinement, or starting over; the "success" of a prompt is in the hands of the writer, but what is consistent is the process of refinement. Jux-taposing the meaning-making and valuation process of integration (hyper-specific, yet intractable) with AI content generation (seemingly infinite, highly directed) we can extrapolate unique threads of interest, specifically analyzing gatekeeping, accessibility, (non) sense-making, and knowledge regulation. We also present the early results of a preliminary study ($N=34$) to assess whether contemporary Generative AI tools hold value for self-identified users of psychedelics toward integration practice.

Keywords. Psychedelic, image generation, machine learning, non-ordinary states of consciousness, abstraction, generative AI

1. Introduction

In this work, we discuss the potential for Generative AI's (GenAI) use in conjunction with integration practices, with particular interest in those which considered inadequate for industry-grade photorealistic image generation. As neither author is a trained practitioner, we instead speak to this as a prospective, determined as viable due to the historical utilization of psychedelic substances prior to Western medical practices discriminatory regulation and co-opting of traditional practices ascribed to various Indigenous communities. This cultural significance acts two-fold, as Xavier Nokes is an Indigenous person with ties to tribes and lands wherein plant medicines have and continue to be utilized, as well as being a major component of their academic research. Pavan Turaga is a person of Indian (subcontinent) origin, with expertise in AI methods. We do not aim to speak for any particular tribe, nation, or community, but solely as individuals, aiming to position our current understanding of the significance of integration when utilized alongside psychoactive substances for clinical and personal (i.e. recreational and/or self-healing) use, as well as look at the prospect of creating new studies.

Psychedelic assisted treatment (PAT) is a burgeoning subcategory of the medical field – particularly for use in mental health (additionally, treatments with ketamine have begun showing results in assisting with physical pain [1]), which emphasizes the use of psychoactive substances to increase the effectiveness, speed, and overall impact of treatment. The efficacy of PAT is generally attributed to the use of entheogens/plant medicines (such as psilocybin mushrooms ("magic mushrooms"), peyote, ayahuasca),

¹ Corresponding Author: Xavier Nokes, xavier.nokes@asu.edu

synthetic substances (such as LSD and MDMA), and chemical substances that induce non-ordinary states of consciousness (NOSC); while primarily known for their emotional and cognitive impact, NOSCs can also be categorized by their somatic, temporal, and social components [2]. For mental health-related use cases, the process of *integration* is a fundamental component of treatment as it provides means by which to disseminate meaning from one's NOSC [3], often via creative practices. While integration can and does exist in other modalities, this paper focuses on the utilization of integration as it pertains to PAT.

Integration is intrinsically tied to the (most) successful processing of NOSCs, noted for their sensorial impact, sometimes be noted as “awe-inducing” or “ineffable” [4], often the reasoning behind their challenging nature to integrate. Integration is poised as a meaning-making system, intrapersonal and instantiated by the actionable processes that define it [5]. While multi-modal, it also is a dedication, with successful cases of integration continuing beyond the initial and becoming regularly operative elements of participants' lives. Here we begin to see the correlation in the processes of integration and GenAI, conceptually similar not in the practice but in their intent. Delineation, then, comes from the consensus of measured *success*; for integration, success is attributed to the individual, whereas success of GenAI is in relevancy to the prompt and legibility to the human user.

In the remainder of the work, we trace the developments in generative algorithms specifically for visual media over the past few years and discuss how there is value in algorithms for generation that create physically implausible imagery (i.e. nonsense). We discuss the value in which integration in PAT might find value and support in the seemingly nonsensical creations of GenAI. It is through this lens that we propose new potentials alongside existing traditional practices, moving toward self-determined meaning making instead of utopianism, as well as co-creation with emerging technologies.

2. Understanding PAT and Generative AI

2.1. Current Practice of PAT in Clinical Studies

While entirely dependent on the facilitator of the integration of a NOSC, there are various means by which an individual may integrate an experience. Typically, the success of integration correlates with the participant's sensory experience in the NOSC, the means by which they best process and sense-make, as well as their general life experience. Dance, music making, and painting are but some of the means by which integration is deployed [2]. For many Indigenous cultures, from whom the utilization of plant medicine is derived, there are societally ingrained formats of “integration” that lack distinction due to the inherent function it has in these cultures [6]. Sense-making in such communities can take the form of a leader who can elucidate one's experience, the individual being tasked with deciphering meaning, or something else entirely.

As previously stated, the processes in which PAT and integration are performed are dependent on multiple factors; “set and setting” is a commonly used phrase concerning NOSC experiences, referring to one's *mindset*, and the environment they are in under an induced state. While variation is important to note, there is a general procedure attributed to the clinical use of psychedelics; to start, participants are subject to an intake evaluation, wherein the facilitators will gauge the mental and physical state of the participant. Participants may be “outpatient” and come and go from the facilities, or they may be “inpatient”, staying onsite even outside of sessions. Facilitators will determine the means of providing the substances to participants, which may take the form of smaller doses over a longer duration, and/or lead to a larger “heroic/therapeutic” dose. Many Western medical facilitators utilize a therapeutic dyad (male/female) during the process, though recently there has been concern and re-evaluation of this as a standard [7]. The process of integration may take place in conjunction with when an NOSC takes place, and/or continue on in various forms *after*, such as either a self-guided process or alongside the facilitator or integration therapist [2].

2.2. Current State of Generative AI

Current state-of-the-art generative models are several, however, one major category referred to as “Generative Adversarial Networks (GANs)” offer many insights in how to use them for exploring the balance between photorealism and abstraction. A standard GAN [8] consists of a generator and a discriminator subnetwork. The generator network is characterized by parameters that allow it to generate random pictures at first. The generated pictures are examined by the discriminator, which is also provided access to a separate training-set of pictures that are considered ‘real’. The discriminator compares the generator’s pictures to the training set and provides its ratings to the generator’s pictures, roughly intended as a ‘1’ if a picture looks similar to the training-set, otherwise ‘0’. Clearly, there are many detailed choices involved in how pictures are measured against a training-set, and what aspects of the pictures are considered essential to be matched. Generally, these questions do not have very satisfactory answers, as neural networks are parameterized by millions/billions of parameters, and it has not been feasible to interpret what neural networks are paying attention to. As training progresses, the generator learns to create pictures that look as-if they could be a plausible picture from the training-set.

This basic architecture is further enhanced with a text-processing layer, that can map textual entries to neural features learned by a correlational approach, via the so called ‘Contrastive Language Image Pairs (CLIP)’ approach [9]. CLIP embeddings are trained to map textual entries and neural network features to a common space, so that when we type in ‘Van Gogh’ as a style prompt, the network learns to create an intermediate representation that correlates highly with image-features that would be found if one were to actually process Van Gogh art via a neural network. These correlations are usually learned by paired images and image-captions from web-scale data. By ‘features’, the approaches mean concepts like colors, textures, strokes, etc.

Currently there is significant interest in creating photorealistic generative models, aimed to create images indistinguishable from real-world pictures. This endeavor is motivated by the broader autonomous systems industry for applications including self-driving vehicles, defense applications, surveillance, human-robot teams, etc. Examples of perceptual tasks that are important in this endeavor are object recognition, 3D model estimation, human activity recognition and more. This ‘autonomous surveillance’ industry has been active for decades, and the bottleneck has been the availability of large annotated datasets to train these perception modules on. Generative methods are now being seen as a way forward, when acquiring real data is expensive or otherwise not feasible [10]. In these cases, the end-goal is photorealism with detailed and controllable generation. However, this is a shifting goal, and methods for generation still seem to have difficulties in this process.

With the understanding that success of GenAI, is predominantly when the output showcases adherence to specificity, we can surmise that the initial attempt is often *not* capable of generating the desired outcome, instead producing more symbolic representations; in consequence, AI prompt writing has become a practice in and of itself, necessitating alteration and refinement in extrapolating the exactly defined from the digital ether. Outside the implicit value of redirection and refinement with which to hone the outcome, we might then ask what value might there be in these so-called “failures”, that generated prior to a final “success”?

2.3. Engaging with Abstraction

Perception comes to the forefront in both of these spaces; for integration, it is through one’s own perspective that they experience and engage with the material that comprises a NOSC. For GenAI, often the model itself is built upon a number of inputs by which it can generate an output. Comparable is a method by which some might integrate an NOSC via shared narrative, wherein participants are asked to share within a larger group. This process is meant as explicitly uni-directional, as participants are not asked to make sense of someone else’s NOSC, but instead serves as a space where one might pull fragments into their own meaning making process. Reflecting on Karol Berger’s text “The Arts Today” from *Are the Arts Essential*, discussing the evolution of art practice through the 18th and 19th century, this approach offers something interesting. Berger states that artists of the time “did not have to worry about the subject matter of their next work”

[11], often merely tasked with the “mimetic” practice of creating representations. This shift in the archetype of the artist lambasted the status of the creator as merely a specialized transcriber, tasking them now with a more substantially challenging role of conlanger. Those utilizing integration are not making *something* from *nothing*, but instead using symbolic “nonsense” as a material for self-directed meaning making.

We can see a clear example of *representation* in GANs, tasked with simply understanding what something is or isn’t, visually speaking. A generated outcome would be technically classified as representational, as it has no means of perceptually delineating something that may instead be more accurately described as an abstraction. We might even find this so-called representation to be lacking, which could derive from any number of factors, such as the dataset informing the model. It is important to be reminded that the datasets that these generative systems are trained upon are lacking both in quality and quantity: this ties directly into integration as a process, extracting meaning for oneself in lieu of accepting what is presented. Visuals experienced in a NOSC are typically more in line with the general understanding of abstraction, with participants in studies sharing everything from smoking gargoyles [4] to the embodiment of gender dysphoria [12], much like the “nonsense” generated prior to the “successful” output. From one person to another, this imagery may be difficult to make sense of, even incomprehensible; but what it does do, is offer the potential to serve as a more personal avenue to explanation and understanding.

3. Preliminary Survey Study

In this initial survey, we chose to pose questions to better understand the *general* feelings of respondents regarding GenAI’s use as a tool for integration in PAT. The questions posed in the survey are as follows:

- Age/Gender
- Have you used psychedelic substances (“magic” mushrooms, LSD, MDMA, entheogens, etc) with the intent of healing, wellness, treatment, or similar?
- In what capacity was the substance(s) used? Select all that apply (Responses: Medicinal/Therapeutic/PAT, Personal, Ceremonial, Other)
- Did you participate in any form of integration?
- (If “Yes”) Having participated, do you feel that your participation was helpful? (Responses: Yes, No, Unsure, Other)
- (If “No”) Having not participated, do you wish you had or had the opportunity to participate in some form of integration? (Responses: Yes, No, Unsure, Other)
- Have you used any form of generative AI? Examples include Dalle, Midjourney, ChatGPT, Runway ML, or similar.
- In what capacity have you used such platforms? Please select all that apply (Responses: Work/Professional, Academic, Personal, Other)
- Try DreamStudio (Responses: Image Upload, Prompt Used, Negative Prompt Used (if any))
- Do you feel that generative AI may provide participants of PAT (or other, similar practices) with additional benefit in creating imagery and integrating their experience(s) with psychedelics? (Response: Yes, Maybe, No, Other)
- Please provide any additional comments relating to the ideas of this survey.

3.1. Respondents Demographic Information

In our survey, we asked participants to provide their opinion regarding GenAI-aided integration. This survey is ongoing, but as of this writing we have received 34 responses. Respondents’ ages show a mean \pm std. dev. of 37.64 ± 11.67 . Gender was split 13:20 (female/male) with 1 respondent choosing to withhold. Participants were asked their familiarity and/or experience with three areas related to this writing; psychedelic substances for health, integration in relation to the use of psychedelics, and GenAI. Additionally, respondents were optionally allowed to use the platform “DreamStudio”

[13] to create and upload their own GenAI image(s), along with any prompts used. We show some sample imagery created by participants, along with prompts, in **Figure 1**.



Figure 1. (Top Left) Prompt: “a hyperrealistic illustration of a fork stuck in the soil with flowers growing out of it at the tip of the fork, along with other flowers around”, Negative Prompt: “upside-down fork, cartoon style, deformed fork, multiple forks, plain background”. (Top Right) Prompt: “Life energy flowing through veins, translucent skin, glowing” (Bottom Left) Prompt: “faceless beings in a warm (purple) environment pushing cart full of money, boxes are on the flat cart. you can see the money because they’re enclosed. There are an army of them. they rule the outer dimensions. they are like blobs, but human shaped so to speak. they are in a line. they are not wearing clothes and their faces are the same (color) as their bodies. the outside is slightly warmer than their bodies. instead of the details in the clothes, make them human shaped “bubbles”, Negative Prompt: “dont darken the face, and don’t expose the money” (Bottom Right) Prompt: “Watercolor of a duck enjoying a day at the pool”, Negative Prompt: “No rubber ducks or toy ducks.”

3.2. Use of Psychedelics

Of the 34 collected responses, 17 of the respondents noted use of psychedelics in this capacity, with only 5 of those respondents noting participation in some form of integration. Of those who had not, 7 responded that they would have/would like to, 4 responding that they were unsure, with the remaining responses left blank.

3.3. Use of GenAI

Regarding the use of GenAI platforms, all but 5 respondents stated they had used GenAI. A majority of 23 respondents stated use of GenAI personally, 22 stating use in work/professional capacities, and 13 stating academic use. It should be noted that the majority of respondents had reported use of GenAI in multiple capacities.

3.4. Use of GenAI for Integration

Regarding respondents’ feelings towards the potential of GenAI-aided integrative practices, responses were varied. Of the 34 surveyed, 16 responded favorably, 13 responded maybe, 4 responded in the negative, and one written response. Of those responding affirmatively, respondents stated that they felt the idea to be “very interesting and promising”, as well as able to “assist (ones) journey”. The 4 who responded negatively, stated that they felt the tools (GenAI platforms) to be inadequate, stating that they “struggled...to get my image to look like what I was envisioning”, and another who stated that their negative prompt was not accounted for. Additional responses reflected that a “need to teach good prompt writing” may have an impact, while one respondent

was very explicit in their frustration over their inability to create the imagery that they were wanting.

4. Conclusions

For those familiar with GenAI output, it is not surprising to see that the general push back regarding the use of GenAI (within this particular context) was more heavily focused on the *inability* of the tool to create exact imagery. In light of this, we believe that our initial question as to the value of abstraction within the context of treatment will present an interesting counterpoint to explore in further studies. Through the initial survey detailed in this writing, we see the responses as a precursor to the capabilities of presumed incapability, particularly as more and more AI platforms present their metric of success through their ability to create high-fidelity and prompt-accurate outcomes.

References

- [1] Yang Y, Maher DP, Cohen SP. Emerging concepts on the use of ketamine for chronic pain. *Expert Rev Clin Pharmacol*. 2020;13(8):835–854. doi: <https://doi.org/10.1080/17512433.2020.1804555>
- [2] Aixalà M, Carlos B. *Psychedelic integration: Psychotherapy for non-ordinary states of consciousness*. Santa Fe, NM: Synergetic Press; 2022.
- [3] Timmermann C, Watts R, Dupuis D. Towards psychedelic apprenticeship: Developing a gentle touch for the mediation and validation of psychedelic-induced insights and revelations. *Transcult Psychiatry*. 2022. doi: <https://doi.org/10.1177/13634615221082796>.
- [4] Pollan M. *How to Change Your Mind*. New York: Penguin Press; 2018.
- [5] Gorman I, Nielsen EM, Molinar A, Cassidy K, Sabbagh J. Psychedelic harm reduction and integration: A transtheoretical model for clinical practice. *Front Psychol*. 2021;12:648243. doi: <https://doi.org/10.3389/fpsyg.2021.648243>.
- [6] Carod-Artal FJ. Hallucinogenic drugs in pre-Columbian Mesoamerican cultures. *Neurología (English Edition)*. 2015;30(1):42–49. doi: <https://doi.org/10.1016/j.nrleng.2011.12.012>
- [7] Belsler A, Cavnar C, Labate B. *Queering Psychedelics: From Oppression to Liberation in Psychedelic Medicine*. Santa Fe, NM: Synergetic Press; 2022.
- [8] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville AC, Bengio Y. Generative adversarial networks. *Commun ACM*. 2020;63(7):139–144. doi: <https://doi.org/10.1145/3422622>
- [9] Radford A, Kim JW, Hallacy C, Ramesh A, Goh G, Agarwal S, Sastry G, Askell A, Mishkin P, Clark J, Krueger G, Sutskever I. Learning transferable visual models from natural language supervision. In: Proceedings of the 38th International Conference on Machine Learning; 2021; PMLR.
- [10] Zewi A. In machine learning, synthetic data can offer real performance improvements [Internet]. MIT News; 2022 [cited 2024 Jan 12]. Available from: <https://news.mit.edu/2022/synthetic-data-ai-improvements-1103>.
- [11] Berger K. *The Arts Today - Are the Arts Essential?* New York: New York University Press; 2022.
- [12] Angieri T. *Integrate.Me* [Internet]. PBS; 2021. Available from: <https://www.pbs.org/pov/films/integrateme/>.
- [13] DreamStudio. About DreamStudio [Internet]. 2024. Available from: <https://dreamstudio.com/about/>.

