In press: ***American Psychologist***

On Conducting Ethically-Sound Psychological Science in the Metaverse

Tracey Cockerton[1], Ying Zhu[2], Mandeep K. Dhami[3]

1. Department of Management, Leadership and Organizations, Faculty of Business and Law, Middlesex University, and the European Centre for the Study of Ethics, Law and Governance in Health Information Technology, U.K. and Ravensbourne University London.
2. Faculty of Management, University of British Columbia-Okanagan campus, Canada
3. Department of Psychology, Faculty of Science and Technology, Middlesex University, London, U.K.

Author Note

Correspondence concerning this article should be addressed to Tracey Cockerton, Department of Management, Leadership and Organizations, Faculty of Business and Law, Middlesex University.

Email: T.Cockerton@mdx.ac.uk

https://osf.io/5tms2/?view_only=e89ee34ba6ad42a1ae7452b043bbb2f4

On Conducting Ethically-Sound Psychological Science in the Metaverse

Abstract

As the next generation of the internet, the metaverse is an immersive three-dimensional (3D) world that incorporates both physical and virtual environments. The metaverse affords numerous advantages for advancing our theoretical and practical understanding of human cognition, emotion, and behavior, as well as shaping our methodological approach to conducting psychological science. However, undertaking research in a world which merges the physical and virtual, also presents new and unique ethical challenges that are not addressed by current ethical guidelines such as the Belmont Report, the American Psychology Association Code of Ethics, and the Association of Internet Researchers Internet Research Ethical Guidelines. We discuss the different domains of the metaverse relevant to psychological research, and consider how three categories of ethical challenges (i.e., 'respect for persons', 'beneficence', and 'justice') may arise when conducting research in the metaverse. We also provide recommendations for addressing these challenges which include reconfiguring existing ethical guidelines as well as creating new ones. Together, these can inform and assist researchers and institutional review boards in making decisions about conducting ethically-sound psychological science in the metaverse.

Keywords: Metaverse, research ethics, virtual reality (VR), digital twins (HDTs), Belmont Report

Public Significance Statement

The metaverse is an immersive three-dimensional (3D) world that merges physical and virtual worlds. It potentially enables psychological scientists to gain a better understanding of human cognition, emotion, and behavior. However, doing research in the metaverse has unique ethical challenges. We identify these and provide recommendations for addressing them so researchers and ethics review boards can make informed decisions about conducting ethically-sound psychological science in the metaverse.

Introduction

Psychological scientists have embraced technological advances (e.g., internet, social media, big data), which have afforded new ways of conducting research (e.g., analyses of 'big data' through machine learning) and offered opportunities for new insights into human cognition, emotion and behavior (Alexander et al., 2020; Ha et al., 2022; Wegrzyn et al., 2017). However, such developments have required Institutional Review Boards (IRBs) to adapt and update their guidance to ensure that research employing new technologies is ethically-sound. We argue that serious consideration of the unique ethical challenges associated with conducting psychological research in the metaverse is now required.

The term metaverse first appeared in the 1992 science fiction novel 'Snow Crash' by Neal Stevenson, and it means beyond universe (i.e., 'meta' = beyond and 'verse' = universe). The metaverse is the next generation of the internet, linking physical reality with digital virtuality (Mystakidis, 2022). The widespread availability of the metaverse and its associated technologies (e.g., virtual reality [VR], augmented reality, digital twins,brain-computer interface, and wearable devices with sensors and geolocation functions are likely to reshape and transform human behavior (for definitions of metaverse-related technology and layers see Cockerton, Zhu & Dhami, 2023).

Although previous technological advances (e.g., facial recognition for emotion research [Wegrzyn et al., 2017]; eye-tracking for attention studies [McCall et al., 2016]; and machine learning for social interactions [Alexander et al., 2020]) have accelerated ethical considerations in conducting research, the metaverse presents unique and new challenges (Ozair, 2022). This is because, unlike other advanced technologies, the metaverse has the ability to combine and integrate these technologies and can collect real time data on all facets of being human (including private and intimate information) without a users' knowledge, in a continuous manner, from users' private spaces. Furthermore, the integration of the aforementioned technologies means that the metaverse affords a significant opportunity to collect a substantial amount of data to craft a holistic picture of an individual's psychological profile, including their mental state, as well as the causes and consequences

of changes in their cognition, emotion, and behavior.

Until now, psychological studies collecting eye-tracking, facial expression, biometric, or even brain wave data, tend to be conducted in a research facility, with data collection occurring for a relatively short and time-limited duration, and with researchers being unlikely to have access to other data (e.g., geolocation of the participant's home, layout of their living room, and their medical conditions as revealed by their body movements). However, this changes when people participate in research in the metaverse – the availability of research data is centralized and is relatively easily obtainable by researchers either through conducting experiments in the metaverse or modeling secondary data downloaded from the metaverse.

Psychology is one of the most regulated behavioral science disciplines. Numerous efforts have been made to prevent unethical research practices and ensure research integrity. However, the unique ethical issues presented by using the metaverse as a research laboratory necessitate urgent consideration. We take a systematic and holistic approach to examine these ethical issues. Specifically, we identify key ethical challenges that may arise when conducting psychological research within specific domains of the metaverse, and provide recommendations to IRBs and psychological scientists for addressing these challenges. Our discussion is organized around the Belmont Report's three ethical principles (e.g., Respect for persons, Beneficence, and Justice) and draws on the American Psychological Association's (APA) ethics code (APA 2017) as well as the Association of Internet Researchers (AoIR) internet research ethics guidance (Franzke et al., 2020). First, we begin by highlighting the opportunities the metaverse affords psychological scientists.

Opportunities for Psychological Science in the Metaverse

Immersive virtual environment technology has been used in psychological research for some time (e.g., Loomis et al., 1999; Gilbert, 2011), however, the metaverse affords psychological scientists' greater opportunities to fully realize the research potential of such technologies. Indeed, widescale accessibility of the metaverse may drive a paradigm shift away from traditional approaches to conducting psychological research, which have often constrained the equity, diversity

and inclusivity of participants, researchers, and research in several ways.

First, many previous studies have been conducted on a narrow demographic of the human population, namely Western, educated, industrialized, rich and democratic (WEIRD) samples, thus leading to a potentially limited (or distorted) view of at least some aspects of human psychology (Henrich et al., 2010). The metaverse, by contrast, has the potential to be more inclusive (Zallio & Clarkson, 2022). Researchers have access to a wider pool of participants worldwide (e.g., augmented reality or VR allow participants to overcome constraints of physical distance and space), enabling the study of more diverse samples (e.g., research with dementia patients, [Coelho et al., 2020]; individuals with disabilities [Stendal et al., 2011]). In addition, the participant pool can extend beyond physical locations to 3D virtual worlds and cyberspace (Dwivedi et al., 2022) and to digital representations of real people (e.g., human digital twins and avatars).

Second, psychological research typically requires experimental control and manipulation of one or more variables, thus reducing its external validity and limiting the generalizability of its findings, as well as potentially threatening its internal validity (Dhami et al., 2004). Immersive VR environments offer researchers control over multiple factors, standardizing stimuli across participants (Loomis et al., 1999), while allowing flexibility in observational research (Mawer, 2016), and providing enhanced external validity in the clinical, affective, and social neurosciences (Parsons, 2015). Such technology allows researchers to subtly manipulate characteristics of the 'physical' and social environment (e.g., visual cues to prime participants and a person's race; Maister et al., 2013) to investigate influences on human cognition, emotion, and behavior. This allows for replication of well-known findings (e.g., Slater et al.'s [2006] VR simulation of the Milgram's experiments) as well as original research that would be difficult or impossible to conduct in the physical world (e.g., body changes in VR research, [Slater et al., 2020]).

Third, the metaverse offers a relatively low cost, open, shared, accessible and persistent research environment that can facilitate multi-disciplinary, collaborative research (Gürerk et al., 2019). For instance, in an 8-week study of human interactions (via avatars), Han et al. (2023) created

virtual environments in the metaverse and found that people moved more synchronously and felt more restored, present, and realistic in spacious environments than constrained ones, with preferences for outdoor virtual environments over indoor ones. Also, the metaverse provides a real-time research environment where researchers do not need to be in the same physical space as others but can remotely embody an avatar and interact with participants and colleagues (Bombari et al., 2015), thus enabling the diversification of both research participants and research teams (e.g., different disciplines, countries, and cultures).

