frontiers | Frontiers in Virtual Reality

TYPE Policy and Practice Reviews
PUBLISHED 05 December 2024
DOI 10.3389/frvir.2024.1443384

Check for updates

OPEN ACCESS

EDITED BY
Maria V. Sanchez-Vives,
August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Spain

REVIEWED BY
Deborah Richards,
Macquarie University, Australia
Jorge C. S. Cardoso,
University of Coimbra, Portugal

*CORRESPONDENCE
Verity McIntosh,
✉ verity.mcintosh@uwe.ac.uk

RECEIVED 03 June 2024
ACCEPTED 20 November 2024
PUBLISHED 05 December 2024

CITATION
McIntosh V and Allen C (2024) What do policymakers need to know about harassment in the metaverse?
Front. Virtual Real. 5:1443384.
doi: 10.3389/frvir.2024.1443384

What do policymakers need to know about harassment in the metaverse?

Verity McIntosh[1,2]* and Catherine Allen[3]

[1]Bristol VR Lab, University of the West of England, Bristol, United Kingdom, [2]Digital Cultures Research Centre, University of the West of England, Bristol, United Kingdom, [3]Limina Immersive, Bristol, United Kingdom

As immersive technologies and spatial computing paradigms move into the mainstream, public and political interest in the metaverse is growing. In some respects, the metaverse offers an exciting view of the future, one in which a global community can meaningfully connect regardless of where they are in the world. In contrast, however, early instances of "proto-metaverse" spaces have been plagued by reports of harassment and abuse. Policymakers around the world are now considering the role that governments might play in the regulation and governance of metaverse spaces, seeking to secure protections for citizens, and criminal accountability for offenders in this fast-evolving space. This paper introduces some of the key issues for governments engaging with this topic, including the suitability of existing legislative frameworks, and consideration of a new category of harm that seeks to recognise the distinctive impact of "conduct" abuses in metaverse environments.

KEYWORDS

metaverse, governance, abuse, harassment, immersive, policy

1 Introduction

There is yet to form one coherent definition of the term "metaverse" and there remains much debate about both what is inferred by the term, and how it is applied. This paper broadly utilises the X Reality Safety Intelligence (XRSI) definition of the metaverse as:

> "A network of interconnected virtual worlds with the following key characteristics: Presence, Persistence, Immersion and Interoperability" (XRSI, 2023)

Two of the stated defining characteristics, "presence" and "immersion" are well established in virtual reality scholarship (Bailenson, 2018; Lee, 2004; Slater, 2009) etc., "Persistence" may be less well examined in the literature, but is conspicuous in the design of many contemporary offerings such as "VRChat" or "Rec Room". These apps present users with a panoply of persistent worlds that can be visited at any time, and that will continue to exist (persist) both before and after the user is in attendance. Such applications build on the popular and well-established games formats e.g., MMORPG (massively multiplayer online role-playing game) and social simulation games such as "The Sims" or "Second Life". The fourth, metaverse-defining characteristic, "Interoperability" remains largely elusive. Tony Parisi's influential "Seven Rules of the Metaverse" argues that the metaverse must be "built upon interoperable technologies and tools, connected via rigorously defined and broadly agreed-upon free and open communications standards" (Parisi, 2018) At the time of writing, multiple XR standards are being developed and proposed by different consortia.

Multiple technical (Steed, 2024) and commercial challenges (Paul, 2022) remain unresolved, and the extent to which future metaverse spaces might prove "interoperable" still spans a wide spectrum of possibilities. We might deduce that the "metaverse", if it is to meet the criteria of the posited definition, is not yet fully realised.

Policymakers looking to develop strategies with resonance into the future may have concerns that it is too early, therefore, to take an interventionist standpoint. Leading some to favour "self-governance" from technology companies; hoping and waiting for best practice to emerge. It could also be argued that instances of harassment and abuse are evident now in proto-metaverse spaces, and that legislators have a responsibility to act on behalf of citizens, regardless of the relative immaturity of an industry (Almond et al., 2024). Policymakers in the immediate term might be well advised to look to, and learn from proto-metaverse experiences that are already garnering large user-bases, and exhibiting most, if not all of the features as set out by the XRSI definition. They may also wish to retain some degree of practical and epistemological flexibility to accommodate future adaptations, allowing legislation and governance strategies to keep pace with this fast-moving sector.