Bridging the Gap Between Perceived and Actual Personalities in Real and Virtual Worlds Through EEG Analysis

Timothy MCMAHAN^{a,1}, Chotepong Victor TANGTON^a, Advay VYAS^a, Chloe HUANG^a, and Aron YEWOOR^a

^a University of North Texas, Department of Learning Technologies ORCID ID: Timothy McMahan
<https://orcid.org/0000-0002-4904-7307>

Abstract. Individuals are spending more time playing video games and socializing in virtual worlds. This opens the question of what personality are they showing in the virtual world? Virtual worlds offer the opportunity to be the person you are not in real-life, but which personality is their real personality. To analyze personality complexes, we developed a dynamic video game that utilizes data standardization based on multiple industry-standard personality tests. Moreover, an electroencephalography (EEG) system is used to indicate any abnormalities or identifiers of misleading results. After consulting standardized personality tests such as the Big-Five Personality Test, NEO Five-Factor Inventory (NEO-FFI), and Myers–Briggs Type Indicator (MBTI), our team crafted three real-world scenarios. These scenarios, embedded in a CLVG, are designed to answer the questions commonly found in standardized exams. There are thirty unique pathways the user can experience depending on their choices, each with a unique summation of calculated values for the six categories. We have hypothesized that the complex decision-making opportunities provided in the game will elicit a range of behaviors reflective of the players' true personalities, as opposed to their idealized self-reports.

Keywords. Neurogaming, personality assessment, video games, electroencephalography (EEG)

1. Introduction

Personality assessment has been a long basis of psychological research, offering invaluable insights into human behavior and correlations between actions and what it means about one's personality [1]. Traditional methods of personality evaluation, such as the Big five or the Myers–Briggs Type Indicator, aim to capture the core aspects of an individual's character through a questionnaire where individuals would respond to various questions that gauge their agreement to statements [2,3]. These assessments can be used in various settings, from clinical, occupational, to educational, influencing things such as therapeutic approaches to hiring decisions [4]. However, a notable discrepancy often arises between how individuals perceive how they think they would behave in a situation rather than how they would behave [5]. This causes their personality to conform into what they imagine their ideal self to be as it is hard to always accurately assume

¹ Corresponding Author: Timothy McMahan, fred.mcmahan@unt.edu.

what one would do in certain situations which are asked during these personality test questionnaires [6]. This gap becomes even more pronounced when contrasting behaviors in real-world settings with those in virtual environments, such as video games. For example, in real life, individuals might adhere more closely to social norms and personal ideals, whereas virtual settings such as video games can offer anonymity and freedom that encourage different aspects of personality to surface. A person who is typically reserved may become more assertive during a video game [7]. Virtual environments offer a unique perspective which challenges the traditional mold of traditional personality assessments, and this exploration becomes increasingly important with virtual interactions becoming more prevalent in our daily lives [8].

Our study aims to address these issues by developing a dynamic video game designed to assess personality in ways that traditional tests might not capture. By integrating real-time data collection tools such as electroencephalography (EEG), we seek to provide a more nuanced understanding of the fluidity and context-dependent nature of personality [9]. The use of electroencephalography (EEG) in video games is similar to methods in recent research, which shows a thorough analysis of EEG data in the context of gaming. A study using EEG uncovered 31 different processing techniques which could be applied to analyze brain activities during gameplay [10]. Using EEG in our study, we hope to find correlations and patterns we can derive that lead to certain decisions and what it may entail about one's personality as performed by similar work incorporating EEG and personality assessment [11, 12, 13].

Furthermore, recent research also delves into the realm of personality types and games, specifically into the gamification adapted for different personality tests [14]. Using the Myers-Briggs Type Indicator [15], the research showed the potential of tailoring game design to align with distinct personality traits. Although some projects use the Big Five [16] personality model instead, the principles derived from the BTI- focused studies provide valuable insights. This highlights the potential and importance of creating games that resonate with specific personality traits. The objective of this research is to investigate the variances between self-reported personality traits and how they differ from their in-game/virtual environment personality test which is deduced through our developed simulation. In developing our program, we adapted scoring methodologies from existing personality tests and created unique scoring systems for specific decisions made within the game. This approach allows us to explore personality dimensions in an immersive environment.

2. Virtual Personality Simulator (VPS)

In the development of our video game simulation, the Virtual Personality Simulator (VPS), our primary objective was to design an experience that did not overtly appear to be assessing the player's personality. Utilizing a Role-Playing Game (RPG) system, we aimed to craft an engaging game that subtly integrated personality assessment mechanisms within its gameplay. This design strategy was intended to ensure that participants remained unaware of the underlying psychological evaluation, thereby eliciting more natural and spontaneous behaviors. The game's scenarios were constructed by cross-referencing existing personality tests, focusing on identifying significant behavioral traits such as greed, competitiveness, and morality. VPS uses a single-player model with non-player characters (NPCs) designed to mimic social interactions as

inspired by what a player might find in an online multiplayer game. Figure 1 shows a mapping of the player's game decisions to the scoring of the personality assessment.

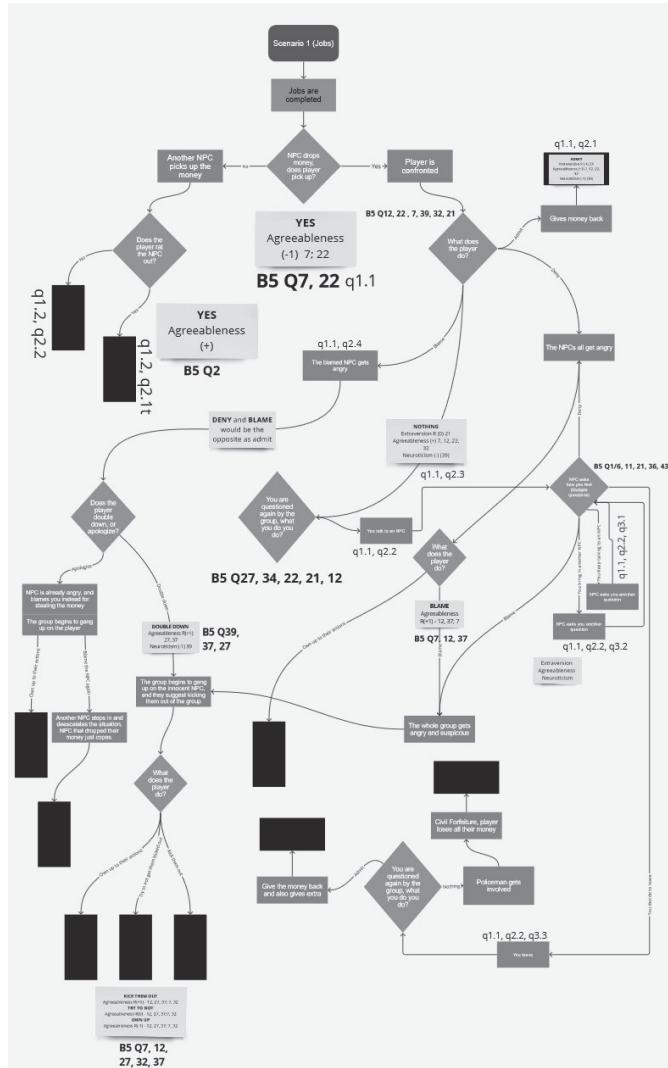


Figure 1. Mapping of scenario 1 decisions to personality assessment calculations.

Utilizing the Unity game engine, the simulation is set within a corporate environment, revolving around the goal of career advancement and financial success. This setting naturally causes situations where players must navigate morally complex decisions, such as choosing whether to collaborate or compete, deceive, or sabotage others to advance their own status. A critical aspect of VPS involves incorporating morally ambiguous choices to probe the depths of the players' ethical boundaries. We

also test the extent of the player's morality by seeing how far they would go to lie and deceive to save themselves even after being caught. The consequences of such choices are designed to simulate real-world repercussions, enhancing the immersion and psychological realism of the scenario. NPCs in the game are endowed with distinct personalities and behavioral traits, ranging from kindness and timidity to rudeness and aggression. This diversity allows us to observe if and how a player's interactions vary with NPCs based on their perceived personalities. By blending immersive gameplay with nuanced psychological assessment, our simulation aims to provide a comprehensive analysis of personality in a manner that traditional tests cannot, leveraging the dynamic and interactive nature of video games to explore the complexities of human behavior in a controlled yet realistic setting.

2.1. Scenario One

Scenario one initiates with the player's introduction to their NPC coworkers, effectively setting the stage for the exploration of workplace dynamics. Figure 2 shows the starting point of the player in the office complex. This initial meeting is crucial as it showcases the varied personalities within the NPC cohort, ranging from the timid and amiable to the overtly competitive and abrasive. The player's first assignment involves choosing a job from a variety of available tasks that differ in complexity and potential earnings. Following job selection, players engage in the tasks they have chosen. The outcome of these tasks directly affects their compensation. Following this segment of Scenario one, the focus shifts to examining the player's inclinations towards greed and competitiveness through a series of subtly designed decisions. The situation unfolds when an NPC, following the game's script, inadvertently drops money. This moment presents a crucial decision point for the player, framed through an interactive menu with options that are not clearly morally binary, enhancing the ambiguity and complexity of the choices. If the player opts to return the money, it can reflect positively on the player's agreeableness score. However, if the player bypasses the money, another NPC will collect it. The player is then prompted to decide whether to report this action to authorities, which subtly tests their inclination towards justice or non-interference, further affecting their agreeableness. If the player chooses to take the money, the ensuing decisions become increasingly intricate and morally grey. They are soon confronted about their actions, leading to a sequence of choices that test their ability to navigate social and moral pressures.



Figure 2. Scenario 1 scene of the office complex where the player first starts.

2.2. Scenario Two

Scenario two delves into the player's behavioral tendencies within a collaborative setting by focusing on interactions within a team environment. The scenario begins with players being given the choice to either become the team captain or remain as a team member, which sets the stage for different developmental paths and personality assessments.

The role of team captain is designed to evaluate leadership qualities such as extraversion, agreeableness, and leadership skills, placing the player in a position to make decisions that affect the whole team. Conversely, choosing to be a team member assesses traits such as extraversion, agreeableness, and openness, focusing on the player's ability to cooperate and respond to leadership. As the scenario progresses, the player's interactions vary based on their chosen role. Team members face a critical moment when the team leader suggests a poor decision. Here, the player can either challenge the idea, highlighting their conscientiousness and willingness to lead, or remain silent to avoid conflict, which can indicate a preference for maintaining group harmony. This decision not only affects the team's dynamics but also offers insights into the player's personal conflict management style.

2.3. Scenario Three

Scenario three introduces a more reflective component in the simulation, serving as an evaluative pause where the player can process and express their reactions to the events of the previous scenarios. This scenario is modeled after a common break room or watercooler area found in workplace environments, where colleagues gather for informal interactions and downtime. This setting is intentionally chosen to foster a casual and relaxed atmosphere, encouraging the player to engage more openly with the NPCs. In this environment, NPCs are positioned around the room. This allows the player to choose whom to interact with. The player will be asked about how he feels, have options in whether to keep talking and getting to know the player, confront past decisions, etc. As the player navigates the break room, they are given the option to initiate conversations with various NPCs. These interactions are crucial as they allow the player to verbalize their feelings and thoughts about the events that have transpired. The dialogue choices presented to the player are designed to reflect a spectrum of emotional responses, from expressing satisfaction or frustration to showing remorse or defiance. This variety ensures that players can accurately convey their emotional state, which is integral to assessing their emotional intelligence and adaptability.

2.4. Electroencephalography

EEG has been shown to be an excellent tool for real-time feedback of a user's cognitive and affective state [17]. Using EEG, in conjunction with VPS, will allow us to assess the user's decision they made during all phases of the game. Based upon the EEG measurements, we hope to correlate the choices with the user's real personality traits. Our system incorporates the five EEG bands, Alpha, Beta, Delta, Theta, and Gamma, which we obtain by performing a Fast-Fourier Transform (FFT) on the raw EEG data after proper signal cleaning (see Table 1). In addition to these bands, calculated measurements are used to measure the users' levels of engagement, arousal, and valence [18]. A number of studies have successfully utilized these attributes to assess personality

through EEG [11, 12, 19]. The EEG will serve as a predictor and standard for the personality complex profile assessed by the video game.

Table 1. EEG band waves and calculated measurements to assess personality.

Metrics	EEG Measurement
Delta	1 – 4 Hz Dominant wave during deep sleep
Theta	4 – 7 Hz Dominant wave during light sleep
Alpha	7 - 13 Hz Dominant wave when in a state of relaxation
Beta	13 – 25 Hz Dominant wave when active, busy, or concentrating
Gamma	25 – 43 Hz Dominant wave when forming ideas or memory processing
Engagement	<i>Beta(Alpha + Theta)</i> [20]
Arousal	<i>(BetaF3 + BetaF4)(AlphaF3 + AlphaF4)</i> [21]
Valence	<i>(AlphaF4/BetaF4) – (AlphaF3/BetaF3)</i> [21]

3. Conclusion

This study presents a new approach to personality assessment by integrating video game environments with electroencephalography (EEG) technology. Our methodology leverages the immersive and interactive capabilities of video games to probe deeper into the personality traits of individuals in a manner that traditional assessments might fail to capture. By embedding psychological evaluation within a video game, we aimed to create a setting where participants' true behaviors could surface more naturally, free from the self-censorship often associated with survey assessments. The use of EEG in this context provides a direct measurement of neurophysiological responses that underlie the players' decisions and interactions within the game. This approach allows us to correlate specific brain activity patterns with different behavioral responses, offering a window into the complex interplay between neural processes and manifest personality traits. Our game design was intentionally crafted to avoid the appearance of a conventional personality test. This choice was predicated on the hypothesis that an engaging and seemingly non-evaluative environment would elicit more genuine behaviors, thereby providing a more accurate reflection of the player's personality. The integration of gaming with real-time neurological monitoring opens new avenues for research into psychological assessment, cognitive neuroscience, and interactive technology.

Future work will focus on analyzing the collected data to validate and refine our assessment model, with a view to expanding the applicability of our methods across diverse settings and populations.

References

- [1] Krishnamurthy R, Hass GA, Natoli AP, Smith BL, Arbisi PA, Gottfried ED. Professional practice guidelines for personality assessment. *J Pers Assess.* 2022;104(1):1–6. doi: <https://doi.org/10.1080/00223891.2021.1958834>.
- [2] Meyer J, Jansen T, Hübner N, Lüdtke O. Disentangling the association between the Big Five personality traits and student achievement: meta-analytic evidence on the role of domain specificity and achievement measures. *Educ Psychol Rev.* 2023;35(1):12. doi: <https://doi.org/10.1007/s10648-022-09682-8>.
- [3] King SP, Mason BA. Myers-Briggs type indicator. In: *The Wiley Encyclopedia of Personality and Individual Differences: Measurement and Assessment*. Wiley; 2020. p. 315–9.
- [4] Kedar SV, Bormane DS. Automatic personality assessment: a systematic review. In: 2015 International Conference on Information Processing (ICIP). IEEE; 2015. p. 326–331. doi: <https://doi.org/10.1109/ICIP.2015.7350791>.
- [5] Sherman RA, Hogan R. Personality assessment. In: *Talent Assessment: Embracing Innovation and Mitigating Risk in the Digital Age*. Routledge; 2023. p. 84.
- [6] Piedmont RL. Social desirability bias. *Encyclopedia of Quality of Life and Well-Being Research*. Springer; 2014. p. 6036–7. doi: https://doi.org/10.1007/978-94-007-0753-5_2730.
- [7] Worth N. Players and avatars: the connections between player personality, avatar personality, and behavior in video games. 2015.
- [8] Ihsan Z, Furnham A. The new technologies in personality assessment: a review. *Consult Psychol J Pract Res.* 2018;70(2):147. doi: <https://doi.org/10.1037/cpb0000116>.
- [9] Malik AS, Osman DA, Pauzi AA, Khairuddin RNHR. Investigating brain activation with respect to playing video games on large screens. In: 2012 4th International Conference on Intelligent and Advanced Systems (ICIAS2012). IEEE; 2012. p. 86–90. doi: <https://doi.org/10.1109/ICIAS.2012.6306218>.
- [10] Cabanero-Gómez L, Hervas R, Bravo J, Rodríguez-Benítez L. Computational EEG analysis techniques when playing video games: a systematic review. In: *Proceedings*. vol. 2. MDPI; 2018. p. 483. doi: <https://doi.org/10.3390/proceedings2110483>.