Fourth, researchers may apply data analytics to vast amounts of data from metaverse platforms such as Roblox, which use artificial intelligence (AI) and real-time data collection to generate datasets (Adjerid & Kelley, 2018). These data can also be user-generated content from a range of biometric, social media, and brain-computer interface applications, thus opening up further avenues for passive and secondary data analysis. Indeed, studies utilizing agent-based modeling for hypothesis testing and theory building (Madsen et al., 2019) and capturing cognitive aspects of creative human decision-making as part of a Pi-Mind ('Patented intelligence') are already underway (Terziyan et al., 2018) and informing the development of digital cognitive clones that can be used as an individual's responsible representative when the human is unavailable (Golovianko et al., 2021).

Given the research opportunities afforded by the metaverse, which should arguably be exploited by psychological scientists, we believe there is an urgent need to understand the unique ethical challenges it presents. In the next section, we define main features of the metaverse and the domains most relevant to psychological research, followed by the Belmont Report's ethical principles. Then, we present a research ethics framework that identifies key ethical issues within specific domains of the metaverse, focusing on issues unique to the distinctive characteristics of the metaverse, and we provide recommendations for addressing these issues (see Cockerton et al, 2023 for a summary).

The Metaverse and Ethical Psychological Research

Despite no agreed upon definition of the metaverse (Sun et al., 2022), Mystakidis (2022)

provides an overview and analysis of key metaverse-related concepts and discusses contemporary developments and innovative applications. Early prototypes of the metaverse include Second Life (2003), Roblox (2006), Minecraft (2011), and Decentraland (2017). These virtual worlds include different representative features of what may be considered key components of the metaverse as it currently stands (Duan et al., 2021). For example, initial concepts of immersive virtual environments (Gilbert, 2011) specified requirements such as a 3D graphical interface, multi-user remote interactivity that is persistent (allowing the virtual environment to continue operating after users disconnect), immersive (whereby users experience a sense of psychological presence in the digital environment), and allowing users to define their own identities, activities, and goals.

More recently, numerous commentators have attempted to define layers and characteristics of the metaverse (see Dwivedi et al., 2022; Park & Kim, 2022). They generally agree on seven layers (Setiawan & Anthony, 2022) that represent the technologies needed to develop the metaverse and the experience people expect to have in every virtual world. The seven layers from top to bottom include: (1) Experience (e.g., e-sports, games), (2) Discovery (e.g., searching tools, social curation), (3) Creator economy (e.g., commerce, design tools), (4) Spatial computing (e.g., combining real and virtual spaces and using sensors such as visual, touch and sound to create realistic virtual environments), (5) Decentralization (e.g., decisions on rules of behavior and management of virtual assets etc. in the metaverse are collective - involving users instead of a single entity), (6) Human interface (e.g., haptics enabled wearables, VR goggles), and (7) Infrastructure (e.g., 6G/sixth generation wireless).

The seven-layer model, however, obscures the domains of the metaverse most relevant to psychological research. These domains are: (1) digitization (i.e., digital information from users); (2) data integration (i.e., combining information about users from all devices), (3) virtualization (i.e., representations of users and their environments); (4) social networks (i.e., information between users), and (5) operating in virtual worlds (referring to information from human beings and the digital entities they created). We focus on the ethical issues that are distinctive to these five domains

(which we define further below) in order to raise awareness and provide guidance to psychological scientists and IRBs.

Domains of the metaverse relevant to psychological research

Digitization. As an immersive digital environment, the metaverse relies on the process of transforming objects, audio, text, images, and other data into a cyber, hyper-realistic simulated world (Zallio & Clarkson, 2022). It also includes digitizing interactions between people, as well as between individuals and organizations through augmented reality, VR, and brain-computer interface enabled digital platforms. Digitalization penetrates our socio-cultural world as the technology nestles itself in us (e.g., through brain implants), between us (through social media like Facebook), knows increasingly more about us (via big data, machine learning, and techniques such as emotion recognition), and is continually learning to behave more like us (AI enabled robots exhibit intelligent behavior and can mimic emotions; Royakkers et al., 2018). The digitization of everyday information into computer readable form creates data for research and other purposes.

Data integration. The metaverse infrastructure relies on advanced cloud computing services and technologies leading to the availability of big data platforms (Sun et al., 2022). Data collected and integrated from the metaverse goes far beyond traditional forms (e.g., demographic, behavioral, relationship, and usage data), to data collected via various devices which is integrated in a central location. For example, data on head movement and brain wave data from VR goggles, eye-tracking data from augmented reality glasses, hand and body movement from a VR controller, haptic feedback, and biometric data (e.g., body temperature) from haptic vests and gloves, and geolocation data from computer or smartphone, can be integrated into one database. This data can be pooled for remote access and analysis, for example, allowing novel insights into holistic and personalized health (McKeown et al., 2021).

Virtualization. Virtualization can allow the self, others, and objects to be represented. Users can experience a sense of “being there” (presence), with other human representations as “social entities” (social presence), as well as “behavioral realism” (supporting natural interaction with

objects and others indistinguishable from the real world; [Kort, 2003, p. 361]). VR enables experimental manipulation of any variable, and behavior (e.g., head, hand, and body movements) can be measured in a fine-grained manner, covertly and continuously (Yaremych & Persky, 2019). Psychological research using VR has to-date fostered clinical applications (e.g., exposure therapy, Marloth et al., 2020). Augmented reality adds virtual features to the real environment, and together virtual and augmented reality (i.e., mixed reality or extended reality) systems provide an enhanced sense of immersive realism (Slater et al., 2020), which can benefit experimental research.

Social networks. The metaverse provides a multisensory, multi-user persistent digital world with interoperability of human digital twins, virtual humans and avatar personalization (Mystakidis, 2022). Thus, social networks in the metaverse enable a broader and richer range of social interactions (e.g., 3D enabled interactions with haptic feedback) that resemble in-person experiences more than the current two-dimensional internet applications (e.g., social media) offering insights into the feelings of social connectedness of users of augmented reality (Miller et al.'s,2019).

Virtual worlds. This domain of the metaverse constitutes a digital platform that combines all advanced technologies such as augmented reality, VR and sensor-enabled wearables. Among them, VR applications provide an integrated network of persistent online computer-generated simulated environments where multiple users in remote physical locations can interact in real time (Dionisio et al., 2013). Virtual worlds provide an immersive experience and a compelling sense of presence (Loomis et al., 1999), inhabited by avatars controlled by their real-world users (Sadler, 2017). Engagement in virtual worlds is predicted to dramatically reshape our conception and experience of self (Gilbert & Forney, 2013), providing insights for social and cultural innovation (Vanacker & Heider, 2012). Psychological research in virtual worlds has already commenced, for example, suggesting that co-morbidity between internet addiction and various forms of real-world addictions may apply to 3D virtual worlds (see Gilbert et al., 2013).

The Belmont Report's ethical principles applied to psychological research in the metaverse

Principle 1: Respect for persons. Respect for persons includes treating individuals as autonomous agents and protecting those with diminished autonomy. This is recognized in the Belmont Report through the application of informed consent procedures which specify processes for protecting participants' anonymity, confidentiality and privacy. This principle is expanded in both the APA ethics code (2017), in terms of respecting people's rights and dignity, and in the AoIR internet research ethics guidance (Franzke et al., 2020), in terms of 'respect for users' in online contexts. Psychological research relating to specific domains of the metaverse faces several unique challenges in order to abide by the principle of respect for persons, which we detail later.

Principle 2: Beneficence. Beneficence requires that research participants are protected from harm and that efforts are made to secure their well-being. This obligation is reflected in the APA (2017) General Principle A of beneficence and nonmaleficence as well as in the AoIR guidance (Franzke et al., 2020). These guidelines draw researchers' attention to understanding their participants and their own responsibility to protect those who may be more vulnerable than others (e.g., minors, women, minorities, those suffering emotional states, and LGBTQ individuals etc.). The principle of beneficence is demonstrated when determining how to assess and mitigate the probability and magnitude of potential harm. Several unique ethical challenges to this principle arise when using the metaverse for psychological research, which we discuss later.