Forms of harassment discussed in this paper are generally limited to behavioural activity occurring in real time in what might be considered a metaverse, or proto-metaverse environment. Proto-metaverse environments may also be referred to as "social VR" apps or "collaborative virtual environments" (CVEs). These are multi-person, co-present shared virtual environments. Visitors may connect to these environments from anywhere in the world, each person manifesting as an embodied avatar whose movements synchronize with the user's physical actions via the sensors in a virtual reality head mounted display (HMD) and controllers. Additional input may include eye tracking, facial expression mapping and/or body tracking systems to further enhance the correlation between the movements of the natural person and their virtual avatar. Vocal communication is often an important feature of these spaces, typically taking the microphone input from the HMD and relaying it spatially and naturalistically within the virtual environment, such that the user's voice can be experienced as natural, person-to-person speech.

In this context, harassment may take the form of verbal or behavioural abuse, and/or the positioning and movements of avatars and virtual objects to enact behaviours experienced as aggressive, violating, offensive or demeaning by other users.

An understanding of harassment and abuse could be extended to include areas such as data and privacy abuses, identity cloning, social and political manipulation, fraud, theft, exploitation, and coercive or discriminatory AI avatar behaviours. For the purpose of clarity, this paper will focus solely on real time, peer-to-peer encounters involving one or more natural persons embodied as avatars in a virtual environment. It should be noted, however that instances of harassment in metaverse contexts may form part of a wider pattern of abuse, taking place both on- and offline and should be considered in such a context when abuses are reported.

2 Harassment and abuse

In recent years, occurrences of harassment and abuse within proto-metaverse platforms, sometimes referred to as social VR platforms, have been well documented in the media. They have received less scrutiny in a scholarly context, and in relation to the roles and obligations that governments may be considered to have to intervene in this space.

Reports suggest that instances of peer-to-peer harassment in social VR are a relatively common, and tend to increase in virtual environments devoid of managed hosting or a clear purpose, with female users and minoritized people most likely to be targeted (Limina Immersive, 2018). A survey of over 600 + users in 2018 suggested that 49% of regular female VR users reported experiences of sexual harassment or abuse in virtual social spaces (Outlaw, 2018). Since then, with the rise in public adoption of VR headsets, the issue appears to have persisted and perhaps escalated. In 2021 the Center for Countering Digital Hate asserted that users of popular social VR platform, VRChat were exposed to abusive behaviour once every 7 min (Center for Countering Digital Hate, 2021). Numerous reports of sexual harassment and abuse within the metaverse have been reported in the media (Eccles, 2022; Patel, 2021; Rifkind, 2022).

A report produced by the authors of this paper for the United Kingdom's Institution of Engineering and Technology signals the prevalence of racist language, homophobic language, transphobic language, non-consensual touch and simulated violence including sexual violence in social VR spaces (Allen and McIntosh, 2022). Research by the same authors for United Kingdom charity the National Society for the Prevention of Cruelty to Children (NSPCC) additionally identifies instances of criminals using proto-metaverse environments to gain access to, abuse and exploit children (Allen and McIntosh, 2023; McIntosh and Allen, 2023a).

3 Impact

Although the nature of harassment and abuse in VR differs from real-world instances, the impact on individuals can be significant. Slater calls attention to the confluence of psychologically convincing Place Illusion (PI) and Plausibility Illusion (Psi) in virtual reality, giving users a strong sense of presence, and implicating their body in the virtual space. "If you are there (PI) and what appears to be happening is really happening (Psi) then *this is happening to you!* Hence you are likely to respond as if it were real. We call this "response-as-if-real" RAIR." (Slater, 2009).

Several researchers have pointed to the compounding impact of "social presence" (Lee, 2004; Ratan, 2012) i.e., the awareness of being co-present with other users, conversing and taking consequential action in a shared, virtual environment. This attribute is often understood in combination with "self-presence" and "environmental presence", cumulatively forming a powerful sense of "being there" that has been identified as particular to virtual reality (Bailenson, 2018). Ratan has suggested that social presence might be considered to be the most impactful of the three, as the participation of other natural persons in a virtual space adds complex social cuing to the simulative environment, further convincing users of the veracity, immediacy and embodied nature of their experience (Ratan, 2012).

The United Kingdom's Cyberpsychology Research Group call attention to the contiguous emotional impact of negative experiences in metaverse environments "Just because these events happen online rather than offline does not mean they are not being experienced as real" (Askham, 2022). Madary and Metzinger take it a step further, introducing the possibility that "[t]orture in a virtual environment is still torture. The fact that one's suffering occurs while one is immersed in a virtual environment does not mitigate the suffering itself" (Madary and Metzinger, 2016).