- [11] Stenberg G. Personality and the EEG: arousal and emotional arousability. *Pers Individ Dif.* 1992;13(10):1097–113. doi: [https://doi.org/10.1016/0191-8869\(92\)90028-8](https://doi.org/10.1016/0191-8869(92)90028-8).
- [12] Zhao G, Ge Y, Shen B, Wei X, Wang H. Emotion analysis for personality inference from EEG signals. *IEEE Trans Affect Comput.* 2017;9(3):362–71. doi: <https://doi.org/10.1109/TAFFC.2017.2655184>.
- [13] Kuper N, Käckenmester W, Wacker J. Resting frontal EEG asymmetry and personality traits: a meta-analysis. *Eur J Pers.* 2019;33(2):154–75. doi: <https://doi.org/10.1002/per.2184>.
- [14] Butler C. A framework for evaluating the effectiveness of gamification techniques by personality type. In: Nah FFH, editor. *HCI in Business*. Springer; 2014. p. 381–9. doi: https://doi.org/10.1007/978-3-319-07293-7_36.
- [15] Myers IB. The Myers-Briggs Type Indicator: Manual (1962). 1962.
- [16] Gosling SD, Rentfrow PJ, Swann WB. A very brief measure of the Big-Five personality domains. *J Res Pers.* 2003;37(6):504–28. doi: [https://doi.org/10.1016/S0092-6566\(03\)00046-1](https://doi.org/10.1016/S0092-6566(03)00046-1).
- [17] McMahan T, Parberry I, Parsons TD. Evaluating player task engagement and arousal using electroencephalography. *Procedia Manuf.* 2015;3:2303–10. doi: <https://doi.org/10.1016/j.promfg.2015.07.377>.
- [18] Parsons TD, McMahan T, Parberry I. Classification of video game player experience using consumer-grade electroencephalography. *IEEE Trans Affect Comput.* 2020;13(1):3–15. doi: <https://doi.org/10.1109/TAFFC.2017.2682119>.
- [19] Klados MA, Konstantinidi P, Dacosta-Aguayo R, Kostaridou VD, Vinciarelli A, Zervakis M. Automatic recognition of personality profiles using EEG functional connectivity during emotional processing. *Brain Sci.* 2020;10(5):278. doi: <https://doi.org/10.3390/brainsci10050278>.
- [20] Freeman FG, Mikulka PJ, Prinzel LJ, Scerbo MW. Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biol Psychol.* 1999;50(1):61–76. doi: [https://doi.org/10.1016/S0301-0511\(99\)00002-2](https://doi.org/10.1016/S0301-0511(99)00002-2).
- [21] Giraldo S, Ramirez R. Brain-activity-driven real-time music emotive control. In: The 3rd International Conference on Music & Emotion; 2013; Jyväskylä, Finland. Jyväskylä: University of Jyväskylä, Department of Music; 2013.

Democratizing Psychophysiological Research: Leveraging Open Resources for Experimental Settings and Data Analysis of Flow, Engagement and Performance

Elena SAJNO^{a,b,c,1}, Samantha CHAN^c, Federico DAGOSTIN^d, Stefano DE GASPARI^{a,b}, Chiara PUPILLO^{a,b}, and Giuseppe RIVA^{b,e,2}

^a*University of Pisa, Pisa, Italy*

^b*HTLab, Università Cattolica del Sacro Cuore, Milan, Italy*

^c*MIT MediaLab, Massachusetts Institute of Technology, Cambridge, USA*

^d*BMI Labs "Mario Stefanelli", University of Pavia, Italy*

^e*ATNPLab, IRCCS Istituto Auxologico Italiano, Milan, Italy*

ORCID ID: Elena Sajno <https://orcid.org/0000-0002-9621-8981>

Abstract. Contemporary research faces challenges due to economic constraints and a lack of specialized software. However, advancements such as affordable wearable sensors and collaborative platforms like GitHub are revolutionizing this landscape. This work proposes the development of an experimental setup that integrates game stimuli presentation and synchronized recording of facial movements and cardiac data, facilitating resource sharing and collaboration. The objective of this framework is to elicit and detect states of flow and engagement along with associated physiological activations. Using different difficulty levels of a Tetris-style game as stimuli, the framework synchronizes players' activity with Interbeat interval (IBI) data from the Polar H10 cardiac chest band and webcam videos. Key to the approach is precise data synchronization using timestamps and established libraries like LabStreamingLayer (LSL). Three Python scripts manage Tetris presentation, Polar H10 Bluetooth connection for cardiac data collection, and video frame tagging for facial expression analysis. These scripts adhere to FAIR principles and provide a practical implementation for studying psychophysiological responses. By leveraging open resources, this framework addresses economic and software limitations and allows for adaptation to various research questions or theoretical constructs.

Keywords. Flow, video game, open source, ECG, emotions, FAIR

1. Introduction

Contemporary research often grapples with economic constraints and a shortage of specialized and updated software tailored to specific research needs, or cheap and reliable hardware. The landscape is, however, evolving: the emergence of affordable wearable sensors, increasingly often considered the gold standard for research due to their reliability and the high sampling frequency is a resource for the researcher in psychology. Meanwhile, platforms such as GitHub, Hugging Face, PhysioNet, and Kaggle facilitate the democratization of science by enabling the sharing of scripts and datasets, thereby fostering collaboration and accelerating progress. Yet, assembling resources for studies, particularly those integrating psychophysiology, remains complex, especially for the crucial need for synchronized data. Within this context, this work proposes the development of an experimental setup stemming from shared resources, that allows the presentation of game stimuli and the synchronized recording of facial movements and cardiac data, with the aim of creating a new sharable instrument. The full code is available on GitHub at the address: <https://github.com/elenasajno/TetrisInterbeats>.

Objectives: The experimental goal for which this framework is designed is to elicit and detect states of flow and engagement and the physiological activations associated with them, as described in [1,2,3]. As stimuli, several levels of different difficulty of a

¹ Corresponding Author: Elena Sajno elena.sajno@phd.unipi.it

² Research partially supported by the Italian Ministry of Health

Tetris-style video game are employed; the player's activity needs to be saved and synchronized with data coming from the cardiac wearable device Polar H10 and videos received from the webcam. The final framework can also be adapted to other research questions or theoretical constructs.

Why using videogames: Video games are seen as effective tools for evoking various specific emotions [4], which can be measured through psychophysiological signals [5,6]. Additionally, different difficulty levels within the same video game, such as Tetris, have been found to provoke distinct mental states: for instance, varying levels of difficulty in tasks can evoke feelings of boredom, anxiety, and engagement [4]. Video games that offer various levels, similar in structure but different in difficulty, enable a more tailored match to the abilities of each participant [4,7]. Specifically, Tetris was used for studies inherent in the states of flow [8,9] and engagement [4,11].

Cardiac Psychophysiology: Psychophysiological signals, such as the cardiac one, offer insight into the sympathetic and parasympathetic activities, emotions, and mental activation [12,13]. Moreover, the use of wearable devices to keep track of ECG/BVP signals allows a comfortable for the participants' data collection, and research which can be set also outside the laboratories. Specific features calculated from heart beating are also particularly informative: Heart Rate Variability (HRV), for example, can be decomposed into frequencies, each one more related to distinct activation of the autonomic nervous system or mental/affective state [15]. HRV values are usually computed by selecting the R peaks in ECG, and calculating the time between an R peak and the following [15]; to reach a "gold standard" HRV an at least 250Hz ECG is required [14]. Another effective way to derive HRV from Interbeat Intervals (IBIs), that already measure the millisecond from a beat and the following [15]: in this work, we propose to extract directly the IBIs, calculated by the Polar H10 band in milliseconds (so at 1000Hz), which is reported as a reliable instrument for the R-R detection [22]. Since the recording of a baseline is necessary for good experimental design involving HRV [16], we have included a specific function for selecting the chosen duration for a baseline acquisition, timestamping its start and end moments.

Facial expressions & Emotions: Since Ekman's studies [17] facial expressions, through microexpressions, are one of the more direct ways to understand which kind of emotion is perceived by a person. Attention [18] and mental workload [19] can also be inferred through eye movement parameters. Software such as Noldus' FaceReader, GazeRecorder, or other AI video applications can detect from videos of a face, the microexpressions, and ocular movements linked to emotions, affect, and level of attention. Recording the participant's facial activity can be easily achieved with a laptop camera: we propose a code for this function, putting in overlay the timestamp of each frame.

Need for Synchronization: The synchronization of signals from different sensors is a crucial step in signal pre-processing, particularly in the design of multimodal affective computing experiments [20]. Typically, synchronization is achieved by collecting markers when stimuli are presented [21]. This process allows for precise alignment of data streams, enabling the investigation of the temporal relationships between stimuli presentation and physiological responses. The objective of this work is to provide scripts to create a complete experimental setting, in which each section can be synchronized with the others: these Python 3 codes are meant to be run from the terminal on the same machine to ensure local timing consistency. Key moments of stimulus presentation will be recorded, video acquisition will be timestamped for each frame, and each IBIs will be recorded with its corresponding timestamps. We have decided to use the LSL timestamp, which calculates the timestamp from the machine's boot time down to the microsecond, the machine's proper tick, and the Unix Time Stamp, which records down to the microsecond referring to the reference timestamp (from which the precise date and time of the recording can be extracted). This approach ensures that it is possible to trace back the various moments of the data collection and that they can be synchronized to the millisecond.

2. Methods

Key to our approach is the punctual synchronization of data streams, through precise timestamps and the utilization of established libraries, such as LSL. By recording and storing essential metadata alongside physiological and behavioral data, our framework enables the reconstruction of temporal sequences, crucial for in-depth analysis and interpretation. Drawing from shared codes on GitHub, our framework is divided into three Python 3 scripts, all intended to run on the same machine, ensuring consistency in local timestamp management. 1) The Tetris component allows the presentation of the stimuli and the collection of gaming performances and timestamps for various in-game events; it is also possible to insert a countdown of desired length to record an initial physiological baseline. 2) The PolarH10 module enables the Bluetooth connection with the chestband, plus the detection and saving of cardiac IBIs with a 1000 ms precision for a reliable HRV calculations [22]. Additionally, the following synchronization with gameactivities and facial expressions enables the assessment of both the resting baseline state and the reactions to specific stimuli/events. 3) The video component ensures that each frame is tagged with multiple timestamps, facilitating the synchronization to perform analysis through facial expression recognition softwares. All three scripts allow for including the subject number (*subject number* = `input("Enter the subject number: ")`), task number (*task number* = `input("Enter the task number: ")`), and date of data collection (*date str* = `input("Enter the date (YYYY-MM-DD): ")`) calling before the main function. Those data are used to save a file named after them (shaped like *filename* = `f'tetris data {date str} subject{subject number} task{task number}.csv`). In the same function are defined the rules for writing the .csv file, defining multiple types of timestamps (LSL, tick, and Unix's: `timestamp = datetime.datetime.now().timestamp(); tick = lib.GetTickCount64(); local time = pylsl.local clock()`), that are reported for each saved data.

Tetris Component: The Tetris framework was started forking tetris.py from <https://github.com/Datamine/Tetris>, a custom Tetris-style game based on pygame. A countdown function was defined, to allow a custom-length baseline recording, that can be selected before the main function (*t* = `input("Enter baseline length (in seconds): ")`):

```
def countdown(t):
    print('start_baseline')
    write_to_csv("start_baseline")
    while t:
        mins, secs = divmod(t, 60)
        timer = '{:02d}:{:02d}'.format(mins, secs)
        print(timer, end="\r")
        time.sleep(1)
        t -= 1
    print('end_baseline')
    write_to_csv("end_baseline")
```

The commands *print(text)* and *write to csv(text)* were added in multiple positions in the script, in the functions defining the gameover, the pause, the resume, the proper gaming phases (the start of the new level and which level was selected), when a complete line is achieved (and if it is single one or more), the exit to menu, and the actual start of the application. For example, the first part of the function defining the game is here reported:

```
def game(screen,startinglevel):
    print("inizio livello")
    write_to_csv("inizio livello")
    cleared = 0
    tetrimino = newtetrimino()
    bestscore = int(getmaxlines())
    board = [[''] * 20 for n in range(10)]
    background = pygame.image.load("Grid.PNG")
    backgroundcolor = getrgb(gridline)
```

```

timestep = time.time()
bottom = pygame.font.Font('BebasNeue.ttf',20)
white = getrgb("#FFFFFF")
print("starting level", startinglevel)
write_to_csv("starting level:"+ str(startinglevel))

```

Polar H10 IBI Component: The project started from <https://github.com/markspan/PolarBand2lsl/blob/main/Polar2LSL.py>. It allows connecting a PolarH10 chest band to the machine running the Python code, receiving the ECG signal (130Hz), and streaming it in a way compatible with softwares able to detect LSL streams. The device connection is established using its Bluetooth Low Energy (BLE) framework through the Bleak library, which utilizes Predefined UUIDs (Universal Unique Identifiers) - based on the Heart Rate GATT (Generic Attribute Profile) service protocol - to access its services. This allows for obtaining a targeted response input from the device, essentially serving as an API. The code establishes a connection to the device, retrieves device information (such as model number, manufacturer name, and battery level), and starts the data stream. Here an extract is reported:

```

import asyncio
from bleak import BleakClient
uuid16_dict = {v: k for k, v in uuid16_dict.items()}
# UUIDs for device information
MODEL_NBR_UUID = "0000{0:x}-0000-1000-8000-00805f9b34fb"
    .format(uuid16_dict.get("Model Number"))
MANUFACTURER_NAME_UUID =
    "0000{0:x}-0000-1000-8000-00805f9b34fb".format(uuid16_
dict.get("Manufacturer Name"))
BATTERY_LEVEL_UUID = "0000{0:x}-0000-1000-8000-00805f9b34fb"
    .format(uuid16_dict.get("Battery Level"))
# UUIDs for connection and data streaming
PMD_SERVICE = "FB005C80-02E7-F387-1CAD-8ACD2D8DF0C8"
PMD_CONTROL = "FB005C81-02E7-F387-1CAD-8ACD2D8DF0C8"
PMD_DATA = "FB005C82-02E7-F387-1CAD-8ACD2D8DF0C8"
PMD_CHAR1_UUID = "fb005c81-02e7-f387-1cad-8acd2d8df0c8"
PMD_CHAR2_UUID = "fb005c82-02e7-f387-1cad-8acd2d8df0c8"
U1_CHAR2_UUID = "6217ff4d-91bb-91d0-7e2a-7cd3bda8a1f3"

```

Due the need to obtain data with a higher sampling rate, to follow HRV guidelines [14], we decided to retrieve the IBIs signal. To achieve this, we modified the script with information coming from <https://github.com/kbre93/dont-hold-your-breath/blob/master/PolarH10.py>. The UUID for the RR interval is UUID 00002a370000-1000-8000-00805f9b34fb [23] and, the raw signal retrieved from the Polar H10 needs to be converted from byte to float [23]. The various UUIDS information was added to our script, alongside the conversion parameters:

```

IBI_WRITE = bytearray([0x02, 0x00, 0x00, 0x01, 0x37, 0x00, 0x01,
    0x01, 0x0E, 0x00])
ibi_SAMPLING_FREQ = 1000
HEART_RATE_MEASUREMENT_UUID = "00002a37-0000-1000-8000-
00805f9b34fb"
byte0 = data[0] # heart rate format
uint8_format = (byte0 & 1) == 0
energy_expenditure = ((byte0 >> 3) & 1) == 1
rr_interval = ((byte0 >> 4) & 1) == 1
if not rr_interval: return
first_rr_byte = 2
if uint8_format:
    hr = data[1]
    pass
else:

```

```

        hr = (data[2] << 8) | data[1] # uint16
        first_rr_byte += 1
    if energy_expenditure:
        ee = (data[first_rr_byte + 1] << 8) | data[first_rr_byte]
        first_rr_byte += 2
    for i in range(first_rr_byte, len(data), 2):
        ibi = (data[i + 1] << 8) | data[i]
        ibi = np.ceil(ibi / 1024 * 1000)
        ibi_stream_values.extend([ibi])
        ibi_stream_times.extend([time.time_ns() / 1.0e9])

```

In the same function in which the IBIs are calculated, the corresponding timestamps are retrieved, and each IBI and its related timestamps (*through local time = pylsl.local clock() e current datetime = datetime.datetime.now()*) are written to a row of a CSV file (nominated as already described).