Principle 3: Justice. The principle of justice refers to fairness in the distribution of burdens and benefits of research. It may be formulated in five accepted ways to each person according to an equal share, individual need, individual effort, societal contribution, and individual merit. This principle is reflected in the AoIR guidance (Franzke et al., 2020) and is expanded in the APA ethics code (2017) to prevent unjust practices derived from psychologists' potential biases, the boundaries of their competence, and the limitations of their expertise. The metaverse may support researchers in enhancing equality, diversity and inclusivity (EDI) in psychological research. However, it also presents new ethical issues related to the principle of justice as outlined in the next section.

Ethical Challenges for Psychological Research in Metaverse Domains

Here, we discuss the novel ethical challenges of conducting psychological research within each of the five relevant metaverse domains (i.e., digitization, data integration, virtualization, social networks, and operating in virtual worlds). These ethical issues, along with proposed solutions, are discussed according to the Belmont principles of respect for persons, beneficence, and justice.

Digitization and Ethical Research

When entering the metaverse, each user's interactions with the virtual world results in data (Sun et al., 2022) that could be used for research (Gilbert, 2011), although unique ethical issues arise when doing so (Royakkers et al., 2018). Once individuals' biometrics (e.g., facial, voice, iris recognition) are digitized, this becomes a permanent record within databases that can be used for comparison and matching purposes (Royakkers et al., 2018). The development of the brain-computer interface and its integration into VR goggles provides information on the frequency and magnitude of a participant's brain waves (Ha et al., 2022). Thus, the digitization of personal and sensitive information raises privacy concerns and has implications for compliance with the principle of respect for persons. The fact that the metaverse enables collection of continuous real-time data from users also raises concerns about both a new concept of mental privacy (Wajnerman Paz, 2021) and the existing issue of an individual's identity being easily revealed by the digitalization of comprehensive data. Furthermore, dynamic and real-time data collection could result in a consistent invasion of participants' privacy, at least for the duration of their time in the metaverse.

The risk of revealing participants' identities may expose them to harm from sophisticated identity fraud involving health, biometric, and brain data at a level that has not been encountered before in psychological research (Winston & Hemanth, 2019). This has implications for compliance with the principle of beneficence i.e., biometric data are biologically unique to each individual, so once compromised cannot be redressed, and the individual remains at risk of identity fraud. In addition, the application of biometrics may result in misclassifications, exploitation and discrimination (Royakkers et al., 2018), and unjust exclusion may arise due to unfair and inequitable

access to digital technologies and potential interventions (e.g., in health, Brall et al., 2019). This raises ethical issues with regard to the principle of justice.

Proposed solutions. A greater understanding of the ethical pitfalls of digitization of our material, biological and socio-cultural worlds is needed. Researchers and IRBs must ensure accurate informed consent (Royakkers et al., 2018) and increased data security. This requires actions including training and expert guidance for members of the research team and IRBs, with increased ethical scrutiny to identify and remedy violations of justice, as well as new models of informed consent while retaining voluntary participation, not undermined by incentives. These consent procedures may include: (i) 'blanket' agreement in advance of future analyses; (ii) 'opt-out' consent where subjects actively withdraw consent; (iii) 'waiver' consent where the research involves no more than minimal risk to subjects, could not be carried out practically without the waiver or an alteration to consent); (iv) 'no consent' (i.e., where the rights of subjects are not infringed by the research and obtaining consent is impractical); (v) 'broad', 'open or categorical consent' (i.e., where consent is sought for a range of uses but not assumed for all purposes and in constrained; McKeown et al., 2021); (vi) 'dynamic or opt-in consent' (i.e., where consent is sought on a case-by-case basis; McKeown et al., 2021); and (vii) a 'meta' consent model in which individuals choose how they prefer to provide consent (e.g., open, opt-in etc.; McKeown et al., 2021). Indeed, different approaches to consent may ensure and promote autonomy for everyone in line with fair data use (Brall et al., 2019). IRBs should also promote the minimal data guideline i.e., only collect or access data that is relevant to the proposed research.

Data Integration and Ethical Research

Research using large volumes of data has already increased in psychology and other fields (Leonelli, 2020). Such 'big data research' is likely to surge with to the availability of multi-modal measures through the advanced technological devices associated with metaverse (Sun et al., 2022). When biometric data is included in research, it may include information from individuals that they are unaware of, such as indicators of their health risks (e.g., an iris scan can determine diabetes)

with unforeseen consequences (e.g., access health insurance; Royakkers et al., 2018). Digital information that includes biometric data, integrated with big data and machine learning algorithms, may be used for tracking, invasion of mental privacy and pervasive monitoring, as well as user consent for a "God View" of individuals' lives (Royakkers et al., 2018). This has ethical implications in relation to the principle of respect for persons.

The metaverse's unique ability to automatically collect and generate comprehensive big data about participants and apply machine learning and AI on this data adds complexity, uncertainty and unpredictability to the potential harms and benefits of conducting research in the metaverse. Potential harm may arise from the automatic processing and integration of audio, visual, body movement, other biometric and geolocation data during research that researchers and users may or may not expect, be aware of, or have access to (Wang et al., 2022). This has implications for compliance with the principle of beneficence.

Explaining data ownership and processing (e.g., storage and sharing) in reasonably understandable language during the informed consent procedure to comply with the principle of respect for persons, may challenge researchers as well as the existing regulatory and ethical governance tools, and current practices and expertise of IRBs (APA, 2017; Ferretti et al., 2021; Franzke et al., 2020). Researchers may also find it difficult to specify responsibility for AI models, how such models operate, what pre-determined principles were used in their development, as well as the implications of AI-generated comprehensive datasets (e.g., including biometric, geolocation, real-time data across multiple metaverse platforms) in respecting privacy and avoiding bias embedded in the algorithms (Franzke et al., 2020). However, the potential negative implications for participants and the unforeseen research gains of complex big data may be very difficult for researchers to fully anticipate (McKeown et al., 2021) and then communicate clearly to potential participants.

A further issue relating to data integration is the risk of harm from data breaches, third-party processing and disclosures, which is uniquely different from traditional psychological laboratories,

since data ownership may be distributed across technology providers and metaverse platforms. The harmful consequences for some participants (e.g., those whose actual or perceived sexual orientation and/or gender identity is lesbian, gay, bisexual, transgender, intersex or queer [LGBTQ]) from specific jurisdictions where certain identities and behaviors are deemed illegal, has implications the principle of beneficence. The distribution of burdens and benefits of research may not be fair, thus also having implications regarding the principle of justice.

Proposed solutions. Ethical issues related to data integration highlight the need for "new thinking about consent" (McKeown et al., 2021, p. 3) when conducting research in the metaverse. In addition to the consent-based solutions proposed above, here ongoing consent (Franzke et al., 2020) may be helpful for researchers to clarify limits to confidentiality in a dynamic environment at different stages of the research. Identifying threats to data security is also important in defining limits to confidentiality, particularly in relation to different legal jurisdictions.

The regulation of big data and transparency of algorithms used within integrated metaverse data platforms may support researchers in providing accurate information for informed consent (Royakkers et al., 2018) to meet the requirements of respect for persons. In addition, IRBs ought to require researchers to clarify how they plan to identify, monitor and act in order to address issues relating to equality, diversity and inclusion in AI decision-making in order to avoid harm and comply with the principle of justice. The composition of IRBs would benefit from expertise in advanced technologies relevant to the metaverse and psychological research (e.g., big data processing of eye tracking, facial expressions, heart rate, pulse, breathing rates, blood pressure, electroencephalography [EEG], biometric and brain-computer interface data).

The role of IRBs may be further assisted by developing risk assessment tools for ethical decision-making concerning big data studies (Ferretti et al., 2021). Indeed, given that such tools need to be regularly updated, IRBs may play an important role in facilitating collaboration between researchers and the tech industry in developing integrated ethical frameworks and guidance.

Virtualization and Ethical Research

The fast pace of metaverse developments has resulted in a lack of understanding of the long-term effects of VR exposure (Madary & Metzinger, 2016; Kellmeyer et al., 2019). There are many potential negative effects of immersive experiences that are not fully understood. For example, what are the physical and psychological effects of transitioning into/out of the VR environment and in-experience user control (Han et al., 2022b)? What are the psychological impacts and potential detrimental effects on users' health and well-being that may result from self-indulgent escapism in VR experiences (Han et al., 2022a)? What is the potential addictive nature and temporary sense of isolation from VR experiences (Merkx & Nawijn, 2021)? And, are individuals able to retain a distinction between virtual and physical worlds (Weber et al., 2021)?