In the context of all of the above it seems likely that the immersive and embodied nature of social, metaverse environments will significantly intensify the impact of harassment and abuse such as physical threats or simulated violence. In metaverse environments, non-consensual instances of touching, verbal harassment or invasion of personal space may put users at particular risk of psychological and emotional distress. Future developments such as haptic technology clothing may further heighten this affect by adding a physical sensation to abuse enacted in metaverse contexts in the future.

Even without the use of specific haptic technology, there is some evidence to suggest that some people, using only a headset and controllers, experience uncanny physical sensations upon being touched in virtual environments (Pilacinski et al., 2023). Some hypothesise that the psychologically convincing nature of metaverse environments can lead users to partially transfer their phenomenal self-model (PSM) into that of their avatar, an effect akin to the Rubber Hand Illusion (Botvinick and Cohen, 1998). Consequently, they may report feeling pronounced physical sensations when they observe their avatar being touched or harmed, even though their physical body remains uncontacted (Desnoyers-Stewart et al., 2024; Madary and Metzinger, 2016; McIntosh and Allen, 2023a).

A comparable phenomenon, "phantom touch", is frequently discussed by users of social VR. Although largely under-researched in a formal setting, this sensation appears to involve users perceiving a touch sensation on their bodies that directly corresponds to a simulated act of touch in VR. Some users actively cultivate this sensation, observing themselves (or more accurately their avatar) being touched, patted and stroked in virtual mirrors as a means to more viscerally associate virtual touch with tactile sensation. There would appear to be enhanced likelihood that those who experience a form of "phantom touch" could be at greater risk of traumatic impact in the event of harassment and abuse (McIntosh and Allen, 2023a).

One often-posed question in regard to VR abuse, from those not familiar with the technology is, "why did not you just take the headset off?". Preliminary research suggests that rapid disengagement from VR, particularly under stress or anxiety, can provoke panic or dissociative episodes, therefore, the solution may not be as simple as disconnecting (Allen and McIntosh, 2022). This question also signals a tendency towards victim blaming, failing to account for well understood trauma-response behaviours such as freeze and appeasement in response to high stress, high risk encounters (Cantor and Price, 2007).

4 Design responses

In response to apparent abuses in proto-metaverse spaces, many app developers and platform owners have sought design solutions to mitigate the risk or severity of potential harms. Some have turned to social science research that may not have been initially conceived in relation to technology paradigms, drawing on research exploring physical and relational behaviour as a route into understanding the needs of social, virtual spaces.

Hideaki Matsui, a design lead at Google has publicly discussed their use of Hall's theory of Proxemics (Hall, 1966). Proxemics suggests that people will maintain differing amounts of distance from one another depending on the social setting and their cultural backgrounds. Google use this framework as a schematic, encouraging designers to construct virtual environments that conserve distances between users that are appropriate to the social context and levels of intimacy that might be anticipated in a particular encounter. As per Hall's design, they distinguish between public, social, personal and intimate space and design experiences accordingly. Their approach notably does not incorporate Hall's framing of such boundaries being informed by background and cultural context.

Michelle Cortese, Design Lead Manager at Meta also advocates for the use of Proxemics, extending the use to reference certain BDSM practices such as negotiated, mutual consent. She writes about the significant number of people, particularly women, who reported being sexually harassed or assaulted in multi-person virtual reality spaces in the late 2010 s, and calls for an approach to personal space management that involves explicit and informed mutual consent.

> "we suggest designers build granular controls that are easy to access and surface before intimate interactions begin. It's important that people can customize and control the types of experiences they're willing to have with other people in these close quarters before they happen" (Cortese and Zeller, 2019)

In the intervening years, many of these recommendations have been adopted, with features such as "personal space bubbles" now available in most social VR apps. Personal space bubbles enforce an invisible boundary around the user, keeping other avatars at a designated distance, or rendering them invisible and inaudible when the allotted space is impinged. In some instances, users can choose only to be perceptible to those pre-designated as "friends" to minimise the risk of harassment.

Whist such design features may prove useful, they can also create an imbalance of power that favours the aggressor. The onus is on the victim to apply extreme caution entering into a metaverses space, configuring complex safety features and limiting their own experience prior to entry, or attempting to do so in the moment whilst experiencing harassing behaviours. Those persistently harassing other users, notionally violating the terms of use of the platform, encounter no such barriers.