Video Acquisition Component: The Video acquisition module started forking <https://github.com/maksimKorzh/open-cv-tutorials/blob/main/src/record-webcam/webcam.py> based on the cv2 library, which permits to display and save videos recorded through the webcam. To ensure that the video can be aligned with the rest of the data streams, we chose to record the video by superimposing on each frame the timestamp at which they were recorded. To achieve this, the libraries and functions for displaying the various timestamps were added, as in the previous modules, and a script that superimposes the various timestamps (*timestamp = time.time(), current time = datetime.datetime.now(), tick = lib.GetTickCount64(), local time = pylsl.local clock(), tick = int(str(tick))*) in a precise position, each frame, before saving the video was created.

```

superimposes
cv2.putText(frame, timestamp_str+" "+tick_str,
            (10, 30), cv2.FONT_HERSHEY_SIMPLEX, 0.8, (0, 0, 255), 2)
cv2.putText(frame, formatted_time_with_ms+" "+local_time_str, (10, 50),
            cv2.FONT_HERSHEY_SIMPLEX, 0.8, (0, 0, 255), 2)
video.write(frame)

```

3. Conclusion

The software here presented provides a practical implementation of shared resources for the elicitation of flow and engagement. It is coupled with the synchronized collection of cardiac data (IBIs) and facial movements for the study of psychophysiological responses by means of a PolarH10 band and a webcam. Feelings of boredom, anxiety and engagement were provoked by tweaking the difficulty settings within the game of Tetris. The code is divided into three separate scripts: the game component, the PolarH10 module and the video recording component, in part adapted from other open source software projects available on GitHub. According to FAIR [10] (Findable, Accessible, Interoperable, Reusable) principles, the scripts have been bundled into a GitHub repository (<https://github.com/elenasajno/TetrisInterbeats>) and are available along with all subsequent versions and evolutions of the toolset, under GPLv3 license terms to all parties, to replicate results or adapt the developed framework to their own research needs. This degree of availability addresses some issues pertaining to replicability, in that it removes the variable of tools used to collect the data, allowing for a more direct comparison of the harvested values. In conclusion, leveraging open resources for experimental setups and data analysis not only addresses resource and software limitations but also promotes collaboration and accelerates scientific progress in psychophysiological research.

Acknowledgment

Special thanks to Matteo “Dalfó” Dalfonso.

References

- [1] Sajno E, Riva G. Follow the Flow: Artificial Intelligence and Machine Learning for Achieving Optimal Performance. *Cyberpsychol Behav Soc Netw.* 2022;25(7):476-7. doi: <https://doi.org/10.1089/cyber.2022.29251.ceu>.
- [2] Sajno E, Beretta A, Novielli N, Riva G. Follow the Flow: A Perspective on the On-Line Detection of Flow Mental State through Machine Learning. In: Proceedings of the 2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE); 2022 Oct 26. IEEE; p. 217-222. doi: <https://doi.org/10.1109/MetroXRAINE54828.2022.9967605>.
- [3] Sajno E. Affective Computing for Detecting Psychological Flow State: A Definition and Methodological Problem. In: Proceedings of the 2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW); 2023 Sep 10. IEEE; p. 1-5. doi: <https://doi.org/10.1109/ACIIW59127.2023.10388109>.
- [4] Chanel G, Rebetez C, Betrancourt M, Pun T. Emotion assessment from physiological signals for adaptation of game difficulty. *IEEE Trans Syst Man Cybern A Syst Hum.* 2011;41(6):1052-63
- [5] Kivikangas JM, et al. A review of the use of psychophysiological methods in game research. *J Gaming Virtual Worlds.* 2011;3(3):181-199.
- [6] Granato M, Gadia D, Maggiorini D, Ripamonti LA. An empirical study of players' emotions in VR racing games based on a dataset of physiological data. *Multimedia Tools Appl.* 2020;79(45):33657-33686.
- [7] Landowska A, Wrobel MR. Affective reactions to playing digital games. In: Proceedings of the 8th International Conference on Human System Interaction (HSI); 2015 Jun. pp. 264-270.
- [8] Keller J, Bless H. Flow and Regulatory Compatibility: An Experimental Approach to the Flow Model of Intrinsic Motivation. *Pers Soc Psychol Bull.* 2008;34(2):196-209. doi: <https://doi.org/10.1177/0146167207310026>.
- [9] Harmat L, de Manzano O, Theorell T, Högman L, Fischer H, Ullén F. Physiological correlates of the 'flow' experience during computer game playing. *Int J Psychophysiol.* 2015;97(1):1-7. doi: <https://doi.org/10.1016/j.ijpsycho.2015.05.001>.
- [10] Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci data.* 2016;3(1):1-9, doi: <https://doi.org/10.1038/sdata.2016.18>
- [11] Chanel G, Rebetez C, Betrancourt M, Pun T. Boredom, engagement and anxiety as indicators for adaptation to difficulty in games. In: Proceedings of the 12th International Conference on Entertainment and Media in the Ubiquitous Era. MindTrek '08. Tampere, Finland: Association for Computing Machinery; 2008. p. 13-17. doi: <https://doi.org/10.1145/1457199.1457203>.
- [12] Sajno E, Bartolotta S, Tuena C, Cipresso P, Pedroli E, Riva G. Machine learning in biosignals processing for mental health: A narrative review. *Frontiers in Psychology.* 2022;13:1066317. doi: <https://doi.org/10.3389/fpsyg.2022.1066317>.
- [13] Cacioppo JT, Berntson GG, Larsen JT, Poehlmann KM, Ito TA. The psychophysiology of emotion. In: Lewis M, Haviland JM, editors. *Handbook of Emotions.* New York, NY: Guilford Press; 2000.
- [14] Malik M, Bigger JT, Camm AJ, Kleiger RE, Malliani A, Moss AJ, Schwartz PJ. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Eur Heart J.* 1996;17(3):354-381. doi: <https://doi.org/10.1093/oxfordjournals.eurheartj.a01486>.
- [15] Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. *Frontiers in Public Health.* 2017;5:258. doi: <https://doi.org/10.3389/fpubh.2017.00258>.
- [16] Laborde S, Mosley E, Thayer JF. Heart Rate Variability and Cardiac Vagal Tone in Psychophysiological Research – Recommendations for Experiment Planning, Data Analysis, and Data Reporting. *Frontiers in Psychology.* 2017;8:213. doi: <https://doi.org/10.3389/fpsyg.2017.00213>.
- [17] Ekman P. An argument for basic emotions. *Cogn Emot.* 1992;6(3-4):169-200. doi: <https://doi.org/10.1080/0269939208411068>.
- [18] Hoang Duc AK, Bays P, Husain M. Eye movements as a probe of attention. *Prog Brain Res.* 2008;171:403-411. doi: [https://doi.org/10.1016/S0079-6123\(08\)00659-6](https://doi.org/10.1016/S0079-6123(08)00659-6).
- [19] de Greef T, Laféber H, van Oostendorp H, Lindenberg J. Eye Movement as Indicators of Mental Workload to Trigger Adaptive Automation. In: Schmorow DD, Estabrooke IV, Grootjen M, editors. *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience.* FAC 2009. Lecture Notes in Computer Science, vol 5638. Berlin, Heidelberg: Springer; 2009. doi: https://doi.org/10.1007/9783-642-02812-0_26.
- [20] Bota PJ, Wang C, Fred ALN, Da Silva HP. A review, current challenges, and future possibilities on emotion recognition using machine learning and physiological signals. *IEEE Access.* 2019;7:140990-141020. doi: <https://doi.org/10.1109/ACCESS.2019.2944001>.
- [21] Koelstra S, Muhl C, Soleymani M, Lee J-S, Yazdani A, Ebrahimi T, et al. DEAP: a database for emotion analysis using physiological signals. *IEEE Transactions on Affective Computing.* 2012;3:18-31. doi: <https://doi.org/10.1109/T-AFFC.2011.15>.
- [22] Gilgen-Ammann R, Schweizer T, Wyss T. RR interval signal quality of a heart rate monitor and an ECG holter at rest and during exercise. *Eur J Appl Physiol.* 2019;119(7):1525-1532. doi: <https://doi.org/10.1007/s00421-019-04117-1>.
- [23] Pramukantoro, E.S. A Study of Continuous Heartbeat Monitoring System Using The Internet of Things Approach. PhD dissertation. Okayama University; September 2022.

Exploring Ethical Considerations of Serious Game: a Qualitative Secondary Analysis on Improving COVID-19 Vaccination Among Youth

Lihong OU ^{a,3} and Thomas D. PARSONS ^{a, b}

^a*Computational Neuropsychology & Simulation Laboratory, Arizona State University*

^b*Edson College, Arizona State University*

ORCID ID: Lihong Ou <https://orcid.org/0000-0002-4104-7935>

Abstract. Serious game-based interventions are increasingly valuable for enhancing public health by boosting vaccination knowledge and shaping young individuals' attitudes and behaviors. These tools significantly impact how young people perceive and respond to vaccinations, but it is crucial to consider their ethical implications, particularly on the neurocognitive, affective, and social aspects of youth engaging with educational technologies. This study continued a prior randomized controlled trial assessing a game-based intervention for COVID-19 vaccine education among unvaccinated 11–14-year-olds. The study retained all 32 parent-youth pairs, equally randomized into intervention and control groups. Surveys were administered before, immediately after, and two months post-intervention, with post-intervention semi-structured interviews conducted with 7 intervention dyads. The previous study showed significant improvements in vaccine knowledge and beliefs among participants. Two months later, 75% of the intervention group reported vaccinating their children compared to 37.5% in the control group. Despite the positive impact, ethical concerns were identified. Neurocognitive concerns include the game's simplicity, potentially hindering the acquisition of critical skills such as critical thinking and problem-solving. Affective concerns highlight the lack of emotional depth, potentially resulting in disengagement and inadequate emotional development. Social concerns arise from the single-player format, which may inadequately promote collaboration, empathy, and appreciation of diverse perspectives. Serious game-based interventions show promise in improving youth vaccine knowledge and uptake, highlighting the benefits of educational technologies in public health. However, their insufficient design may hinder neurocognitive, emotional, and social development, necessitating a more nuanced, ethically responsible approach to enhance their effectiveness.

Keywords. Serious game-based learning, youth vaccine education, ethical

1. Introduction

Serious games have evolved beyond entertainment, aiding users in various domains since the 1970s [1]. They support public health initiatives, from promoting healthy lifestyles to addressing specific health challenges [2-6]. Designed to increase vaccination knowledge, awareness, and influence the attitudes and behaviors of young individuals [7,8], serious games demystified misconceptions, encourage informed decision-making, and potentially increase vaccination rates. This paper explores a 2019 coronavirus disease (COVID-19) game-based intervention's impact on youths' cognitive and behavioral responses to vaccination education, emphasizing the ethical considerations in their design and implementation.

³ Corresponding Author: Lihong Ou, lihongou@asu.edu.

2. Methods

This research reanalyzed interview data from a prior study to explore new themes, ensuring transparency and rigor following Ruggiano and Perry [9]. The current study used data from a prior randomized controlled trial on a serious game-based COVID-19 intervention's feasibility, acceptability, and efficacy. Participants were US parent-youth dyads with unvaccinated youths aged 11-14 who spoke and read English. Recruitment used flyers and videos via newsletters, organizations, marketplaces, libraries, and social media. Participants completed a QR code screening survey, followed by verification and snowball sampling. Eligible dyads received the intervention and were followed up after two months. The parent study included online surveys at baseline, post-intervention, and a 2-month follow-up, plus semi-structured interviews post-intervention. This secondary analysis used only the interview data, conducted via Zoom from April 25th to July 25th, 2023, as the surveys didn't align with the current research questions.

The dataset comprised transcripts from 7 interviews, focusing on experiences with a game-based COVID-19 intervention among parent-youth dyads with unvaccinated youths aged 11-14. Statements were coded using interpretive description [10] to inform serious game use in vaccination education. An interpretive description qualitative design was chosen to understand gaming experiences, guided by Parsons' brain-based cyberpsychology framework [11,12]. This approach integrates neuroscience and cyberpsychology to explore cognitive, affective, and social processes in human-technology interactions, providing a foundational structure for the study. Data were analyzed using constant comparative analysis [13], with emergent themes identified and discrepancies resolved through consensus discussions.

3. Results

Seven parent-youth dyads participated in the semi-structured interview immediately after playing the game. These dyads comprised three mother-son dyads, three father-son dyads, and one mother-daughter dyad. Among the parents, five identified as Black or African American, while two identified as non-Hispanic White. In terms of educational attainment, four parents had achieved at least a bachelor's degree, one had some high school education, one had completed high school or obtained a GED, and one had a graduate degree, respectively. Regarding employment status, six parents were working for pay. Participants shared insights into learning, gaming, and their interactions, engagement, and responses to the game elements.

Within four overarching themes, a total of nine subthemes were identified. For the neurocognitive theme, subthemes included: (1) previous experience with educational gaming, cognitive engagement through educational gaming, cognitive challenge and skill development; (2) in the affective domain, emotional connection and motivation, aesthetic and narrative impact on enjoyment; (3) the social theme encompassed social learning and behavioral influence; (4) lastly, the user experience and interface design theme included user interface adaptability, accessibility and engagement, and user-friendly design.

3.1. Cognitive Engagement through Educational Gaming

None of the participants had previous experience with educational gaming, with only one male youth mentioning that he had played alphabet games at school. The educational game was found to be an effective tool for delivering age-appropriate content about the COVID-19 vaccine, engaging players, and enhancing their understanding of vaccination requirements, vaccine eligibility, including the necessity of receiving two doses. The game also communicated the importance and safety of the COVID-19 vaccine, prompting players to reevaluate their beliefs and attitudes. Its design, by offering direct feedback in response to player choices, facilitated an understanding of the consequences of those choices within the game environment.

This active involvement, coupled with the strategic navigation of questions and persuasion of in-game characters to receive the vaccine, not only made the experience immersive but also enabled a participant who initially found the game confusing to

become actively engaged. Even for those who perceived the information as a reminder, the game served as a platform for applying existing knowledge in new and meaningful ways. However, one participant noted effective strategies for navigating vaccine-related queries, suggesting that while the game challenged players to think critically, some found ways to streamline this process. They commented that the information provided in the game failed to fully engage them or meet their expectations for specificity and depth, which would facilitate learning.

3.2 Cognitive Challenge and Skill Development

The game was reported to help players gain a better understanding of vaccination details, such as the necessity of completing the vaccine series and information on pain management post-vaccination. The participants also appreciated the implementation of a five-minute time limit for completing the game, as it introduced an additional layer of challenge and excitement.

There was a notable critique from participants regarding the repetitiveness of questions. This repetition seemed to not add substantial value to the learning experience, as the predictability of answers and lack of novel learning opportunities did not adequately engage users in the problem-solving process. Participants also expressed the need for the game to provide clear, attainable goals and to present information on vaccine efficacy in a manner that was cognitively engaging without being overwhelming. Additionally, the game's content delivery needed to accommodate users' varying levels of patience and reading ability were reported.

3.3 Affective Connection and Motivation

The participants found that convincing characters to be vaccinated was helpful, engaging them emotionally and motivating further learning. The incorporation of interactive elements also encouraged them to take action to administer vaccination. They showed a desire to improve as they retried the game for correct answers. The dyads also suggested that providing details about the disease could potentially influence behavior by ensuring players are aware of the risks associated with COVID-19 and the benefits of vaccination. Moreover, including background music that was either thematic or contained elements of storytelling that children find appealing was shared by the participants as something that could enhance the overall experience.

3.4 Aesthetic and Narrative Impact on Enjoyment

Participants reported that the game's vibrant animations and character designs enhanced their experience, viewing the game as an effective medium for conveying messages about COVID-19. Additionally, they highlighted the value of interactive features, such as moving in-game avatars to isolation centers and engaging in dialogue, which further bolstered the game's appeal. They expressed a preference for scenarios like a police officer transporting an isolated person to an ambulance, valuing the visual representation of such relevant events. The inclusion of personalized elements like character names further contributed to the game's appeal, with its colorful and engaging content capturing and retaining player attention effectively.

3.5 Social Learning and Behavioral Influence

The study findings suggested that the educational game was effective in enhancing parents' understanding of COVID-19 vaccinations. The participants noted that the game served as a valuable educational resource, particularly for those lacking sufficient information, by providing comprehensive knowledge about vaccination, thereby potentially reducing vaccine-related anxieties. It was also seen as an aid for governmental public health outreach.

Furthermore, the game was praised for its interactive elements that mimic real-life interactions, such as conversations with healthcare professionals, which participants found to be an authentic and engaging feature of the gameplay. Nonetheless, they

suggested a feature where vaccinated avatars could be used to influence other characters, simulating real-life social encouragement for vaccination. They also noted that the game concluded after characters received their second vaccine dose, suggesting a need for further content on ongoing management, such as yearly boosters, and for more interactive elements post-vaccination to enhance realism and reflect current public health practices.