There is also a need to understand the combined effects of virtual and augmented reality, and extended reality exposure on both general and specific populations, such as individuals who are vulnerable to disturbed perceptions of reality (Marloth et al., 2020). Patients (e.g., individuals with anxiety disorders, paranoia, or dementia) may be at increased risk of 'deceptive illusions' without the capacity to exercise informed consent (Kellmeyer et al., 2019). VR also has the potential to induce 'virtually real trauma' with researchers' focus on external validity achieved through 'context realism' (Ramirez, 2019) that may expose participants to psychological risks of harm. Indeed, while a desirable feature of VR is the immersive feeling of 'being there' (or presence), cybersickness (i.e., motion sickness) is a well-known side-effect, which is inversely related to an individual's sense of presence (Marloth et al., 2020; Weech et al., 2019). These examples raise ethical concerns related to the principles of respect for persons and beneficence.

A specific issue relating to VR and representation in the metaverse is the illusion of embodiment created when users are exposed to simulations such as virtual body ownership and when they experience a sense of loss of control (Kellmeyer et al., 2019) or feel that they are not the author of their own actions (Marloth et al., 2020). Psychological research has investigated the impact of virtual embodiment on implicit social cognitions related to age (Banakou et al., 2016), body shape and size (Preston & Ehrsson, 2014), gender groups (Slater et al., 2020), and race (Maister

et al., 2013; see Parsons, 2015 for a review). Relatedly, the 'Proteus effect' (Yee & Bailenson, 2007) refers to when users behave in ways consistent with the negative characteristics of their avatar (Ratan et al., 2020). Although this research area may develop with increased access to the metaverse, the illusion of embodiment raises ethical concerns related to the principle of beneficence.

Proposed solutions. Madary and Metzinger (2016) recommend that informed consent for VR experiments includes an explicit statement that their immersive nature can have lasting behavioral consequences for participants, some of which are presently unknown. We encourage psychologists using VR (e.g., psychotherapists use it to treat patients; for review see Cieślik et al., 2020) to share their informed consent procedures so that these can be constantly improved. IRBs are also encouraged to consult researchers, within and outside their own institutions, who have conducted VR studies. We recognize that this may at times be infeasible and highlight that it does not absolve researchers from their responsibility to protect research participants. It is essential to carefully screen participants in order to minimize the risks of aggravating an existing psychological disorder or an undetected psychiatric vulnerability (Gregg & Tarrier, 2007) such that exclusion criteria specific to possible risks posed by VR are specified. VR should be used with caution when individuals have reduced capacity for reality testing, such as those who suffer from dementia or experience symptoms of psychosis (Kellmeyer et al., 2019; Marloth et al., 2020).

The Equivalence Principle may also provide a useful ethical guide. It states that "if it would be wrong to allow subjects to have a certain experience in reality, then it would be wrong to allow subjects to have that experience in a virtually real setting" (Ramirez, 2019, p. 226). This principle is a counterargument to the use of VR to investigate the psychological effects of real-world experiences which would be unethical to (re-)create physically (e.g., experiments in obedience; Slater et al. [2006]). Debates continue regarding which situations are permissible for individuals to be exposed to (e.g., understanding people's psychological responses to natural disasters). Virtual environments can be easily manipulated, allowing modifications to be introduced in real-time as direct reactions to

user behavior (Marloth et al., 2020). Thus, researchers could adapt research procedures and protect users should unexpected negative experiences arise.

Risk assessments that consider users of all ages and abilities who will have access to the metaverse are also required, along with avatar proteus effect assessments (Yee & Bailenson, 2007) in order to measure the extent to which an avatar-self is equivalent to the real-self it represents (Hendaoui et al., 2008). This is because close equivalence of an avatar to a research participant presents further risks to those who are vulnerable to disturbed perceptions of reality, 'virtually real trauma' and/or changed self-perceptions (Marloth et al., 2020). Finally, effective reporting procedures and tools need to be developed to facilitate safe research practices (Franzke et al., 2020).

Social Network and Ethical Research

The social network domain of the metaverse presents researcherswith numerous virtual communities to study. While online social media research is not new (Gosling et al., 2011) nor discussions of the ethical issues raised by such research (Arigo et al., 2018; Willis, 2019), the metaverse combines these with features of real-world field research in a simulated virtual environment. Thus, the ethical issues related to this domain of the metaverse build on social media research and real-world field research, along with additional complexities related to the digitization of information, data integration, and VR technology (as discussed above).

Sadler (2017) suggests that when conducting online field research, researchers should consider the level of risk associated with group accessibility, public versus private spaces, and perceived privacy. For example, groups with very low accessibility and a high perception of privacy would be considered high-risk, such as a support group for survivors of sexual abuse in Second Life (Sadler, 2017). In the offline world and traditional internet, researchers can focus solely on participants who consent, whereas in the metaverse, all communications and conduct are processed digitally. This raises challenges for researchers if they attempt to exclude users who refuse consent or if they include bystanders (Nebeker et al., 2017). Including metaverse users who have not given

their consent or do not fully understand their consent is a risk to their privacy and the principle of respect for persons. Even unintentional breaches of privacy due to the default setting of metaverse platforms may damage the reputation of research communities and trust in researchers (Fairfield, 2012) as well as undermine the value of research involving the metaverse.

The ability to target specific groups based on personal information and justifiable inclusion and exclusion criteria is also challenging when undertaking research in the metaverse. Currently, little guidance is available for researchers or IRBs on how to do this ethically, i.e., by maintaining the principles of justice, respect for privacy and investigator transparency (Parsons, 2019). Researchers may find it difficult to enhance the equality, diversity and inclusivity of data collection in the metaverse to avoid biases and social exclusion of different groups, and abide by the principle of justice in ensuring a fair distribution of burdens and benefits of research. Researchers may need to consider providing access to the metaverse for hard-to-reach groups while remaining accountable for safeguarding users who may be less experienced or motivated to engage in a virtual world study that they perceive as having no benefit for them.

Providing users with the option to refuse consent or withdraw from social network research without consequences may be unachievable in a virtual world. However, procedures that do not allow participants to withdraw consent for future uses of their personal data may not comply with the principle of respect for persons. McKeown et al. (2021), on the other hand, argue that the ethical basis for data reuse in research may outweigh an individual's preferences, particularly in health research and the probable health gains. They argue that acceptance of a change in research norms requires public engagement, trustworthy institutional data governance, and honesty regarding the pros and cons of data reuse.

Proposed solutions. Psychological research in the social network domain of the metaverse requires clarity and transparency to "avoid deception and refrain from [researchers] fabricating online identities to gain access to … online communities" and "proactively disclose their presence" when collecting information (Gelinas et al., 2017, p. 7; word in brackets added). Researchers need to

carefully plan consent procedures that facilitate refusal and withdrawal without penalty. This requires evaluating user involvement and potential risks to bystanders and virtual communities (Sadler, 2017). Researchers may benefit from collaborating with platform and application designers and other professionals to develop a globally recognized framework that respects metaverse users' privacy (Wajnerman Paz, 2021). Researchers may also need to provide guidelines for appropriate behavior, and monitor conduct and communications as part of the research to safeguard users (Franzke et al., 2020).

While guidance on justifiable inclusion and exclusion criteria may be helpful, the principle of justice requires researchers to consider the representation of users and/or their avatars across dimensions such as gender, age, race, ethnicity, disability and religion, as part of their participant recruitment procedures. In time, the metaverse may provide a more inclusive approach to human research participation with access to participants who once may have been excluded or neglected. Different ethnic groups, minors, and users with physical or learning disabilities or with health conditions who may have been over-protected from participating in research previously could now be accommodated and appropriately supported by customizable research settings and the adoption of advanced technological devices. This can only occur, however, once access to the metaverse is in place and the potential benefits of the research are clarified.

When assessing participant recruitment, IRBs ought to consider requirements for researchers to undertake an equality, diversity and inclusivity impact assessment of their planned research in the metaverse. This would require clarifying inclusion and exclusion criteria to avoid undue reliance on or exclusion of specific groups. A transparent account of who is involved in the research and who may benefit, as well as identifying potential adverse effects on different groups, and acceptability of devices to diverse communities, may reduce researchers' reliance on opportunity samples of users and inferred characteristics of users. The outcome of equality, diversity and inclusivity assessments may also facilitate academic-industry partnerships in order to address inequalities, although any conflicts of interest with industry partners would need to be declared

(Schoentgen & Wilkinson, 2021).