In addition to personal space and visibility configuration; block, mute and report tools are frequently available to users in social VR spaces. These are designed be deployed *ad hoc* in the event of unwanted attention or abusive behaviour. However, for victims of abuse; such reporting features can be difficult to navigate in the moment, especially when abuse is ongoing. It is also generally unclear what responses or punitive measures might follow from the reporting of such instances. To date platforms are not obliged by any regulatory authority to make transparent their internal monitoring, evidentiary and justice systems, to disclose actions taken to investigate or remediate reports of abuse, or to notify the complainant of any actions taken (Allen and McIntosh, 2022; Center for Countering Digital Hate, 2021).

5 Regulation and governance

5.1 Suitability of existing laws

Around the world, governments are seeking advice on whether existing legislature is sufficient to ensure that their citizens are afforded the same rights, protections and freedoms in metaverse spaces as they might expect in comparable physical and digital spaces.

One key challenge to efficacy, is that many current legal frameworks around the world related to abuse and harassment make clear distinctions between "content" abuses which can include the posting and sharing of abusive materials such as text, imagery and video, and physical "contact" abuses, which generally involve unwanted physical touch.

Many governments have sought to improve protections for citizens in 2D online platforms in recent years. New criminal designations are being written onto the statute books for online criminal behaviour such as "cyber-flashing" and the posting of "revenge porn" (Online Safety Act 2023, 2023). In the relatively new field of multi-person, metaverse environments, there is currently little legislative provision to account for abuses that might take place in psychologically convincing, simulative environments where multiple natural persons are co-present and interacting with one another.

Given what is understood about the immersive, embodied and relational qualities of metaverse environments, governments may need to specify a new category of harm. Perhaps one that recognises certain forms of user "conduct" as harassment and abuse, even where there is no physical contact, or associated production or proliferation of content.

In a report examining risks to children engaging with XR platforms, Pettifer et al. (2022) recommend the delineation of risk in metaverse contexts along similar lines; content risks, contact risks and conduct risks. These categories are offered as a means to illustrate the range, and distinctive nature of threats and harms identified in their research. In the case of Pettifer et al., the term "conduct" is utilised to explicate the additional vulnerability of children using metaverse platforms, rather than to label the behaviour of abusive users. They explain that as part of normal child development, many children currently use the internet to "engage actively in risk-taking behaviours [...] purposefully accessing and/or downloading inappropriate and illegal content or sharing intimate personal information or images" (Pettifer et al., 2022). They express concern that these types of predictable behaviours or "conduct" when translated into metaverse contexts, can significantly increase a child's risk of being victimized and targeted for abuse.

The term "conduct", then, may hold value as a legal category or classification with which to describe and evaluate peer-to-peer, embodied encounters in XR. The ability to assess abuses in relation to the interpersonal "conduct" of those involved, beyond the sharing of content, or the physicality of contact might assist responsible bodies in both identifying those at risk, and holding to account those who would seek to exploit and abuse other users.

5.1.1 Case study

As an early test of the suitability of existing legislature, United Kingdom police announced in January 2024 that they were investigating an alleged instance of "sexual attack" of a girl who is under 16, and has reported being abused by a group of men in a social VR setting (Camber, 2024).

In an interview with LBC News, The United Kingdom's Home Secretary, James Cleverly said "I know it is easy to dismiss this as being not real, but the whole point of these virtual environments is they are incredibly immersive. We're talking about a child here, and a child has gone through sexual trauma. It will have had a very significant psychological effect and we should be very, very careful about being dismissive of this." (Taylor, 2024).

In response to this case, the chairman of the United Kingdom's Association of Police and Crime Commissioners, Donna Jones was quoted as saying "We need to update our laws because they have not kept pace with the risks of harm that are developing from artificial intelligence and offending on platforms like the metaverse." (Taylor, 2024).

The statements of two such prominent public figures suggests an appetite at policy level to apply some of the principles discussed in this paper at the highest levels of governance. This specific case is understood to be ongoing at the time of publishing. It will be interesting to see how existing legislation is applied and reconciled, and the epistemological frameworks used to explain the outcome of this apparently unprecedented case.

5.2 Accountability

Policymakers may wish to consider creating stronger links between activity in the metaverse and national law enforcement agencies. This would ensure that serious crimes committed in metaverse worlds do not remain under the exclusive jurisdiction of the platform's internal justice system, which is arguably better suited to technology-related issues than serious criminal offences. Public confidence will also need to be built such that anyone reporting abuses to civic authorities can expect to be understood, believed, and for their complaint to be acted upon.