3.6 User Interface Adaptability

The participant feedback on interface adaptability highlighted a need for full-screen visual elements and consistent positioning across devices, to maintain context and orientation within the game environment. Participants suggested enhancements like keyboard integration for character interaction and color-coded text to attract younger players' attention. A preference for a split-screen layout was expressed, to prevent overlay boxes from obstructing gameplay visuals, alongside a call for introductory pop-ups to set the stage for engagement. There was also a desire for improved depth in graphics, indicating the importance of a three-dimensional feel to increase realism and player immersion.

3.7 Accessibility and Engagement

Some dyads addressed the necessity for interactive components that provide deeper learning opportunities. The suggestion to replace static question boxes with characters voicing their concerns was deemed an effective approach to humanize the educational content. Others proposed the use of visuals, like pictures and animations, to replace dense text, thereby streamlining comprehension and engagement, particularly under time constraints. Additionally, the inclusion of direct informational links and local vaccine resources was highlighted as a means to enrich the educational capacity of the intervention, ensuring immediate responsiveness and support within the gameplay, and enhancing the overall interactive experience for youth users.

Feedback on the user-friendly design highlighted the game's positive reception due to its approachability and straightforward format. However, it also identified areas requiring improvement to enhance authenticity and player immersion. Suggestions included enhancing graphical quality to avoid a monotonous aesthetic, optimizing the interface to reduce screen clutter, and incorporating elements like chairs in vaccination scenes for a more realistic depiction. These adjustments could be essential for maintaining player interest and interaction.

4. Discussion

This qualitative secondary analysis explored the ethical considerations of using serious games in health education, with a specific focus on educating youth about COVID-19 vaccinations. The interactive nature of the game facilitated informed decision-making about vaccinations, contributing to a broader understanding and encouraging vaccine uptake among young players. Notably, most participants were new to educational gaming, which suggests an opportunity to engage a previously unreached audience in public health initiatives. The game's capacity to present relevant COVID-19 vaccine information in an engaging and comprehensible manner could be particularly beneficial in addressing vaccine hesitancy. While the game was educational and interactive, concerns arose about its simplicity and repetitiveness. These issues might not challenge players' cognitive skills like critical thinking, potentially hindering neurocognitive development and player autonomy [14]. Enhancing the game with more engaging content, diverse experiences, and complex scenarios tailored to various literacy levels is essential to promote informed decision-making and effective vaccination behaviors.

The character design, animations, and interactivity in the gameplay tailored to engage youth were key for connecting them to content meaningfully. By assigning characters actual names and involving them in vaccination communication and

administration scenarios, the game enhanced relatability and real-world relevance. These personalized features, alongside vivid animations, visually rich content, and fact-rich content, could further contribute to an emotionally rich environment and enhance player engagement, encourage informed decisions, and shape vaccination intentions and behaviors [12]. Additionally, incorporating realistic scenarios with background music and story elements may also magnify immersion, making the educational process both enjoyable and impactful.

The integration of healthcare dialogues in the game-based intervention complemented vaccination outreach efforts and showcased its potential as a reliable source of vaccine knowledge. Incorporating peer and social influence into the future updates could portray community norms and diverse perspectives on vaccination, enhancing both learning about vaccination and its practical application in real-world contexts. However, the game concluding after the second dose reveals a gap in ongoing education and its single-player setup might not sufficiently cultivate collaboration and shared decision making. Expanding the narrative to cover yearly boosters and adaptive vaccine reference guide could provide continuous learning and mirror the evolving nature of vaccination recommendations, thereby maintaining relevance and educational impact over time.

Moreover, ethical design in educational games prioritizes user experience and interface design to empower youth users, fostering an accessible and inclusive environment that accommodates diverse abilities and learning styles. This approach seeks to educate in a constructive and respectful manner, considering the cognitive and emotional well-being of youth. Thoughtful design considerations regarding full-screen visibility, adaptable layouts, voiced dialogues, keyboard inputs, and enriched visual representations mirroring real-world scenarios, coupled with prompt access to supportive materials, could facilitate active learning while mitigating cognitive strain.

In conclusion, educational gaming interventions could improve vaccine knowledge, beliefs, and intentions among youth, showcasing the value of digital tools in health education. While effective, the current intervention may lack in supporting comprehensive neurocognitive, affective, and social development. A carefully designed approach could enhance learning, adhering to ethical principles and developmental needs.

References

- [1] Djaouti D, Alvarez J, Jessel J.-P, Rampnoux O. Origins of serious games. In: M Ma, A Oikonomou, LC Jain, editors. *Serious games and edutainment applications*: Springer; 2011.p.25-43, doi: https://doi.org/10.1007/978-1-4471-2161-9_3
- [2] Andrew L, Barwood D, Boston J, Masek M, Bloomfield L, Devine A. Serious games for health promotion in adolescents – a systematic scoping review. *Educ Inf Technol.* 2023 May; 28(5): 5519–50. doi: <https://doi.org/10.1007/s10639-022-11414-9>
- [3] Haruna H, Zainuddin Z, Okoye K, Mellecker RR, Hu X, Chu SKW, et al. Improving instruction and sexual health literacy with serious games and gamification interventions: an outlook to students' learning outcomes and gender differences. *Interact. Learn. Environ.* 2023; 31(4): 2392–2410. doi: <https://doi.org/10.1080/10494820.2021.1888754>
- [4] Lussier-Drochers D, Massé L, Simonato I, Lachapelle Y, Godin-Tremblay V, Lemieux A. Evaluation of the effect of a serious game on the performance of daily routines by autistic and ADHD children. *Adv. Neurodev. Disord.* 2023; 7(4): 566–578. doi: <https://doi.org/10.1007/s41252-023-00319-4>
- [5] Damasceno EF, Da Silva AP, Barbosa JD. A serious game-based platform for measuring treatment adherence. *Multimed Tools Appl.* 2024 Jan;83(4):12017–33.doi: <https://doi.org/10.1007/s11042-023-15988-5>
- [6] Vonthonr F, Yuen A, Pellerin H, Cohen D, Grossard C. A serious game to train rhythmic abilities in children with Dyslexia: Feasibility and usability study. *JMIR Serious Games.* 2024 Jan 11;12:e42733.doi: <https://doi.org/10.2196/42733>
- [7] Choi J, Tamí-Mauri I, Cuccaro P, Kim S, Markham C. Digital health interventions to improve adolescent HPV vaccination: A systematic review. *Vaccines.* 2023 Jan 22;11(2):249.doi: <https://doi.org/10.3390/vaccines11020249>
- [8] Hakim H, Driedger SM, Gagnon D, Chevrier J, Roch G, Dubé E, et al. Digital gamification tools to enhance vaccine uptake: Scoping review. *JMIR Serious Games.* 2024 Feb 29;12:e47257.doi: <https://doi.org/10.2196/47257>
- [9] Ruggiano N, Perry TE. Conducting secondary analysis of qualitative data: Should we, can we, and how? *Qualitative Social Work.* 2019 Jan;18(1):81–97. doi: <https://doi.org/10.1177/1473325017700701>
- [10] Thorne S, Kirkham SR, O'Flynn-Magee K. The analytic challenge in interpretive description. *Int J Qual Methods.* 2004;3(1):1–11. doi: <https://doi.org/10.1177/160940690400300101>.
- [11] Parsons TD. *Cyberpsychology and the brain: The interaction of neuroscience and affective computing.* Cambridge: Cambridge University Press; 2017, doi: <https://doi.org/10.1177/1473325017700701>

- [12] Parsons TD. Ethical challenges in digital psychology and cyberpsychology. Cambridge: Cambridge University Press; 2019. doi: <https://doi.org/10.1017/9781108553384>
- [13] Corbin JM, Strauss A. Basics of qualitative research: Techniques and procedures for developing grounded theory. 3rd ed. Thousand Oaks, CA: SAGE Publications Inc; 2008. doi: <https://doi.org/10.4135/9781452230153>
- [14] Nagel SK, Reiner PB. Technologies of the extended mind: Defining the issues. In: Neuroethics: Oxford University Press; 2017. p. 108–122, doi: <https://doi.org/10.1177/1473325017700701>

Harmful Online Content and Cross-Cultural Challenges Faced by Adolescents: Findings From Focus Groups Conducted in Brazil and Australia

Alex S. G. PESSOA^{a,4}, Andrew J. CAMPBELL^b, and Brad RIDOUT^b

^a*Federal University of São Carlos (UFSCar), Laboratory of Analyses and Prevention of Violence, Graduate Program in Psychology, São Paulo, Brazil.*

^b*The University of Sydney, Cyberpsychology Research Group, Biomedical Informatics and Digital Health Theme, School of Medical Sciences, Faculty of Medicine and Health, New South Wales, Australia.*

ORCID ID: Alex Sandro Gomes Pessoa <https://orcid.org/0000-0002-9271-8575>

Abstract. Internet and social media content offers a psychological risk and protection challenge for the development of children and adolescents, depending upon the quality of supervision they receive. Moreover, understanding the levels of online content risk between diverse cultures is an under investigated area of research. To address this, the current study presents partial findings from an ongoing research project involving researchers from Brazil and Australia. Phase I of the study aimed to identify the self-reported contemporary concerns and experiences of adolescents in Brazil regarding harmful online content, as well as the perceived concerns of parents, educators and health service professionals. Thirty-three people from Brazil participated: 13 adolescents; 11 parents or legal guardians of adolescents; and 9 professionals who work with adolescents in different services. Data collection included a questionnaire to characterize sociodemographic variables and separate Focus Groups (FGs) with the above-mentioned groups. Descriptive statistics were used to analyze the sociodemographic variables, and qualitative data from FGs were analyzed through Thematic Analysis. 15 themes related to the risks and concerns raised by the participants were found, and 9 of them were similar among the three groups. This research may inform future interventions aimed at improving adults' guidance and supervision regarding adolescents' use of the internet and social media. Phase II will be replicating this study with an Australian population and will allow a cross-cultural analysis to determine whether there are any differences regarding concerns of the groups by country of origin of the participants.

Keywords. Social media, online violence, misinformation, bullying, youth, mental health

1. Introduction

The internet has profoundly changed social dynamics, interpersonal relationships, and developmental processes (social and psychological) of children and adolescents [1]. Researchers have been concerned about the influence that excessive internet use can have on the development of adolescents, especially when its use is accompanied by harmful behaviors and poor supervision by parents or other adults [2]. Unregulated content such as sexual content, inappropriate health advice, strategies for committing suicide or hurting oneself, and content related to drug use, all pose clear risks to the healthy psychosocial development and physical safety of adolescents [3].

While literature demonstrates the importance of supervision when adolescents are online [2], there is a scarcity of evidence-based programs designed to help parents and professionals to provide structured support in order to better psychosocially prepare them for the content they may encounter [4]. There is also a lack of recent cross-cultural studies into the concerns adolescents and adults have about adolescents' use of the internet and social media [5, 6], particularly for newer social media platforms such as TikTok and OnlyFans.

⁴ Corresponding Author: Alex Pessoa, alexpesso@ufscar.br

As part of an ongoing research project involving researchers from Brazil and Australia, this work-in-progress research aimed to identify the self-reported contemporary concerns and experiences of adolescents in Brazil regarding harmful online content, as well as the perceived concerns of parents, educators and health service professionals, using Focus Groups.

2. Methods

This research received approval from the Federal University of São Carlos Human Research Ethics Committee (Protocol number 79002224.5.0000.5504). The fieldwork involved recruiting three groups: (i) adolescents of all genders, aged 12-18 years old; (ii) parents or legal guardians of adolescents; (iii) professionals who had been directly involved with adolescents in public services and policies for at least one year. Recruitment of all participants was conducted through posting flyers in virtual environments (i.e., Facebook, Twitter, Instagram, and WhatsApp) and through snow-ball sampling. Adults had to sign a consent form to participate, while adolescents had to complete an assent form. After obtaining their signatures, a link was sent to potential participants to inform them about the subsequent steps.

Data collection occurred in two distinct stages via online platforms. In stage 1, participants completed a questionnaire to capture sociodemographic variables and adolescents' use of the internet and social media. Stage 2 involved participating in the Focus Groups (FG). The FG question script was predefined and included questions about adolescents' internet and social media usage time, their most frequent activities, online risks they identified and perceived, negative experiences they or other people they knew had faced, and challenges in supervision (in the case of FGs with parents and professionals), among other topics.

3. Results

3.1. Demographic

Thirty-three people participated in the research: 13 adolescents (FG1-A); 11 parents or legal guardians of adolescents (FG2-Pa); and 9 professionals who work with adolescents in different services (FG3-Pr). FG1-A consisted of adolescents (7 girls and 6 boys) aged between 14-17 years ($M = 15.76$; $SD = 1.36$), studying in public schools in São Paulo state. 76.9% reported being Black, 15.4% White, and 7.7% Indigenous. All participants identified as cisgender, with 84.6% identifying as heterosexual and 15.4% as bisexual. 38.5% reported spending more than 5 hours per day on the internet or social media, 30.8% said they spend up to 5 hours daily, 15.4% up to 3 hours, 7.7% up to 4 hours, and 7.7% were unsure. The social media platforms most used by participants were Instagram (100%), WhatsApp (100%), TikTok (92.3%), YouTube (53.8%), and websites or virtual platforms for school tasks (38.5%).

FG2-Pa was comprised of 9 mothers, 1 father, and 1 grandmother, residing in 3 different Brazilian states (São Paulo, Paraná, and Pará). Ages ranged from 33 to 69 years old ($M = 45.81$; $SD = 7.00$), with 6 participants identifying as white and 5 as lack people. Ten participants (90.9%) reported being responsible for only 1 adolescent, and 1 participant (9.1%) reported being responsible for 2 adolescents. All identified as cisgender and heterosexual. When asked about how much time they believed the adolescents they were responsible for spent daily on the internet or social media, 36.4% were unsure, 27.3% said up to 5 hours, 18.2% up to 3 hours, 9.1% up to 4 hours, and 9.1% more than 5 hours per day. Regarding their preparedness to supervise adolescents' internet and social media use, 54.5% said they sometimes felt prepared, 36.4% felt prepared, and 9.1% reported not feeling prepared at all.

FG3-Pr consisted of 6 women and 3 men, all residents of the state of São Paulo. The age of participants ranged from 30 to 48 years ($M = 37.33$; $SD = 5.43$), with 88.9% being White and 11.1% Black. 55.6% were psychologists, 22.2% social workers, 11.1% educators, and 11.1% lawyers. Regarding their years of experience working with adolescents, 66.7% had over 5 years of experience, 22.2% up to 4 years, and 11.1% up

to 5 years. All participants also identified themselves as cisgender and heterosexual. Based on their previous experiences with adolescents, they were asked how much time they believed adolescents, in general, spent using the internet and social media. 44.4% believed they spent more than 5 hours online per day, 22.2% up to 4 hours, 11.1% up to 5 hours, 11.1% up to 3 hours, and 11.1% were unsure. Regarding their preparedness to provide guidance to adolescents about internet and social media use, 55.6% of FG3-Pr said they sometimes felt prepared, while 44.4% said they did feel prepared. It is important to note that in both FG2-Pa and FG3-Pr, no participant reported feeling completely prepared to provide precise guidance to adolescents on this topic.

3.2. Main Concern and Challenges Revealed in the Focus Groups

The thematic analysis [7] of the data derived from the FGs revealed 15 themes related to the risks and concerns raised by the participants. Among these, 9 themes were similar across the three groups: (1) compulsive use of the internet and social media, as well as the implications for adolescents' mental health (I think he only doesn't use it [social media] when he's sleeping. – FG-Pa-P2); (2) strategies to evaluate and seek the credibility of the information, content, and sources accessed (Some teenagers don't seek medical attention because they follow advice from Google. – FG-Pr-P6); (3) prevalence of cyberbullying in virtual environments (I wear glasses and braces. Several people have already said horrible things to me online. It's sad! – FG1-A-P1); (4) abrupt changes in interpersonal relationships, with implications for face-to-face verbal communication (I can't talk to people anymore. I'm very shy and feel like I depend on my phone [to talk to other people]. – FG1-A-P5); (5) sexting and the non-consensual dissemination of intimate images (I know a girl who had her [intimate] photo posted on Instagram. – FG2-Pa-P8); (6) content associated with suicide induction (There were [online] challenges that led young people to commit suicide. – FG3-Pr-P8); (7) uncertainties about Artificial Intelligence (AI) and its usability (AI ends up harming my classmates because many people don't know how to write – FG1-A-P6); (8) gaming, both in terms of time spent and content (I've been monitoring teenagers who are betting [in online games] and doing it regularly – FG3-Pr-P5); (9) peer pressure exerted in virtual interactions (I'm afraid that soon someone will convince her to say: 'No, leave your mom out of this, you don't need to tell her'. – FG2-Pa-P3).