Virtual Worlds and Ethical Research

In the virtual worlds domain of the metaverse, avatars are an evolving technological development (Aljaroodi et al., 2019) that simulate human communication and maintain social relationships (Dwivedi et al., 2022). They may be autonomous and semi-autonomous (Lawler-Dormer, 2018) or with 'perceived agency' (Nowak & Fox, 2018), acting independently of their user/owner. Metaverse users can be represented by one or more of these artificial-human entities, which can be customized in appearance and character. Avatars may represent the identity of their user, often with the user known by the avatar's assigned name, and are also becoming more life-like – transcending the 'uncanny valley' (Seymour et al., 2018), thus also becoming more personally identifiable (Fairfield, 2012). While the Belmont Report does not cover harm to non-humans, Harris and Anthis (2021) discuss the possibility of artificial entities with the capacity for positive and negative experiences (i.e., sentience) being created which may become an important area of psychological research.

Individuals may identify with their avatar (Vanacker & Heider, 2012) and have a moral attachment to the avatar (Wolfendale, 2007). The implication of human-avatar identity involvement and attachment represents a risk of psychological harm to the human behind the avatar (Vanacker & Heider, 2012). With the advance in augmented reality and VR technologies, the heightened sensory experience in the metaverse may mean that the impact and harm of deviant behaviors on victims in virtual worlds not only feels real but is magnified (Dwivedi et al., 2022). The fact that the user may not be the owner of their avatar presents issues regarding rights and accountability, particularly since human digital twins or clones can learn to make decisions under difficult conditions (i.e., from "*confrontation with a strong, constantly evolving, hardly predictable artificial adversary*") (Golovianko et al., 2021, p. 181). As a research environment, the metaverse poses new challenges regarding the principle of beneficence. Without policies and policing in place to support safeguarding against abuse and grooming, the metaverse may be a high-risk research environment

for users who may be considered marginalized in the gaming and virtual worlds contexts (e.g., women, LGBTQ, and ethnic minorities) (Freeman et al., 2022). Maltreatment is more commonplace than in the real world (Kavanagh et al., 2020) as demonstrated by research on 'embodied harassment' in social 3D VR worlds (Freeman et al., 2022).

Engagement in the digital world has also been found to reduce inhibition (Crowell et al., 2008), meaning that virtual world experiences may include more risky behaviors, self-disclosures, anti-social behavior, cyberbullying, and traumatic experiences, than in the real world (Mystakidis, 2022). While harassment, assault, bullying and hate speech are commonplace in VR games (Kowert & Cook, 2022), there are few mechanisms (e.g., in-game reporting tools) by which such behavior can be reported by users or researchers. Additionally, 'deepfakes' and exposing users to advanced technological devices (e.g., collecting real-time biometric data with geolocation enabled) could result in blackmail, ideological influencing, sabotage and intimidation (De Ruiter, 2021).

Thus, the metaverse presents researchers with challenges regarding their social responsibilities to be alert to the possible unexpected harmful consequences of their research and to always support participants within the limits of their professional competence. However, some potential hazards may be beyond researchers' awareness, competence, and control, such as the misuse of metaverse users' biometric profiles feeding biases into AI algorithms (Mystakidis, 2022). For instance, biases about emotions based on gender or age may be embedded in affective brain-computer interface technology. This could lead to monitoring affective states according to what is considered affectively 'normal' with neurofeedback applied to some individuals to self-regulate their emotions accordingly (see Steinert & Friedrich, 2020). The consequences of this have implications for the principles of beneficence and justice.

Proposed Solutions. IRBs ought to require clarity from researchers regarding the rights and accountability of digital actors in the metaverse. This level of clarity is likely to extend to developers and owners of virtual worlds as well as their occupants, and may require expertise from a range of fields, including computing, law and 'technoself' studies (Luppicini, 2013). IRBs may therefore

benefit from expanding their membership to reflect the multi-disciplinary expertise required to consider issues such as the nature of autonomous agents.

Researchers will also need to remain aware of their social responsibilities when undertaking research in a multi-disciplinary virtual environment that is constantly evolving, in order to serve the best interests of metaverse users (i.e., research participants). To this end, researchers may need to consult with, refer to, and collaborate with other professionals, such as computer scientists and the tech industry. This may include informing advanced computing programming and blockchain technology from a psychological perspective, testing AI models and machine learning algorithms involved in decision-making to detect and avoid biases, and working to develop regulatory tools (He, 2022). IRBs and researchers could work with the tech industry to develop international level research ethics guidance at pace to support ethical research in the metaverse.

IRBs should also require clarification from researchers regarding their competence, training and experience in conducting research in the metaverse, especially where a medium- to high-risk of harm is present. In most cases, it may be sufficient for researchers to clarify how participants will be protected from potentially harmful content or behaviors, and to limit the duration of the research in order to avoid long-term harm and over-immersion in virtual environments. However, this presupposes that a valid and reliable approach to risk/benefit analysis exists. Additionally, as Friesen et al. (2017) note, there may be different interpretations of risk/benefit analyses from the perspective of the IRB, researchers and participants. Indeed, while lessons from research on AI and virtual and augmented reality (Finnegan et al., 2021) can be applied to the metaverse to avoid risks of harm, the metaverse poses new and more challenging issues. It may take time to build the evidence base identifying potential harm associated with specific types of psychological research in the metaverse (e.g., experimental studies using brain-computer interface enabled VR goggles, qualitative research on users' communications using passive data).

Researchers should engage in extensive piloting to test each new procedure prior to a full-scale launch of a study in the metaverse. They also need to utilize their knowledge of experimental

psychology and the results of research on VR in order to develop pre-screening tools (Madary & Metzinger, 2016). A key learning point from current research on neurological injuries resulting from VR technology use is for increased awareness about physical risks to inform users and professionals (Warner & Teo, 2021). The potential psychological effects of the metaverse also require similar awareness-raising. In this respect, it is essential that researchers disseminate their findings, even from pilot studies, regarding the physical and psychological impact of the metaverse.

Conclusion

The metaverse has embraced cutting-edge technologies such as brain-computer interface, human digital twins, sensor-enabled wearables, augmented reality, VR and machine learning, to merge physical and digital worlds in order to provide users with a 3D immersive real time experience. As a psychology laboratory, the metaverse offers many advantages given its highly controllable and simulated nature, affording internally and externally valid research, and multiple channels of data for multi-variate analyses at a relatively low cost. The metaverse also has the potential to increase the demographic representativeness of psychological research on a global scale, accelerate advancements in psychological theorizing, and transform psychological methodologies. Indeed, psychological research within and about the metaverse can play a significant role in the development of the virtual universe itself, thus propelling the discipline of psychology.

However, generalizable advancements, to some extent, require equality, diversity and inclusivity in access to digital technology. This relates to the principle of justice given it refers to the digital inequality that exists on a global scale. Even in developed nations, digital disparities often exist among those with lower incomes and/or education levels, disabilities, ethnic minorities, women, older people and those living in rural areas (Parsons, 2019). Access to technology is essential to support ethical psychological research in terms of equality, diversity and inclusivity.

We have discussed the ethical challenges associated with conducting psychological research in the metaverse that are likely to stretch the expertise and ethical reasoning of researchers and IRBs. As a multi-disciplinary development, the metaverse relies on accumulating knowledge from

psychological science. Thus, there will be times when the benefits of psychological research in the metaverse may be significant for the technology industry while the burden of risk is on users (including research participants) and researchers. For example, research on the use of human digital twins in running massive simulations across multiple domains to predict adversaries' behaviors and tactics in order to create proactive cyber defense strategies (Nguyen, 2022) as well as research on the negative effects of VR escapism (Han et al., 2022a) may each have immediate benefits for the technology industry, but not for users.

IRBs may need to assess the nature and appropriateness of psychological research in the metaverse in terms of burdens and benefits at the individual, group, and societal levels. Schoentgen and Wilkinson (2021) present a framework for assessing ethical concerns in digital technologies that place human and societal impact at the center of technology design, purpose, and usage. Building on the AoIR internet research ethics guidance (Franzke et al., 2020), researchers ought to consider issues relating to different stages of research and methods in the metaverse (e.g., transparency versus vast datasets to reduce bias, versus more privacy and control for research participants).