Criminal prosecution of individuals for abusive conduct in the metaverse is one area that governments certainly need to consider. Another is the relative culpability and accountability of the companies providing metaverse apps, platforms and services. Where frequent instances of criminal activity, such as abusive behaviour are found to be taking place in a particular app or platform, regulators may wish to consider holding providers wholly or partially accountable, particularly if they are failing to uphold legal standards, and either encouraging or turning a blind eye to persistent abusive behaviour.

In the US, holding platforms to account is likely to prove challenging. Section 230 of the Communications Act affords legal immunity for providers of interactive computer services with respect to the actions of their users (Section 230, 1934).

In the United Kingdom, the new Online Safety Act (2023) has some provision for this, extending a "duty of care" to platform owners and managers regarding content that users should be able to encounter online. The challenge of “content” *versus* “conduct” and “contact” is largely unaddressed in the Act, however the metaverse has been deemed explicitly in scope (Local Government Association, 2022). The Institute for Engineering and Technology recently called on United Kingdom government to ensure that new legislation is made fit for purpose in relation to social, spatial environments (Almond et al., 2024).

The EU’s new Digital Services Act (European Union, 2023) goes further still, holding “very large” tech companies legally accountable for the content posted on their platforms. Again, the Act sets out a framework for addressing illegal “content” online, however there is no direct provision for metaverse contexts, and it remains unclear how the more behavioural, conduct-based forms of abuse and harassment might be addressed by this new legal framework.

In many ways, questions of accountability in metaverse environments echo those of Web 2.0 and social media cultures. Much has been learned in recent years about the impact of online abusive behaviours, and the responsibilities of individuals, corporate entities and governments in mitigating harm. Whether or not new legal instruments such as those described above can extend, or be bent to meet the emerging needs of immersive environments, remains an open question.

5.3 Jurisdiction

In most legislative frameworks, sovereign jurisdiction is determined by the geography of where an alleged crime has taken place. For many exponents of the metaverse, the promise of this new paradigm lies in its potential to be borderless and decentralised. Just as cryptocurrency could be conceived as an alternative to centralised banking systems, so the metaverse might be imagined as an alternative to state-based territoriality for interpersonal encounters. What then for state-based authorities looking to respond to reports of criminal activity, including reports of harassment and abuse in the metaverse?

As with the internet before it, questions of jurisdiction in metaverse contexts are proving challenging. Users of such spaces may be encountering one another in what experientially is a common metaverse environment, but connecting from very different territories, each with their own particular legal contexts. To further complicate matters, the metaverse environment visited might be provided by a company in another territory, with the underpinning technology stack hosted across multiple territories. What legal frameworks should then apply when abuses are detected? And which nation(s) should have the jurisdiction to prosecute criminal behaviour?

Laws governing interpersonal behaviour vary considerably between territories, and jurisdictional ambiguity can create a vacuum of legal accountability, a lag in governmental response to evident harms, and a gulf of support for victims of criminal behaviour.

Even in instances when jurisdiction is relatively unambiguous, or where laws can be expected to be common across territories, challenges can remain. For instance, most legal systems descended from English law e.g., Australia, Canada, New Zealand, Singapore and the United States, conform to similar systems of Tort law (civil laws pertaining to interpersonal wrongdoing between private persons). However, it remains unclear whether such laws would be legally applicable in metaverse contexts as the legal “personhood” of an avatar is yet to be determined. Questions remain regarding whether the actions of an avatar in a virtual world should be considered directly analogous to the action of the embodied “natural person” controlling it. Or whether avatar behaviour would be better understood as akin to a playable video game character (Cheong, 2022). Each approach would attract a very distinct legal response, particularly in relation to acceptable levels of interpersonal violence.

In the absence of legal certainty, there is concern that cases of abuse and harassment may become entrenched in costly, intractable disputes regarding which legal jurisdiction applies, risking a drain on resources in multiple territories, and lessening the likelihood of successful conviction (Europol, 2022; Kalyvaji, 2023).

One approach would be to make platforms responsible for ensuring that the legal protections of each user are implemented in the design of the space before they are granted access to a given metaverse environment. Where legal frameworks in different jurisdictions prove incompatible, this may lead to citizens from certain territories being excluded, or companies running multiple instantiations of metaverse environments, the user being directed to the space that is compliant with their domestic legal system. An alternative, or addition perhaps, is to encourage closer working with international agencies such as Interpol to ensure the complementarity of different governmental approaches, and to enhance international cooperation agreements, enabling cross-jurisdictional prevention strategies and joined-up response to crimes involving metaverse technologies and environments.