Additionally, professionals and adolescents shared concerns regarding (10) the induction of self-harming behaviors (There are [online challenges] of self-mutilation. There are [online] groups that do the opposite of our work in promoting mental health. – FG3-Pr-P2) and (11) the need for adolescents to be better prepared to make posts and engage in virtual interactions more critically and consciously (You have to think twice before posting and commenting. If you're going to comment, you have to think: 'If they were commenting this to me, would I like it?' – FG1-A-P20). Both parents and adolescents raised (12) concerns about online scams, such as card cloning and fraudulent sales (I tell him [adolescent]: 'Do you think you're only going to pay this amount?'. This is a scam. Can't you see the price is too low? – FG-Pa-P1). Professionals and parents reported that (13) they sometimes feel unprepared to provide guidance, especially due to the rapid changes occurring in virtual environments (I'm a complete layperson when it comes to technology, I'm terrible. So, I don't have the knowledge to talk about the internet. I think one of the first things we need is to get familiar with the subject. - FG3PrP1).

Finally, two themes that emerged only in FG1-A were (14) the existence of anonymous accounts and fake profiles ([I'm concerned about] anonymous accounts, because they create anonymous profiles, put the school's name, and start posting pictures of people who study at the school. They start speaking badly or well about them. – FG1-A-P2) and (15) the overload of school activities they receive. Regarding this last theme, adolescents highlighted the contradictions between the advice they receive from professionals and their parents (to reduce screen time) and the number of tasks they are exposed to daily with the school routine (One thing I don't understand is that the teachers complained that we were on our phones all the time, but now the state sends a lot of [digital] school activities. In my opinion, this contributes to phone dependency as well. –FG1-A-P1).

4. Discussion

The results showed that the adolescents who participated in the study use the internet and social media intensively on a daily basis, which is similar to findings from other countries [8, 9]. No professional or parent reported feeling completely prepared to supervise or guide adolescents regarding their usage time, which may be explained by the lack of educational programs targeted at parents [10, 11], or the scarcity of training opportunities for professionals regarding technologies and virtual interactions [12]. All groups reported several challenges and problems that adolescents are exposed to online, however, it was noted that some themes seemed to be universal (i.e., all groups raised them) while others appeared were raised by only one or two groups. These results could inform future interventions, specifically targeting each group by ensuring they include information on risks and concerns that weren't raised by that group (e.g., parents did not raise self-harm as an issue, professionals did not raise scams as an issue). It is important to highlight that the adolescents were able to describe in detail the various risks to which they are exposed to, mainly because they had already experienced online challenges or were aware of incidents that occurred to others. These findings can challenge misleading narratives that adolescents are unaware of online risks and, therefore, more susceptible to being victims of different forms of online rights violations [13]. Nonetheless, it is recognized that adolescents need access to self-protection strategies and should be aware of places they can reach out to if they need help [14].

5. Conclusion

This research may inform future interventions aimed at improving parental supervision regarding adolescents' use of the internet and social media, as well as the associated risks. Additionally, the data compiled in this investigation can be utilized in developing training programs for professionals who work with this population, enhancing their knowledge and repertoire on these topics and enabling them to provide more precise and effective guidance to their clients. Replicating this study with the Australian population will allow a cross-cultural analysis to determine whether the concerns of the groups are the same regardless of the country of origin or if there are specific issues related to different cultures.

References

- [1] Senekal JS, Ruth G, Wolfaardt L, Jansen C, Williams K. Social media and adolescent psychosocial development: a systematic review. *South African Journal of Psychology*. 2023;53(2):157–171. doi: <https://doi.org/10.1177/00812463221119302>.
- [2] Ren W, Zhu X. Parental mediation and adolescents' internet use: the moderating role of parenting style. *J Youth Adolesc*. 2022;51:1483–1496. doi: <https://doi.org/10.1007/s10964-022-01600-w>.
- [3] Purba AK, Thomson RM, Henery PM, Pearce A, Henderson M, Katikireddi SV. Social media use and health risk behaviours in young people: systematic review and meta-analysis. *BMJ*. 2023;383:e073552. doi: <https://doi.org/10.1136/bmj-2022-073552>.
- [4] Morgan MHC, Huber Krum S, Willis LA, Shortt JW. A literature review of digital behavioral parent training programs for parents of adolescents. *Prev Sci*. 2024;25:155–174. doi: <https://doi.org/10.1007/s11121-023-01596-0>.
- [5] Wang X, Liu Z. Online engagement in social media: a cross-cultural comparison. *Comput Hum Behav*. 2019;97:137–150. doi: <https://doi.org/10.1016/j.chb.2019.03.014>.
- [6] Xenos M, Vromen A, Loader BD. The great equalizer? Patterns of social media use and youth political engagement in three advanced democracies. *Inf Commun Soc*. 2014;17(2):151–167. doi: <https://doi.org/10.1080/1369118X.2013.871318>.
- [7] Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77–101. doi: <https://doi.org/10.1191/1478088706qp063oa>.
- [8] Squire KD, Steinkuehler C. The problem with screen time. *Teach Coll Rec*. 2017;119(12):1–24. doi: <https://doi.org/10.1177/016146811711901207>.
- [9] Rideout V. The Common Sense Census: Media Use by Tweens and Teens. Los Angeles, CA: Common Sense Media; 2015.
- [10] Symons K, Vanwesenbeeck I, Walrave M, Van Ouytsel J, Ponnet K. Parents' concerns over internet use, their engagement in interaction restrictions, and adolescents' behavior on social networking sites. *Youth Soc*. 2020;52(8):1569–1581. doi: <https://doi.org/10.1177/00441118X19834769>.
- [11] Dias P, Brito R, Ribbens W, Daniela L, Rubene Z, Dreier M, et al. The role of parents in the engagement of young children with digital technologies: exploring tensions between rights of access and protection,

- from ‘gatekeepers’ to ‘scaffolders.’ *Glob Stud Child.* 2016;6(4):414–427. doi: <https://doi.org/10.1177/2043610616676024>.
- [12] Van Den Beemt A, Thurlings M, Willems M. Towards an understanding of social media use in the classroom: a literature review. *Technol Pedagogy Educ.* 2019;29(1):35–55. doi: <https://doi.org/10.1080/1475939X.2019.1695657>.
- [13] Weinstein E. The social media see-saw: positive and negative influences on adolescents’ affective well-being. *New Media Soc.* 2018;20(10):3597–3623. doi: <https://doi.org/10.1177/1461444818755634>.
- [14] Kvádová N, Smáhel D, Machácková H, et al. Who is exposed to harmful online content? The role of risk and protective factors among Czech, Finnish, and Spanish adolescents. *J Youth Adolesc.* 2021;50:2294–2310. doi: <https://doi.org/10.1007/s10964-021-01422-2>.

Leveraging Artificial Intelligence to mitigate User Susceptibility to Malicious Push Notifications during Augmented Reality Immersion

Sarah KATZ ^{a,5}

^a*Department of Cyberpsychology, Capitol Technology University, U.S.A.*
ORCID ID: Sarah Katz <https://orcid.org/0000-0002-9066-5270>

Abstract. Distraction caused by the visual processing of multiple objects during augmented reality (AR) immersion could make users more susceptible to malicious push notifications. This qualitative empirical study used a virtual presentation to simulate the user interfaces of the popular augmented reality applications Instagram, Maps, and Pokémon GO along with an animated push notification entering the user view of 70 participants aged 18-40 who regularly used these applications. Of the two notification themes presented – familiarity and urgency – 62 of 70 participants chose the familiarity theme to engage with. Based on these results, four experts in the field of AR application development were consulted to design a proposed artificial intelligence (AI) equipped feature that could detect possibly malicious artifacts entering the user's line of sight during partial immersion in augmented reality. Participants included a business partner at an AR firm, a security engineering manager, an AI engineer focused on machine learning, and a data analytics specialist. The case study determined that a security application natively implemented into the device could use heuristic analysis of user screen captured activity to assess potentially malicious push notifications in real time, with a focus on familiarity-related content.

Keywords. Cybersecurity, cyberpsychology, augmented reality, artificial intelligence, application development.

1. Introduction

The level of distraction during such experiences as augmented reality (AR) immersion is often believed to make the users vulnerable to social engineering, including real-time malicious push notifications. While existing studies have emphasized distraction as a vector exploited by threat actors for social engineering using tactics such as communication requests and time-sensitive requests [1], a dearth of research remains as to how extended reality-specific distraction affects user judgment of the legitimacy of objects appearing on screen in real-time. Provided the increasing use of augmented reality mobile applications in both a leisure and workplace setting, this two-stage qualitative empirical study aims to assess user susceptibility in order to subsequently propose a hypothetical security solution to help mitigate this threat.

2. Methods

The methodology involved an in-depth questionnaire administered via email to 70 participants (regular users of augmented reality for both leisure and at the workplace, aged 18-40) regarding interaction with push notifications within an AR simulation. Each participant was presented with a link to an online animation-enhanced simulation centering the following AR mobile interfaces suggested as popular AR apps [2], and regularly used by the selected participants: Google Lens, Google Translate, Instagram, Maps, and Pokémon GO. The simulation presented users with the option of engaging

⁵ Corresponding Author: Sarah Katz, skatz@captechu.edu.

with a push notification appearing on screen while using all five applications, with option A denoting a familiarity theme (“Contact calling in”) and option B denoting an urgency theme (“Update now or device will re-start in 5 minutes”). Users were then asked to explain the rationale behind their choices. Collated responses to the questionnaire for both the leisure and workplace settings showed a trending preference for the familiarity theme, with recurring reasoning citing the assumption that a friend or co-worker might be calling in. As predicted by the hypothesis, this sense of obligation tended to supersede the risk of data progress loss upon a forced reboot.

Given this risk to both individual users and organizational reputation, such a threat vector could be mitigated from the developer standpoint. To propose a solution, four experts were consulted in the business, backend, and user experience fields of AR application development to design a hypothetical artificial intelligence (AI) equipped feature that could detect suspicious artifacts entering the user’s line of sight during partial immersion in an AR application at the workplace. Participants included a business partner at an AR firm and a security engineering manager at a top Fortune 500 organization for the market appeal of such a feature; an AI engineer focused on machine learning for a detailed concept of the feature’s functionality, and a user experience (UX) specialist for consideration into potential user-facing setbacks.

Interview questions were administered via email, consulting each respective participant regarding the backend structure and UX challenges of an application disrupting real-time social engineering threats, with emphasis on familiarity-based onscreen notifications.

This mitigation proposal stage of the study determined that a security application natively implemented into the device could use heuristic analysis of user screen captured activity to assess potentially malicious push notifications in real time.

3. Conclusion

Overall participant responses revealed that, while necessary for user and organizational security, such an application has not yet been developed due to challenges in balancing application visibility with user privacy. Therefore, future studies could entail a merging of engineering and privacy expertise to devise a realistic framework for such a feature toward mitigating user susceptibility to social engineering during augmented reality immersion.

References

- [1] Siddiqi MA, Pak W, Siddiqi MA. A study on the psychology of social engineering-based cyberattacks and existing countermeasures. *Appl Sci*. 2022;12(12):6042. doi: <https://doi.org/10.3390/app12126042>.
- [2] Ramic M. Best augmented reality apps of 2023. Takeaway Reality - Spatial Computing Agency - AR/VR. 2024. Available from: <https://www.takeaway-reality.com/post/best-augmented-reality-ar-projects-of-2023>

XR for STEM Education and Social Health Outcomes: Focus on Embodied Mixed Reality

Mina C. JOHNSON-GLENBERGA,^{a,6} and Yanjun LYU^b

^a*University of Arizona, Department of Psychology*

^b*Arizona State University, School of Arts, Media and Engineering*

ORCiD ID: Mina C. Johnson-Glenberg <https://orcid.org/0000-0003-0941-3744>

Abstract. This paper proposes an update to the eXtended Realities (XR) continuum. The continuum's platforms range from Augmented Reality (AR) to Mixed Reality (MR) to Virtual Reality (VR) and this continuum has been primarily based on Milgram and Kishino's (1994) visual dimension. We explore how adding a second one called the embodiment/haptics dimension can expand and further define the XR continuum. We discuss designs for effective experiences in multi-modal platforms. Within each of the three platforms supporting the XR continuum, there can be lower or higher amounts of embodiment designed into the experience. These levels of embodiment can, and should, be quantified and included as predictor variables in theories of change. We focus here on the MR space and describe three Mixed Reality environments that successfully incorporate embodiment. We report some outcomes that suggest how important it is to report on and explicate interactions (mediators and moderators) in MR studies.

Keywords. Mixed reality, XR eXtended realities, stem education, social emotional learning, aptitude by treatment interactions, computer-supported learning

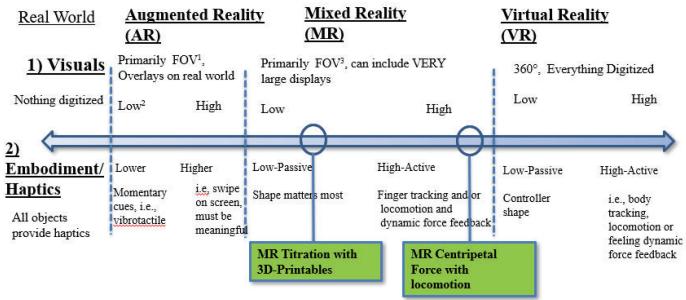
1. Introduction

The first part of this paper defines the spectrum of learning technology platforms called eXtended Realities (XR). These platforms range from Augmented Reality (AR) to Mixed Reality (MR) to Virtual Reality (VR). This article explores how to design efficacious experiences in the three platforms and focuses on the middle range of MR. We take a multimodal design perspective and encourage designers to include more multimodal assessment methodologies [1] as well. The first two learning signals or modalities that we are all used to are those in lectures - the auditory and visual signals. In mediated (computer-driven) environments, it is possible now to easily incorporate more modalities into learning. We focus on a signal or dimension called haptics/embodiment. The second part of the article explores an example of an innovative, multi-modal interaction environment created in MR for children called SensaSea.

The XR Spectrum. The well-known graphic called the “Reality-Virtuality Continuum” [2] starts with the real world on the left end and ends with the fully virtual world on the far-right end. The XR field uses agreed-upon labels for the three main spaces (platforms) along this continuum: AR to MR to VR. The borders between these three platforms, or spaces, are somewhat fluid. Figure 1 represents an extension of Milgram and Kishino’s spectrum. There are two new ideas to highlight in this updated continuum, the first is the additional dimension of embodiment/haptics [3], and the second is the concept of sublevels (or degrees) of haptics/embodiment WITHIN each platform. That is, there can be fewer or more opportunities for embodiment within each type of environment - be it AR on a mobile device, MR in a large room with manipulable, or a fully enclosed VR system. **Figure 1** shows the new continuum with some explanatory notes.

⁶ Corresponding Author: Mina C. Johnson-Glenberg, mina.johnson@asu.edu.

XR Spectrum: Dimensions of Visuals and Haptics



¹ FOV= Field of View affects both *immersion* (objective and technology-induced) and *presence* (subjective, the feeling of "being there").

² Low visual immersion this relates to amount of digitization perceived. This means overlays on a phone display are smaller than those on a larger tablet display.

³ Field of View can be very large in MR. It could be a Giant Screen JMAX film, but by definition, MR is *not* delivered in a head-mounted 360 degree delivery platform; that term is reserved for VR.

Figure 1. The newest XR continuum with embodiment/haptics as a second dimension.

2. STEM Education Gains in Two MR Platforms

Below we explain two different MR systems that fall in two different places along the new embodied XR continuum. The first, MR Titration, is somewhat low because the visuals were constrained to a laptop field of view (FOV), however, the titration experiment also included a tangible device that mimics a real-world device. This device controlled the action in the digitized lesson - that would be considered medium high on the haptics/embodiment dimension. It is not the highest because the printed burette only allowed for passive force feedback. The second lesson listed in Figure 1 was on the topic of physics and centripetal force, it included a very large display and locomotion, so it was rather high on the embodiment/haptics dimension within the platform of MR.

Instructional designers in XR should be creating learning content and healthcare experiences that take advantage of some of the profound affordances of XR, namely embodiment and the agency associated with manipulating content in multiple dimensions [4]. Manipulating objects in multiple dimensions can make the environment seem even more immersive, and user-initiated actions can add to the experience of agency [5]. Several embodied studies report that when action is designed in the lesson learners exhibit higher knowledge gains and better retention of content. However, with an increase in modal signals (that is, beyond the usual auditory and visual signals), there is an increase in the potential for cognitive overload. The hypothesis is that adding a sensory-motor gesture/action signal to a lesson or experience might have negative effects on certain users. It is probably the case that different learners will be differentially affected by varying platforms *and* varying affordances of the systems. This is the definition of an interaction – when levels of X (e.g., prior knowledge) at different levels of Y (e.g., varying platforms) differentially affect outcomes. The field is in need of new methods of assessment based on action and embodiment while keeping in mind affordances along the XR spectrum. Below, we report an example of a more sophisticated statistical analysis of interaction effects in an MR environment.