Greater international collaboration between researchers and the tech industry is encouraged in order to collectively and effectively develop responsible research and innovation guidelines. The recent 'Ethically Aligned Design' from The Institute of Electrical and Electronics Engineers (IEEE) (Ethically Aligned Design First Edition [EADe1], 2019) has made significant progress in this regard as it integrates ethical considerations into the context of software and interactive systems design, prioritizing human well-being. However, such guidelines will need to be integrated into existing psychological research ethics guidance so that psychological scientists can conduct ethically-sound research in the metaverse. In addition, continuing professional development of researchers is required to meet the complexities of undertaking research in the metaverse. We, amongst others (Maloney et al., 2021), recognize the importance and urgency of this endeavor.

Using the metaverse as a psychology laboratory will test the boundaries of existing research ethics policies and reveal policy areas that need to be revisited, as well as highlight important policy

gaps that need to be filled. The fast pace of technological innovation, coupled with peoples' eagerness to engage with new technologies, against a backdrop of new cohorts of tech-savvy psychological scientists, mean that policies and practices for ethical research need to be urgently updated.

References

Adjerid, I., & Kelley, K. (2018). Big data in psychology: A framework for research advancement. *American Psychologist, 73*(7), 899-917. https://doi.org/10.1037/amp0000190

Alexander III, L., Mulfinger, E., & Oswald, F. L. (2020). Using big data and machine learning in personality measurement: Opportunities and challenges. *European Journal of Personality*, *34*(5), 632-648. https://doi.org/10.1002/per.2305

Aljaroodi, H. M., Adam, M. T. P., Chiong, R., & Teubner, T. (2019). Avatars and Embodied Agents in Experimental Information Systems Research: A Systematic Review and Conceptual Framework. *Australasian Journal of Information Systems, 23*. https://doi.org/10.3127/ajis.v23i0.1841

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. Retrieved from https://www.apa.org/ethics/code/

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct (2002, amended effective June 1, 2010, and January 1, 2017). http://www.apa.org/ethics/code/index.html

Arigo, D., Pagoto, S., Carter-Harris, L., Lillie, S. E., & Nebeker, C. (2018). Using social media for health research: Methodological and ethical considerations for recruitment and intervention delivery. *Digital Health*, *4*, 2055207618771757.

Banakou, D., Hanumanthu, P. D., & Slater, M. (2016). Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias. *Frontiers in Human Neuroscience*, 601. https://doi.org/10.3389/fnhum.2016.00601

Bombari, D., Schmid Mast, M., Canadas, E., & Bachmann, M. (2015). Studying social interactions through immersive virtual environment technology: virtues, pitfalls, and future challenges. *Frontiers in Psychology*, *6*, 869. https://doi.org/10.3389/fpsyg.2015.00869

Bradshaw, T. (2016). Neal Stephenson: VR, the Metaverse, and what Snow Crash got wrong. *Financial Times*. Retrieved from https://www.ft.com/content/64350756-db14-11e5-a72f-1e7744c66818

Brall, C., Schröder-Bäck, P., & Maeckelberghe, E. (2019). Ethical aspects of digital health from a justice point of view. *European Journal of Public Health*, *29* (Supplement 3), 18-22. https://doi.org/10.1093/eurpub/ckz167

Cieślik, B., Mazurek, J., Rutkowski, S., Kiper, P., Turolla, A., & Szczepańska-Gieracha, J. (2020). Virtual reality in psychiatric disorders: A systematic review of reviews. *Complementary Therapies in Medicine, 52*, 102480.

Cockerton, T., Zhu, Y. & Dhami, M. (2023) *Conducting Ethically-Sound Psychological Research in the Metaverse.* OFS https://osf.io/5tms2/?view_only=e89ee34ba6ad42a1ae7452b043bbb2f4.

Coelho, T., Marques, C., Moreira, D., Soares, M., Portugal, P., Marques, A., et al. (2020). Promoting reminiscences with virtual reality headsets: A pilot study with people with dementia. *International Journal of Environmental Research and Public Health*, 17(24), 9301. https://doi.org/10.3390/ijerph17249301

Crowell, C. R., Narvaez, D., & Gomberg, A. (2008). Moral psychology and information ethics: Psychological distance and the components of moral behavior in a digital world. In *Information Security and Ethics: Concepts, Methodologies, Tools, and Applications,* 3269-3281. IGI Global. https://doi.org/10.4018/978-1-59904-937-3.ch219

Davidoff, E. J. (2020). Agency and accountability: Ethical considerations for brain-computer interfaces. *The Rutgers Journal of Bioethics*, 11, 9-20.

De Ruiter, A. (2021). The distinct wrong of deepfakes. *Philosophy & Technology, 34*(4), 1311-1332. https://doi.org/10.1007/s13347-021-00459-2

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin, 130*(6), 959-988. https://doi.org/10.1037/0033-2909.130.6.959

Dionisio, J. D. N., III, W. G. B., & Gilbert, R. (2013). 3D virtual worlds and the metaverse: Current

status and future possibilities. *ACM Computing Surveys (CSUR)*, *45*(3), 1-38. https://doi.org/10.1145/2480741.2480751

Duan, H., Li, J., Fan, S., Lin, Z., Wu, X., & Cai, W. (2021, October). Metaverse for social good: A university campus prototype. *Proceedings of the 29th ACM International Conference on Multimedia* (pp. 153-161). https://doi.org/10.1145/3474085.3479238

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., ... & Wamba, S. F. (2022). Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *66*, 102542. https://doi.org/10.1016/j.ijinfomgt.2022.102542

EADe1 (2019) The IEEE global initiative on ethics of autonomous and intelligent systems. First Edition. *IEEE*. https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/autonomous-systems.html

Fairfield, J. A. (2012). Avatar experimentation: Human subjects research in virtual worlds. *UC Irvine Law Review*, *2*, 695.

Ferretti, A., Ienca, M., Sheehan, M., Blasimme, A., Dove, E. S., Farsides, B., ... & Vayena, E. (2021). Ethics review of big data research: What should stay and what should be reformed? *BMC Medical Ethics*, *22*(1), 1-13. https://doi.org/10.1186/s12910-021-00616-4

Finnegan, D. J., Zoumpoulaki, A., & Eslambolchilar, P. (2021). Does mixed reality have a Cassandra complex? *Frontiers in Virtual Reality, 2*. https://doi.org/10.3389/frvir.2021.673547

Fisher, C. B., & Fried, A. L. (2003). Internet-mediated psychological services and the American Psychological Association ethics code. *Psychotherapy: Theory, Research, Practice, Training, 40*(1-2), 103.

Franzke, A. S., Bechmann, A., Zimmer, M., and Ess, C. M. (2020) Internet Research: Ethical Guidelines 3.0 the Association of Internet Researchers (AoIR 2019), 11. https://aoir.org/reports/ethics3.pdf

Freeman, G., Zamanifard, S., Maloney, D., & Acena, D. (2022). Disturbing the peace: Experiencing and mitigating emerging harassment in social virtual reality. *Proceedings of the ACM on Human-*

Computer Interaction, *6*(CSCW1), 1-30. https://doi.org/10.1145/3512932

Friesen, P., Kearns, L., Redman, B., & Caplan, A. L. (2017). Rethinking the Belmont Report? *The American Journal of Bioethics, 17*(7), 15-21. https://doi.org/10.1080/15265161.2017.1329482

Gelinas, L., Pierce, R., Winkler, S., Cohen, I. G., Lynch, H. F., & Bierer, B. E. (2017). Using social media as a research recruitment tool: ethical issues and recommendations. *The American Journal of Bioethics, 17*(3), 3-14. https://doi.org/10.1080/15265161.2016.1276644

Gilbert, R. L. (2011). The prose (psychological research on synthetic environments) project: Conducting in-world psychological research on 3d virtual worlds. *Journal For Virtual Worlds Research*, *4*(1). https://doi.org/10.4101/jvwr.v4i1.2108

Gilbert, R. L., Murphy, N. A., Krueger, A. B., Ludwig, A. R., & Efron, T. Y. (2013). Psychological benefits of participation in three-dimensional virtual worlds for individuals with real-world disabilities. *International Journal of Disability, Development and Education*, *60*(3), 208-224. https://doi.org/10.1080/1034912x.2013.812189

Gilbert, R., & Forney, A. (2013). The distributed self: Virtual worlds and the future of human identity. In *The immersive internet* (pp. 23-36). *Palgrave Macmillan*, London. https://doi.org/10.1057/9781137283023_3