2.1. Titration

For the chemistry titration study [3], a 3D-printed burette device was designed and then tethered to a laptop-based titration lesson. This plastic burette mimicked the look and feel of a real, tangible glass burette. Students were randomly assigned to two conditions: they either 1) physically turned the burette valve (high embodied condition) or 2) tapped computer keys to control the virtual burette on screen (low embodied). The second condition was considered low embodied because users did not do the same twisting gesture that students do in real chemistry labs. (Note: The burette group used significantly more twisting gestures during a scored recall.) Both groups showed equal

chemistry learning gains after the intervention. However, an aptitude by treatment (ATI) analysis revealed that the lower prior knowledge students learned significantly more if they were in the low embodied keyboard condition. This suggests that being in a more haptic and more novel environment was deleterious for the low-knowledge students. The high embodied content differentially helped the higher prior knowledge students. This result has important implications for how we design novel MR environments. Most MR environments will feel novel the first few times they are used. Our results suggest, at least for STEM, it may be helpful to pre-train low-knowledge users before placing them in novel environments. Additionally, we designers need to be thoughtful about embedding the pedagogy of scaffolding into lessons. **Figure 2** shows the regression results on posttest knowledge, the solid (blue) line representing the users in the high embodied burette group; the dashed line is the control group.

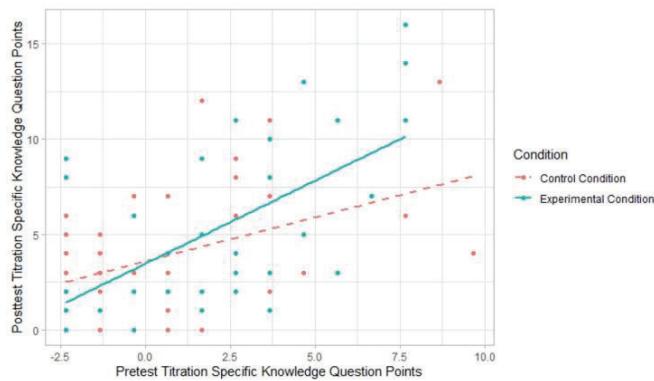


Figure 2. Results showing the significant ATI interaction between lower and higher prior knowledge students in either a low or high embodied MR condition.

2.2. Centripetal Force

This second study is higher up on the level of embodiment as seen in Figure 1. We created an MR environment called *SMALLab*, an immersive 15 by 15 ft floor projection with trackable objects. Johnson-Glenberg et al. [6] published results on a 3 X 2 X 3 mixed design. A full description can be found in the 2016 article. We note here that a randomized control trial study with 109 students used the following conditions:

- 1) **High-embodied**, swing the tethered object overhead, users can locomote and spin, they also feel centripetal force through their whole body.
- 2) **Mid-embodied**, create circles on a large interactive whiteboard with a trackable pen, users feel arm moving and some shoulder activity.
- 3) **Low-embodied** – sit and use a mouse and PC to draw smaller circles, users feel actions in the arm, wrist, and some shoulder activity.

We wanted to understand how these three platforms interacted with two levels of embodiment designed into each platform and if that differently affected immediate post-test learning scores and/or delayed retention (the interaction question). The retention test was given one week after the intervention. There was no main effect on the platform. However, for delayed retention, only those in high embodied conditions retained the STEM material, while those in the lower embodied conditions showed significant losses. This supports the interaction of time by level of embodiment argument in MR. This means that when users act upon or DO the activity, they retain knowledge about it longer than those who did not physically do the activities.

3. *SensaSea*: Fostering Children’s Social Collaboration in Responsive Multisensory Environments

The *SensaSea* System is designed by Yanjun Lyu; it is a new responsive multisensory environment designed and housed at Arizona State University, Tempe. *SensaSea* is a room-sized interactive installation that incorporates wearable devices, interactive visual floor projections, auditory, and tactile modalities. *SensaSea* is designed as a physical environment for embodied interaction [7] and free play suitable for multiple players; the system uses social proximity as the primary mechanism. The main objective is to promote active peer interaction and social connectedness among elementary school children through sensory-guided approaches. The multi-modal system also features an interactive soundscape and innovative real-time haptic feedback. In the current system, visiting children are provided with attractive and customizable vests that include sensors and feedback devices. Three or more children can be tracked and active in *SensaSea* which includes infra-red tracking cameras and colorful projections of the ocean and sea creatures. **Figure 3** shows the ocean-like floor projection.

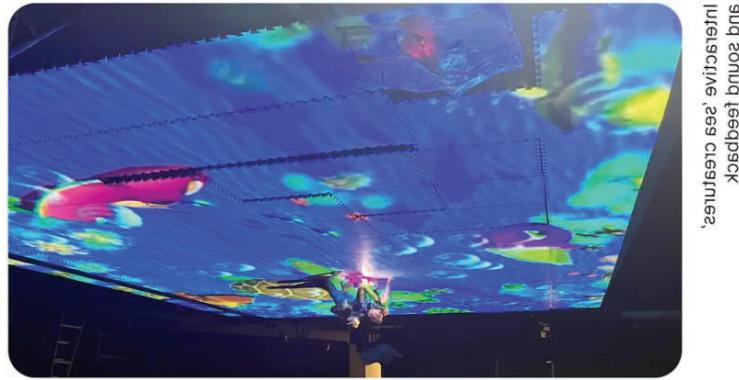


Figure 3. Two children in the middle of the oceanic *SensaSea* MR environment, 20 X 20 ft. projection.

As players move and locomote, their actions “stir up the ocean” causing sea creatures to follow in their wake. In later stages, as the children approach one another, their associated creatures (turtle, sea horse, etc.) begin to also digitally interact in the projection in real-time. When all three players converge in close physical proximity, their corresponding sea creatures (aka digital avatars) ‘leap’ from the ocean and swim in a circular formation. Additional bubbles, a joyful soundtrack, and a buzzing haptic sensation across the shoulders (sensors are sewn into the vests) are linked to the players’ actions and proximity cues. The haptic vests feature multiple actuators and a custom-programmed microprocessor, offering a wide range of vibrotactile patterns. This design elegantly incorporates multiple sensory modalities, including visual, auditory, and vibrotactile patterns with are coupled with playful game mechanisms.

Usability and feasibility tests have revealed that the system results in positive emotions and elicits multiple pro-social behaviors. Particularly when children notice their creature/avatar jumping, the children mimic those behaviors by jumping together with their peers. We have seen spontaneous, collaborative actions such as holding hands and verbally inviting peers to join in. *SensaSea* also promises benefits in health and wellness from the following perspectives:

Therapeutic Play and Social Connectedness: Children may encounter emotional challenges related to their condition or long-term hospitalization [8]. They might face challenges in engaging with their peers and participating in recreational activities. Furthermore, medical procedures can induce stress and heighten anxiety. *SensaSea*’s game-like environment provides safe and relaxing spaces for children to explore and express themselves creatively, fostering resilience, and distracting them from the hospital environment. Its design supports multiple players, and the reward structure encourages active social communication, a sense of togetherness, and inclusion. Its adaptable mobility features can be tailored to different room settings.

Physical and Cognitive Benefits: Understanding the interplay of action, cognition, and the environment is crucial in children's interaction design. Our system fosters active and embodied thinking by individuals thereby enhancing sensory perception, motor skills, and cognitive processing. The dynamic haptic feedback design enhances children's sensory perception and motor skills, facilitating collaborative play through verbal sharing of experiences and embodied interaction.

Rehabilitation: Rehabilitation or physical therapy for children's sensory integration can be tailored to target specific motor skills or rehabilitation goals, providing real-time sensory feedback and rewards through sensory cues. Therapists can design interactive exercises utilizing haptic wearables with multiple motors to support rehabilitation. The system is especially motivating for youth who need to practice balance and walking (locomotion).

Parent and Caregiver Engagement: The future development of *SensaSea* installation should consider designing with and for parents and/or caregivers in the play experience. Having a room in a hospital environment that can encourage joint participation and bonding between children and their caregivers or medical professionals [9, 10] seems like a crucial component in fostering a more positive hospital stay and potentially faster recovery.

4. Conclusion

We encourage designers to create content with multimodality in mind. We encourage learning scientists and investigators to analyze statistical interactions as they seek answers about how different XR platforms affect users. It is certainly the case that learners and users with varying skill sets, and in different developmental stages, are differentially affected by the amount of embodiment designed into any XR system.

References

- [1] Walkington C, Nathan MJ, Huang W, Hunnicut J, Washington J. Multimodal analysis of interaction data from embodied education technologies. *Educ Technol Res Dev*. 2023.
- [2] Milgram P, Kishino AF. Taxonomy of mixed reality visual displays. *IEICE Trans Inf Syst*. 1994;E77-D(12):1321–1329.
- [3] Johnson-Glenberg MC, Yu CS, Liu F, Amador C, Bao Y, Yu S, LiKamWa R. Embodied mixed reality with passive haptics in STEM education: randomized control study with chemistry titration. *Front Virtual Real*. 2023;4:1047833. doi: <https://doi.org/10.3389/fvir.2023.1047833>
- [4] Johnson-Glenberg MC. Immersive VR and education: embodied design principles that include gesture and hand controls. *Front Robot AI*. 2018;5:81. doi: <https://doi.org/10.3389/frobt.2018.00081>
- [5] Geden M, Emerson A, Carpenter D, Rowe J, Azevedo R, Lester J. Predictive student modeling in game-based learning environments with word embedding representations of reflection. *Int J Artif Intell Educ*. 2020;31(1):1–23. doi: <https://doi.org/10.1007/s40593-020-00210-6>
- [6] Johnson-Glenberg MC, Megowan-Romanowicz C, Birchfield DA, Savio-Ramos C. Effects of embodied learning and digital platform on the retention of physics content: centripetal force. *Front Psychol*. 2016;7:2133. doi: <https://doi.org/10.3389/fpsyg.2016.02133>
- [7] Dourish P. Where the action is: the foundations of embodied interaction. Cambridge, MA: MIT Press; 2001.
- [8] Mattsson A. Long-term physical illness in childhood: a challenge to psychosocial adaptation. *Pediatrics*. 1972;50(5):801–811.
- [9] Shin JY, Rheu M, Huh-Yoo J, Peng W. Designing technologies to support parent-child relationships: a review of current findings and suggestions for future directions. *Proc ACM Hum Comput Interact*. 2021;5(CSCW2):1–31. doi: <https://doi.org/10.1145/3479545>
- [10] Yip JC, Clegg T, Ahn J, Uchiduno JO, Bonsignore E, Beck A, Mills K. The evolution of engagements and social bonds during child-parent co-design. *Proc 2016 CHI Conf Hum Factors Comput Syst*. 2016:3607–3619. doi: <https://doi.org/10.1145/2858036.2858389>

A Game of Mindless Swiping: How Gamification and Mindless Swiping Affects Online Dating Users' Psychology

Aleksandra Bombaci FROELICH ^{a,7,8}

^a*Arizona State University*

ORCiD ID: Aleksandra Bombaci Froelich <https://orcid.org/0009-0007-7549-7970>

Abstract. Online dating has become a major source of personal connection for users worldwide. The User Interface and User Experience (UI/UX) of applications can influence, impact, and affect users in a multitude of ways. Using certain design elements such as particular colors, design hierarchies, and implicit associations, the designers of the applications can encourage particular behaviors, beliefs, and feelings in their users. The simple swipe design, in combination with bright colors, simplicity of use, and other factors, encourages users to find Tinder to be more like a game or shopping experience [1, 2, 3]. The gamification and simple swipe create a perfect storm where users are regularly caught in never-ending loops searching for fulfillment and validation, driven by negative feelings [4] and lackluster results, fed by hope and simplicity. There is much speculation on what design elements and user demographics play into how likely users are to get caught in, and the potential side effects of the perpetual loop. This study aims to look at two potential elements that may encourage the looping effect: the swipe feature and the gamification of the experience. An online survey of 18-35-year-old Tinder users ($N = 74$) was conducted, coded, and analyzed. The results of the study provide a cursory understanding of the relationships between and the impacts of the swipe feature, the gamification of Tinder, and user psychology regarding online dating with Tinder.

Keywords. Online dating, user interface and user experience, UI/UX design, swiping, gamification, mental health

1. Introduction

In 2023, dating migrated to a predominantly online activity [5], with three-in-ten adults in the United States having used online dating sites or apps [6,7]. With over 430,000 results when searching Google Scholar for “The Emotional Effects of Online Dating,” it is clear that this has become a popular research topic.

Remarkably absent from these papers is any discussion of the impact of the design or physical attributes of the applications/sites themselves on users. From feeling an immediate sense of satisfaction and gratification to the use of colors, tools, and functionality, the design of applications and sites have immediate quantifiable effects on their users, i.e., generating a sense of success, frustration, or failure [8]. Reviewing the top US dating app, Tinder [9], this paper hopes to address this gap in research to better understand how the UI/UX of dating apps affect a user’s mental health.

Tinder uses design hierarchy, color psychology, gamification and implicit associations to influence how its users interact with the application and one another. Bright colors [1], simplicity of use, words like “keep playing” after matching with other users add to the game-like design [3], and the simple swipe design gamified the matching process [2]. The effects of the game-like swiping design extend beyond just making dating a game for fun. It has created a way of thinking that now has users seeing

⁷ Corresponding Author: Aleksandra Bombaci Froelich, aleksfroelich@gmail.com

⁸ This is derivative work of the author's thesis: <https://keep.lib.asu.edu/items/193366>

dating not only as a game where there are winners and losers but as permission to objectify others and see them as commodities [10, 11, 12, 13].

The UI/UX team at Tinder believes the simple swiping method of rejection helps suppress the guilt their users could have when they are rejecting someone [1]. The effortless task of swiping left or right is also believed to remove the sting of rejection when users do not match [1]. However, when there is not an instant match, and the milliseconds pass users by, it is another hit of rejection. Research shows that even mild rejection can cause psychological and physical pain equal to that of grand rejection [14]. These negative feelings drive users' need for fulfillment and validation, forcing them back into the never-ending cycle of online dating [4]. Online dating has simply created the perfect looping effect to draw users back regardless of their most recent outcome. Didn't find love, "Come back, keep looking." "Found it and lost it", "See it worked, come back, keep trying." So on and so forth, the loop begins, again and again: new updates, new apps, drawing new users.

The purpose of this study was to determine in what ways Tinder's design elements impacted and influenced its' users. The survey asked questions concerning users' beliefs, views, and feelings regarding different UI/UX elements within Tinder. It also asked how they felt those elements impacted and influenced their own psychology and how they interacted with Tinder and other users.

2. Methods

2.1. Demographics

With 74 total participants, just over half self-identified as female (n=43), nearly a third self-identified as male (n=27), and the remaining participants self-identified as non-binary or third gender (n=4). Overall, the average age of a participant was 21. For both female and male participants, the average age was 24, and for non-binary (or third gender) participants, it was 22. The majority (72.97%) of participants self-identified as White (n=54), and 27.03% (n=20) identified as a minority group member (Asian - 12.16%, Hispanic or Latino - 8.11%, Black or African American - 4.05%).

The largest held occupation was student at 43.24%, then business/desk jobs (13.51%), and food services (9.46%). The majority of participants (61.16%) had used Tinder on and off for multiple years. Participants who had used using Tinder for only a year were the next largest group (22.97%), followed by those who had used the Tinder for three or more years (9.46%) and then two years (5.41%).

2.2. Methodology

This study used an online survey consisting of questions ranging from free response and short answer to dichotomous, nominal, ordinal, interval-level, and continuous questions. It took an average of 15 minutes to complete. Participants were collected via a mixed approach of snowball sampling and convenience sampling. The survey began with a consent form that clarified the requirements for participating (e.g. age and use of Tinder), that they had the right to leave the survey at any time and were not obligated to answer any questions they felt uncomfortable responding to.

The survey collected 128 participants (n=128); however, following initial data analysis, only 74 participants' surveys were completed to a satisfactory percentage for further analysis (n=74). The qualitative data was coded inductively to find themes, patterns, categories, and groups in order to better analyze the data. The remaining data was organized and coded for analysis.

3. Results

3.1. User Interface and User Experience

Participants were asked several questions regarding their perspectives on different UI/UX elements within Tinder. Nearly half (45.21%) agreed that Tinder's design influenced how they interacted with the application. When asked to elaborate on how they believed the designs influenced them, at equal positions at the top, participants believed that Tinder encouraged commoditization and a lack of humanity (20.27%), as well as quick thinking and instant gratification (16.22%).

Participants were also asked to rank the order of importance of different elements of the users' profile based on how they believe Tinder intends them to interpret the level of importance versus their personal beliefs about what was the most important. Based on how participants interpreted Tinder to show them what was the most important, the profile was ranked as the most important 100% of the time, and name was ranked as the least 56.4% of the time. Participants' personal beliefs, ranked the profile picture as the most important 92.3% of the time, and name as the least important 89.7% of the time. Overall, participants ranked the profile picture as the most important 96.2% of the time and ranked the name as the least 75.6% of the time.

Participants were also asked about their swiping habits and whether they found themselves swiping without thought. Participants were 4.5x more likely to mindlessly swipe Often (O)/ All The Time (A) than they were to mindlessly swipe Rarely (R)/ Never (N). Non-binary (or third gender) participants, over half (53.48%) of females and (66.66%) males said it happened either Often or All of the time. A low-level correlation between age and swiping habits was found (-0.309, $p=0.007$), implying that the older users got, the less likely they were to mindlessly swipe.

The data showed a low-level positive correlation between mindless swiping habits and whether users agreed with Tinder feeling like a game/shopping experience (0.393, $p=0.001$), suggesting that the more users mindlessly swiped the more they agreed that Tinder felt like a game/shopping experience. Participants who O/A mindlessly swipe were 7.5x more likely to say Online Dating with Tinder is "A Game.". Participants who did not believe Tinder influenced how they interacted with the application were nearly twice as likely to mindlessly swipe O/A and were 7.5 more likely to agree that it felt like a game/shopping experience. When asked directly, nearly 90% ($n = 66$) of the participants agreed that "The way Tinder is designed and how it is used can often make online dating feel like a game/shopping experience".

3.2. User Psychology

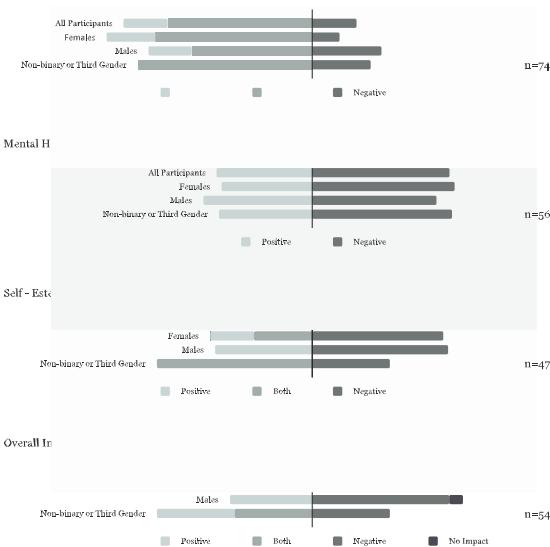


Figure 1. Figure 1 depicts participant responses to questions regarding user psychology by gender.

Participants were questioned regarding different psychological aspects related to using Tinder (See Figure 1). The majority of participants agreed that it had both positive and negative effects on their sense of self. The data suggests that the longer participants used the app, the more likely they were to say that it had had a negative impact on their sense of self. Participants who O/A mindlessly swiped were 2.5x more likely to say it had a negative impact on their sense of self.

Of the participants who responded ($n = 57$), more than 50% implied that using Tinder had a negative effect on their mental health. Females and Non-Binary (or third gender) were more likely to have had negative impacts than males. Participants who agreed that using Tinder felt like a game/ shopping experience were 1.45x more likely to have had a negative impact on their mental health, whereas those who disagreed were twice as likely to say it had a positive impact.

Of the participants who discussed the effects on their ($n = 47$) self-esteem, the overall impression was using Tinder had negatively impacted their self-esteem (53.32%). Participants were nearly twice as likely to have had a negative impact on their self-esteem. Regardless of gender or age, the participants said it had harmed their self-esteem; the only group to have a higher percentage response as a positive impact on their self-esteem were users who had only used Tinder for one year. The data in this study suggests that the more often participants mindlessly swipe, the more likely they are to have had a negative experience. Participants who never mindlessly swipe indicated that they only had positive impacts, whereas those who did it “All of The Time” were 5x more likely to have had a negative impact on their self-esteem.

Participants who agreed with Tinder feeling like a game/ shopping experience were almost twice as likely to say it had a negative impact on their self-esteem.

The majority of participants (51.85%) experienced a negative overall impact on psychological well-being while using Tinder. Younger participants (18-25) were 2.04x more likely to have had a negative overall impact than those ages 26-35. Female participants were nearly twice as likely as their male counterparts to have had negative overall experiences, while non-binary (or third-gender) participants were evenly split. The overall impact on users was negative regardless of how long they had been using Tinder. However, users who had been on Tinder for more than one year or on and off for multiple years were twice as likely to have had negative overall impacts.

4. Discussion

4.1. Further Research

This research was a starting point in understanding how the current UI/UX design of dating applications (Tinder) impacts their users. This is an ever-growing concern for the effects of social media platforms on their users and how designers, users, and even governmental regulation can help reduce the damage these platforms cause. Further research in this area is imperative to creating a better experience and overall safer environment for users online. Using comparative analysis, user testing, and studying more defined user groups, such as those with disabilities or those of different gender identities or sexualities, could assist in a more targeted understanding of the impacts different design elements have on particular parties and how to mitigate those effects. Additionally, longitudinal studies would allow researchers to understand the impact online dating has on users and society over time. Finally, exploring specific aspects of the applications, such as filtering and preferences or safety features, could help to understand how users feel regarding these features, (i.e. discussions regarding filters encouraging discrimination or cruelty), and where improvements can be made to make a positive online dating experience for all.

4.2. Limitations

One limitation of this study was that it was conducted in the Spring of 2022, during the pandemic. This may have impacted the data as many users became more aware of the pitfalls of online dating and made their unhappiness and frustrations well-known. However, it is believed that the data was largely unaffected by this, as the majority of users had been using online dating prior to the pandemic. Another limitation was the study size; a larger, more diverse population of users could aid in a better, more generalizable set of conclusions. Finally, as self-report surveys can introduce response bias, interviews, focus groups, and user testing could aid in a richer, more in-depth understanding of users' personal experiences, perceptions, desires, and frustrations with online dating and its design.

4.3. Conclusion

This research offers support and discourse regarding the use of Tinder, and its impact and influence on its users. The data suggests many of Tinder's UI/UX elements impact and influence how users use and interact with the application. The data also suggests that the majority of the psychological impacts on users were negative. The results of this study support the belief that the gamification and swiping design combined with other UI/UX elements increase the potential negative effects of using Tinder.

References

- [1] Faller P. For the love of UX: Tinder's product designers talk user-centered design for emotional experiences [Internet]. Adobe XD Ideas. 2018 [cited 2021 Apr 9]. Available from: <https://xd.adobe.com/ideas/perspectives/interviews/love-ux-tinders-product-designers-talk-user-centered/>.
- [2] Nik I. I've tried Tinder—here's what I think about it as a UX designer [Internet]. Medium. 2020 [cited 2022 Feb 9]. Available from: <https://uxdesign.cc/ive-tried-tinder-heres-what-i-think-about-it-as-a-ux-designer-aae6fa88c517>.
- [3] Stampler L. Inside Tinder: meet the guys who turned dating into an addiction [Internet]. TIME. 2014 [cited 2022 Sep 6]. Available from: <http://time.com/4837/tinder-meet-the-guys-who-turned-dating-into-an-addiction/>.
- [4] Sales NJ. Tinder and the dawn of the dating apocalypse [Internet]. Vanity Fair. 2015 [cited 2021 Sep 16]. Available from: <https://www.vanityfair.com/culture/2015/08/tinder-hook-up-culture-end-of-dating>.
- [5] Rosenfeld MJ, Thomas RJ, Hausen S. Disintermediating your friends: how online dating in the United States displaces other ways of meeting. Proc Natl Acad Sci U S. 2019;116(36):17753–8. Available from: https://web.stanford.edu/~mrosenfe/Rosenfeld_et_al_Disintermediating_Friends.pdf.
- [6] Vogels EA. 10 facts about Americans and online dating [Internet]. Pew Research Center. 2020 [cited 2020 Feb 6]. Available from: <https://www.pewresearch.org/short-reads/2020/02/06/10-facts-about-americans-and-online-dating/>.
- [7] What is visual hierarchy? [Internet]. Interaction Design Foundation. 2023 [cited 2021 Apr 9]. Available from: <https://www.interaction-design.org/literature/topics/visual-hierarchy>.
- [8] Curry D. Dating app revenue and usage statistics [Internet]. Business of Apps. 2024 [cited 2024 Mar 10]. Available from: <https://www.businessofapps.com/data/dating-app-market>.
- [9] Banks J, Westerman DK, Sharabi LL. A mere holding effect: haptic influences on impression formation through mobile dating apps. Comput Human Behav. 2017;76:303–11. doi: <https://doi.org/10.1016/j.chb.2017.07.013>
- [10] Heino RD, Ellison NB, Gibbs JL. Relationshopping: investigating the market metaphor in online dating. J Soc Pers Relat. 2010;27(4):427–47. doi: <https://doi.org/10.1177/0265407510361614>
- [11] Hobbs M, Owen S, Gerber L. Liquid love? Dating apps, sex, relationships and the digital transformation of intimacy. J Sociol [Internet]. 2017 Sep 5;53(2):271–84. Available from: <https://journals.sagepub.com/doi/full/10.1177/1440783316662718>
- [12] Lawson HM, Leck K. Dynamics of internet dating. Soc Sci Comput Rev. 2006;24(2):189–208. doi: <https://doi.org/10.1177/0894439305283402>
- [13] Miller RS. Intimate relationships. New York: McGraw-Hill; 2015. p. 310.

SUBJECT INDEX

- Abstraction 200
Accentuated Personalities 76
Accessibility 101
Adolescence 106, 150
Affective Processing 125
Anxiety 76, 81
Application Development 230
Aptitude By Treatment Interactions 232
Artificial Intelligence 20, 42, 156, 177, 230
Augmented Reality 230
Autism Spectrum Disorder 161, 171
BBC 143
Body dissatisfaction 138
Body Satisfaction 120
Bullying 225
Character And Application Fidelity 28
Chess players 42
Children 161
Clans 60
Cluster Analyses 120
Cognition 34
Cognitive Behavioral Therapy 150
Cognitive development 42
Cognitive Training 171
Commercial Video Game 171
Computer-Supported Learning 232
Consciousness 45
Consequential Conversations 28
Covid-19 86, 101
Cyber Space 188
Cyberdelics 81
Cyberpsychology 20, 34, 230
Cybersecurity 230
Deception 34
Delight 76
Digital Biomarker 156
Digital Interactions 91
Driving Simulator 113
Eating disorders 138
ECG 213
Elderly 156
Electroencephalography 206
Embodiment 45
Emotion 125, 131, 213
Emotion Regulation 106
Ethical 219
Executive Functions 106
Expectations 138
Extended Cognition 45
FAIR 213
Familiarity 96, 195
Familiarity-Detection 195
Flow 213
Frailty 156
Functional Seizures 131
Game Transfer Phenomena 45
Gamification 237
Gender Identity 143
Gender role 138
Generative AI 200
Generative Images 76
Global Teams 60
Hallucination 81
Hallucination Machine 81
HMDs 113
Human Body Odors 125
Human Interaction 188
Image Generation 200
Immersive Virtual Reality 106
Intelligent Virtual Tutoring Systems 96
Interactive Simulation 166
Internal State Detection 96
Internalizing Symptoms 106
Intervention Development 150
Involuntary Cognitions 45
Large Language Models 177
Machine Learning 96, 131, 156, 200
Memory Assessment 183, 195
Mental Health 71, 177, 225, 237
Misinformation 225
Mixed Reality 232
Motion Sickness 113
Neurogaming 206
News 143
Non-Ordinary States Of Consciousness 200
Onlife 55
Online Counselling 71
Online Dating 237
Online Ethnography 60
Online Relationships 91
Online Violence 225
Open Source 213
Organizational Citizenship 60
Perfectionism 138
Personality 34
Personality Assessment 206
Perspective Taking 161
Persuasion 34
Physiological Monitoring 166
Predictive Processing 45
Preoperative Anxiety 166
Projection System 188
Psychedelic 200

- Psychological Diagnosis 76
Psychological First Aid 86, 101
Psychology 138, 177
Psychomatics 20
Quality 101
Radicalization 55
Radicalized Ecosystems 55
Responsive Virtual Humans 28
Satisfaction 101
Satisfaction With Life Scale 76
Scam Compliance 34
Scenes 195
Self 45
Sentiment Analysis 131
Serious Game 161
Serious Game-Based Learning 219
Social Emotional Learning 232
Social Media 71, 120, 138, 143, 225
Social Stimulation 125
Stem Education 232
Subjectivity 55
Suicide 150
Surgical Outcome 183
Swiping 237
System 0 20
- Technological singularity 42
Telemedicine 183
Telepsychology 86, 101
Third Space 188
Transgender 143
Typical Development 171
UI/UX Design 237
Unity Game Engine 183
User Interface And User Experience 237
Video Game 60, 206, 213
Video Games Effects 45
Virtual Assessment 183
Virtual Environment 125
Virtual Reality 81, 91, 96, 113, 150, 156, 161, 166
Virtual Standardized Patients 177
Virtual Teams 60
Visuospatial Abilities 171
XR Extended Realities 232
Youth 71, 225
Youth Vaccine Education 219
Zone of proximal development 42

AUTHOR INDEX

- ADAMS, 183
ALCAÑIZ, 113, 125, 161, 170
AMON, 70
AMPATZOGLOU, 106
ANONICK, 150
ARGÜERO-FONSECA, 91
ARORA, 166
BERNARDELLI, 81
BLANCHARD, 96
BRIZZI, 81
BURLESON, 131
CAMPBELL, 34, 70, 225
CAPRIOTTI, 156
CARBALLO-MARQUEZ, 106
CASSISTA, 120
CASTILLON, 96
CERVERA-TORRES, 125
CHALENOR, 143
CHAN, 213
CHARTIER, 96
CHICCHI GIGLIOLI, 156
CHOU, 60
CITI, 125
CLEARY, 195
CONTRERAS, 131
CUDO, 44
DAGOSTIN, 213
DE GASPARI, 20, 156, 213
DELGATY, 150
DI LERNIA, 81, 156
DI NATALE, 81,
DRANE, 183
ELGABALAWY, 166
ESCRIBANO-BASTANTE, 113
FERNANDEZ-CAPO, 106
FORSYTH, 70
FRISONE, 81
FROELICH, 237
GAGNON SHAIGETZ, 166
GAGNON-GIROUARD, 120
GALIMBERTI, 55
GAMIZ-SANFELIU, 106
GARCIA-CASANOVAS, 106
GAROLERA, 106
GOODE, 34
GRECO AN., 81
GRECO AL., 125
HEBBARD, 166
HENEGHAN, 150
HERVÁS-ZUÑIGA, 161
HUANG, 206
HUBAL, 27
JOHNSON-GLENBERGA, 232
KARKASHADZE, 76
KATZ, 230
KENNY, 177
LACEY, 34
LEMIEUX, 120
LERNER, 44
LI, 150
LLAMAS-ALONSO, 44
LLOYD, 143
LOGAN, 166
LOMBARDI, 55
LONGONI, 20
LOPEZ-CASTILLO, 170
LUCINI, 55
LYU, 232
MADDALON, 113, 125, 161, 171
MAPLES-KELLER, 166
MARCHIORO, 91
MARIN-MORALES, 125
MARRACCINI, 150
MCMAHAN, 206
MIDDLETON, 150
MINISSI, 131, 125, 161, 171
MONTES-ITURRIZAGA, 86, 101
MOTA, 166
MUTCH, 166
NEISSER, 51
NEWTON, 131
NOKES, 200
OHSUGA, 188
OKADA, 183
OLIVETTI, 156
ORTIZ DE GORTARI, 45
OSORIO, 86, 101
OU, 219
PAEK, 183
PARRA-VARGAS, 171
PARSONS, 161, 177, 219,
PASSALACQUA, 20
PÉREZ-MARAVÍ, 86, 101
PESSOA, 225
PIZZOLI, 81
PONS, 171
PORRAS-GARCIA, 106
PROULX, 166
PUPILLO, 20, 81, 156, 213
RAJEH, 183
RASTELLI, 81
REYNOLDS, 166
RIDOUT, 34, 71, 225
RIVA, 20, 81, 156, 213
ROBERTS, 131
ROCHINA, 113
ROJAS, 106
SAEZ, 171

- SAJNO, 20, 81, 156, 213
SANSONI, 20
SERGEEVA, 76
SERIKOV, 76
SMITH, 166
SOMMER, 166
TANGTON, 206
TURAGA, 200
UBALDI, 81
- VENKATESHA, 96
VERGINE, 55
VLASOV, 76
VOISKOUNSKY, 42
VYAS, 206
WALL, 183
WIEDERHOLD, 20
YEWOOR, 206