Golovianko, M., Gryshko, S., Terziyan, V., & Tuunanen, T. (2021). Towards digital cognitive clones for the decision-makers: adversarial training experiments. *Procedia Computer Science*, *180*, 180-189. https://doi.org/10.1016/j.procs.2021.01.155

Gosling, S. D., Augustine, A. A., Vazire, S., Holtzman, N., & Gaddis, S. (2011). Manifestations of personality in online social networks. *Cyberpsychology, Behavior, and Social Networking*, *14*(9), 483-488. https://doi.org/10.1089/cyber.2010.0087

Gregg, L., & Tarrier, N. (2007). Virtual reality in mental health. *Social Psychiatry and Psychiatric Epidemiology*, *42*(5), 343-354. https://doi.org/10.1007/s00127-007-0173-4

Gürerk, Ö., Bönsch, A., Kittsteiner, T., & Staffeldt, A. (2019). Virtual humans as co-workers. *Journal of Behavioral and Experimental Economics*, *78*, 17-29. https://doi.org/10.1016/j.socec.2018.11.003

Ha, J., Park, S., & Im, C.-H. (2022). Novel hybrid brain-computer interface for virtual reality applications using steady-state visual-evoked potential-based brain-computer interface and electrooculogram-based eye tracking for increased information transfer rate. *Frontiers in Neuroinformatics, 16*, 758537-758537.

Han, D. I. D., Bergs, Y., & Moorhouse, N. (2022a). Virtual reality consumer experience escapes: preparing for the metaverse. *Virtual Reality*. https://doi.org/10.1007/s10055-022-00641-7

Han, D. I. D., Melissen, F., & Haggis-Burridge, M. (2022b). Immersive Experience Framework: A Delphi Approach. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4084598

Han, E., Miller, M. R., DeVeaux, C., Jun, H., Nowak, K. L., Hancock, J. T., et al. (2023). People, places, and time: A large-scale, longitudinal study of transformed avatars and environmental context in group interaction in the metaverse. *Journal of Computer-Mediated Communication, 28*(2).

Harris, J., & Anthis, J. R. (2021). The moral consideration of artificial entities: A literature review. *Science and Engineering Ethics*, *27*(4), 53. https://doi.org/10.1007/s11948-021-00331-8

Haugeland, J. (1985). *Artificial Intelligence: The very idea*. Cambridge, MA: Bradford Books.

He, T. (2022). From single-player games to metaverse: A futuristic analysis of challenging legal issues in the video game industry in China. *GRUR International*. https://doi.org/10.1093/grurint/ikac039

Hendaoui, A., Limayem, M., & Thompson, C. W. (2008). 3D social virtual worlds: Research issues and challenges. *IEEE Internet Computing, 12*(1), 88-92. https://doi.org/10.1109/mic.2008.1

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Beyond WEIRD: Towards a broad-based behavioral science. *Behavioral and Brain Sciences*, *33*(2-3), 111. https://doi.org/10.1017/s0140525x10000725

Huynh-The, T., Pham, Q. V., Pham, X. Q., Nguyen, T. T., Han, Z., & Kim, D. S. (2023). Artificial intelligence for the metaverse: A survey. *Engineering Applications of Artificial Intelligence, 117*, 105581. https://doi.org/10.1016/j.engappai.2022.105581

Kavanagh, E., Litchfield, C., & Osborne, J. (2020). Virtual technologies as tools of maltreatment:

Safeguarding in digital spaces. In *Routledge Handbook of Athlete Welfare* (pp. 221-230). Routledge. https://doi.org/10.4324/9780429201745-25

Kellmeyer, P., Biller-Andorno, N., & Meynen, G. (2019) Ethical tensions of virtual reality treatment in vulnerable patients. *Nature Medicine,25*(8), 1185-8. https://doi.org/10.1038/s41591-019-0543-y

Kort, Y. A. D., Ijsselsteijn, W. A., Kooijman, J., & Schuurmans, Y. (2003). Virtual laboratories: Comparability of real and virtual environments for environmental psychology. *Presence: Teleoperators & Virtual Environments*, *12*(4), 360-373. https://doi.org/10.1162/105474603322391604

Kowert, R., & Cook, C. (2022). The toxicity of our (virtual) cities: Prevalence of dark participation in games and perceived effectiveness of reporting tools. *Proceedings of the 55th Hawaii International Conference on System Sciences*. http://hdl.handle.net/10125/79724

Kumar, S., & Singh, S. K. (2021). Brain-computer Interaction (BCI). *Insights2Techinfo*, pp. 1. https://doi.org/10.3389/frvir.2021.786665

Lawler-Dormer, D. (2018). Self-styling an emotionally intelligent avatar. *Technoetic Arts, 16*(1), 33-42. https://doi.org/10.1386/tear.16.1.33_1

Leonelli, S. (2020). Scientific research and big data. https://plato.stanford.edu/entries/science-big-data/?ref=hackernoon.com

Lin, Y., Chen, L., Ali, A., Nugent, C., Ian, C., Li, R., ... & Ning, H. (2022). Human Digital Twin: A survey. *arXiv preprint arXiv:2212.05937*. https://doi.org/10.21203/rs.3.rs-2369073/v1

Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers*, *31*(4), 557-564. https://doi.org/10.3758/bf03200735

Luger, G., & Stubblefield, W. (1993). *Artificial Intelligence: Structures & strategies for complex problem solving*. Benjamin Cummings Publishing Co. CA, USA.

Luppicini, R. (2013). The emerging field of technoself studies (TSS). In *Handbook of Research on Technoself: Identity in a technological society* (pp. 1-25). https://doi.org/10.1386/tear.16.1.33_1

Madary, M., & Metzinger, T. K. (2016). Real virtuality: A code of ethical conduct. Recommendations for good scientific practice and the consumers of VR-technology. *Frontiers in Robotics and AI*, 3. https://doi.org/10.3389/frobt.2016.00003

Madsen, J. K., Bailey, R., Carrella, E., & Koralus, P. (2019). Analytic versus computational cognitive models: Agent-based modeling as a tool in cognitive sciences. *Current Directions in Psychological Science*, *28*(3), 299-305. https://doi.org/10.1177/0963721419834547

Maister, L., Sebanz, N., Knoblich, G., & Tsakiris, M. (2013). Experiencing ownership over a dark-skinned body reduces implicit racial bias. *Cognition, 128*(2), 170-178. https://doi.org/10.1016/j.cognition.2013.04.002

Maloney, D., Freeman, G., & Robb, A. (2021, March). Social virtual reality: ethical considerations and future directions for an emerging research space. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (pp. 271-277). IEEE. https://arxiv.org/pdf/2104.05030.pdf

Marloth, M., Chandler, J., & Vogeley, K. (2020). Psychiatric interventions in virtual reality: Why we need an ethical framework. *Cambridge Quarterly of Healthcare Ethics*, *29*(4), 574-584. https://doi.org/10.1017/s0963180120000328

Mawer, M. (2016). Observational practice in virtual worlds: Revisiting and expanding the methodological discussion. *International Journal of Social Research Methodology*, *19*(2), 161-176. https://doi.org/10.1080/13645579.2014.936738

McCall, C., Hildebrandt, L. K., Hartmann, R., Baczkowski, B. M., & Singer, T. (2016). Introducing the Wunderkammer as a tool for emotion research: Unconstrained gaze and movement patterns in three emotionally evocative virtual worlds. *Computers in Human Behavior*, *59*, 93-107. https://doi.org/10.1016/j.chb.2016.01.028

McKeown, A., Mourby, M., Harrison, P., Walker, S., Sheehan, M., & Singh, I. (2021). Ethical issues in consent for the reuse of data in health data platforms. *Science and Engineering Ethics*, *27*(1), 1-21. https://doi.org/10.1007/s11948-021-00282-0

Merkx, C., & Nawijn, J. (2021). Virtual reality tourism experiences: Addiction and isolation. *Tourism Management*, *87*, 104394. https://doi.org/10.1016/j.tourman.2021.104394

Miller, M. R., Jun, H., Herrera, F., Yu Villa, J., Welch, G., & Bailenson, J. N. (2019). Social interaction in augmented reality. *PLOS ONE, 14*(5), e0216290. https://doi.org/10.1371/journal.pone.0216290

Minsky, M. (1986). *The Society of Mind*. Simon & Schuster, New York, USA.

Mystakidis, S. (2022). Metaverse. *Encyclopedia, 2*(1), 486-497. https://doi.org/10.3390/encyclopedia2010031

Narayanan, A., Bonneau, J., Felten, E., Miller, A., & Goldfeder, S. (2016). *Bitcoin and cryptocurrency technologies: a comprehensive introduction*. Princeton University Press.

Nebeker, C., Harlow, J., Espinoza Giacinto, R., Orozco-Linares, R., Bloss, C. S., & Weibel, N. (2017). Ethical and regulatory challenges of research using pervasive sensing and other emerging technologies: IRB perspectives. *American Journal of Empirical Bioethics*, *8*(4), 266-276. https://doi.org/10.1080/23294515.2017.1403980

Nguyen, T. N. (2022). Toward human digital twins for cybersecurity simulations on the metaverse: ontological and network science approach. *JMIRx Med, 3*(2), e33502. https://doi.org/10.2196/33502

Nowak, K. L., & Fox, J. (2018). Avatars and computer-mediated communication: A review of the definitions, uses, and effects of digital representations on communication. *Review of Communication Research, 6*, 30-53. https://doi.org/10.12840/issn.2255-4165.2018.06.01.015

Ozair, M. (2022). The Metaverse: Never too soon to discuss ethics. *Nasdaq*. https://www.nasdaq.com/articles/the-metaverse%3A-never-too-soon-to-discuss-ethics

Park, S. M., & Kim, Y. G. (2022). A Metaverse: Taxonomy, components, applications, and open challenges. *IEEE Access*, *10*, 4209-4251. https://doi.org/10.1109/access.2021.3140175

Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Frontiers in Human Neuroscience*, *9*, 660. https://doi.org/10.3389/fnhum.2015.00660

Parsons, T. D. (2019). *Ethical challenges in digital psychology and cyberpsychology*. Cambridge University Press. https://doi.org/10.1017/9781108553384

Preston, C., & Ehrsson, H. H. (2014). Illusory changes in body size modulate body satisfaction in a way that is related to non-clinical eating disorder psychopathology. *PLOS ONE*, *9*(1), e85773. https://doi.org/10.1371/journal.pone.0085773

Ramirez, E. J. (2019). Ecological and ethical issues in virtual reality research: A call for increased scrutiny. *Philosophical Psychology*, *32*(2), 211-233. https://doi.org/10.1080/09515089.2018.1532073

Ratan, R., Beyea, D., Li, B. J., & Graciano, L. (2020). Avatar characteristics induce users' behavioral conformity with small-to-medium effect sizes: A meta-analysis of the proteus effect. *Media Psychology*, *23*(5), 651-675. https://doi.org/10.1080/15213269.2019.1623698

Royakkers, L., Timmer, J., Kool, L., & van Est, R. (2018). Societal and ethical issues of digitization. *Ethics and Information Technology*, *20*(2), 127-142. https://doi.org/10.1007/s10676-018-9452-x

Sadler, R. (2017). Real ethical issues in virtual world research. In *Data Collection in Sociolinguistics* (pp. 59-63). Routledge. https://doi.org/10.4324/9781315535258-13

Schoentgen, A., & Wilkinson, L. (2021). Ethical issues in digital technologies. *23rd Biennial Conference of the International Telecommunications Society (ITS)*. http://hdl.handle.net/10419/238052

Setiawan, K. D., & Anthony, A. (2022, August). The essential factor of metaverse for business based on 7 layers of metaverse–Systematic literature review. In *2022 International Conference on Information Management and Technology (ICIMTech)* (pp. 687-692). IEEE. https://doi.org/10.1109/icimtech55957.2022.9915136

Seymour, M., Riemer, K., & Kay, J. (2018). Actors, Avatars and agents: Potentials and implications of natural face technology for the creation of realistic visual presence. *Journal of the Association for Information Systems*, *19*(10), 4. https://doi.org/10.17705/1jais.00515

Slater, M., Antley, A., Davison, D., Swapp, D., Guger, C., Barker, C., & Sanchez-Vives, M. V. (2006). A

virtual reprise of the Stanley Milgram obedience experiments. *PLOS ONE*, 1, e39. https://doi.org/10.1371/journal.pone.0000039

Slater, M., Gonzalez-Liencres, C., Haggard, P., Vinkers, C., Gregory-Clarke, R., Jelley, S., ... & Silver, J. (2020). The ethics of realism in virtual and augmented reality. *Frontiers in Virtual Reality*, *1*, 1. https://doi.org/10.3389/frvir.2020.00001

Srivastava, K., Das, R. C., & Chaudhury, S. (2014). Virtual reality applications in mental health: Challenges and perspectives. *Industrial Psychiatry Journal*, 23(2), 83–85. https://doi.org/10.4103/0972-6748.151666

Steinert, S., & Friedrich, O. (2020). Wired emotions: Ethical issues of affective brain–computer interfaces. *Science and Engineering Ethics*, *26*, 351-367. https://doi.org/10.1007/s11948-019-00087-2

Stendal, K., Balandin, S., & Molka-Danielsen, J. (2011). Virtual worlds: A new opportunity for people with lifelong disability? *Journal of Intellectual and Developmental Disability*, *36*(1), 80-83. https://doi.org/10.3109/13668250.2011.526597

Sun, J., Gan, W., Chen, Z., Li, J., & Yu, P. S. (2022). Big data meets metaverse: A survey. *arXiv preprint arXiv:2210.16282*.

Terziyan, V., Gryshko, S., & Golovianko, M. (2018). Patented intelligence: Cloning human decision models for Industry 4.0. *Journal of Manufacturing Systems*, *48*, 204-217. https://doi.org/10.1016/j.jmsy.2018.04.019

Vanacker, B., & Heider, D. (2012). Ethical harm in virtual communities. *Convergence*, *18*(1), 71-84. https://doi.org/10.1177/1354856511419916

Wajnerman Paz, A. (2021). Is mental privacy a component of personal identity? *Frontiers in Human Neuroscience, 15*, 773441-773441. https://doi.org/10.3389/fnhum.2021.773441

Wang, Y., Su, Z., Zhang, N., Liu, D., Xing, R., Luan, T. H., & Shen, X. (2022). A survey on metaverse: Fundamentals, security, and privacy. *arXiv preprint arXiv:2203.02662*. https://doi.org/10.48550/arXiv.2203.02662

Warner, N., & Teo, J. T. (2021). Neurological injury from virtual reality mishap. *BMJ Case Reports, 14*(10), e243424. https://doi.org/10.1136/bcr-2021-243424

Weber, S., Weibel, D., & Mast, F. W. (2021). How to get there when you are there already? Defining presence in virtual reality and the importance of perceived realism. *Frontiers in Psychology, 12*, 628298-628298. https://doi.org/10.3389/fpsyg.2021.628298

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: A review. *Frontiers in Psychology*, *10*, 158. https://doi.org/10.3389/fpsyg.2019.00158

Wegrzyn, M., Vogt, M., Kireclioglu, B., Schneider, J., & Kissler, J. (2017). Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLOS ONE*, 12(5), e0177239.

Willis, R. (2019). Observations online: Finding the ethical boundaries of Facebook research. *Research Ethics, 15*(1), 1-17. https://doi.org/10.1177/1747016117740176

Winston, J. J., & Hemanth, D. J. (2019). A comprehensive review on iris image-based biometric system. *Soft Computing*, *23*(19), 9361-9384. https://doi.org/10.1007/s00500-018-3497-y

Wolfendale, J. (2007). My avatar, my self: Virtual harm and attachment. *Ethics and Information Technology*, *9*(2), 111-119. https://doi.org/10.1007/s10676-006-9125-z

Yaremych, H. E., & Persky, S. (2019). Tracing physical behavior in virtual reality: A narrative review of applications to social psychology. *Journal of Experimental Social Psychology*, *85*, 103845. https://doi.org/10.1016/j.jesp.2019.103845

Yee, N., & Bailenson, J. (2007). The Proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research*, *33*(3), 271-290. https://doi.org/10.1111/j.1468-2958.2007.00299.x

Zallio, M., & Clarkson, P. J. (2022). Designing the metaverse: A study on inclusion, diversity, equity, accessibility and safety for digital immersive environments. *Telematics and Informatics*, *75*, 101909. https://doi.org/10.1016/j.tele.2022.101909