5.4 Stakeholder literacy

Perhaps a more manageable short to mid-term strategy for government agencies engaging with this topic, could be to address stakeholder literacy. Terms such as “virtual and augmented reality”, “spatial computing”, “immersive media” and “metaverse” remain quite amorphous and impenetrable in the public imagination. A lack of direct experience and an absence of cultural context beyond science fiction can make it challenging for citizens and government agencies to conceive of the parameters of existing risk. For progress to be made, it may be necessary to invest in the “immersive literacy” (Allen, 2021) of key stakeholders. Jones, Dawkins and McDougall explore how immersive literacy programmes might build upon, and extend approaches developed for media literacy education (Jones et al., 2022). Such initiatives could potentially foster critical public discourse around this emerging socio-technical paradigm (Bulger and Davison, 2018; Livingstone, 2004). Policymakers with aspirations towards higher levels of “metaverse” literacy could include specialist training programmes designed to give stakeholders direct experience of embodied proto-metaverse platforms, providing insight into the current trajectory and pace of technological developments, and unpacking the manner in which the affordances of this medium relate to issues of abuse and harassment. Governments may wish to consider prioritising the literacy of responsible bodies such as legislators, police and the judiciary. Public literacy campaigns may also be valuable in supporting citizens to understand their rights, and empowering them to make informed and empowered choices about their own engagement with the metaverse.

6 Actionable recommendations

The following is offered in support of policymakers interested in assessing the suitability of existing legislative frameworks, and in appraising the need for novel approaches. The author recommends that policymakers:

- Recognise that harassment in the earliest iterations of the 'metaverse' (or proto-metaverse), is a current, well evidenced harm, not a future risk.
- Understand that harassment in embodied, spatial environments is likely to be experienced differently to online and in person abuses.
- Acknowledge that legal frameworks developed to address online and in person and abuses may lack appropriate categories to describe abusive, interpersonal behaviour in social, virtual environments.
- Consider extending existing frameworks to include 'conduct' as well as content and contact categories of abuse.
- Establish a clear and enforceable position regarding levels of criminal accountability for peer-to-peer abuse in virtual environments. Particular attention could be given to defining:
    - o Legal personhood of avatars and accountability for behaviour enacted by natural persons whilst embodied as an avatar in virtual worlds.
    - o Categories of harassment that might apply in virtual environments e.g., verbal threat, hate speech, harassment, psychological trauma, virtual body sovereignty, psychosomatic and tele-haptic experiences of unwanted touch etc
    - o Responsibility and accountability of platforms/services providers to preserve users' rights and civic protections in virtual spaces.
- Assess the existence and suitability of governmental systems, beyond the internal justice systems of platforms/services providers to support citizens. Are sufficient abuse prevention measures in place, as well as evidence gathering and prosecution strategies to enable a timely and proportionate response to reported instances of harassment and abuse in a virtual environment?
- Seek to develop cross-territorial approaches, harmonizing domestic policy with international partners to combat challenges of jurisdictional accountability and authority.
- Design stakeholder and public literacy campaigns that can improve awareness of risks, support ethical, legal and political engagement, and create a culture of informed and empowered citizenship.

6.1 Expanded view

A report commissioned by the Council of Europe 'Risks and Opportunities of the Metaverse' (McIntosh and Allen, 2023b) (written by the authors of this paper) offers further support to policymakers across the EU. It directly addresses the specific circumstances of metaverse and proto-metaverse environments and how they differ from existing digital platforms. Whilst this paper offers a focussed view on harassment and abuse in the metaverse, policymakers may additionally wish to review the impact of other areas affecting democracy, human rights and the rule of law. The report explores how areas such as data, privacy, artificial intelligence, creativity, health, community, sustainability, digital democracy and social and political manipulation are all impacted by the rapid development of metaverse infrastructure. To fully assess issues of harassment and abuse, policymakers may wish to develop a more holistic view of the immersive policy space, appreciating metaverse technologies as both drivers and signifiers of societal transformation along a range of interconnected axes.

7 Conclusion

Although the metaverse is often positioned as a "future horizon" technology, it is evident that early versions of the metaverse are already here, and that instances of harassment and abuse are taking place with potentially significant consequences for citizens. Governments have an opportunity to urgently consider the suitability and efficacy of existing legislature, and to assess whether new legal instruments are needed to reflect the distinctive experience of embodied, immersive, multi-person environments. Policymakers may also wish to consider prevention, reporting and prosecution strategies, as well as the accountability of both individuals and platforms/service providers in relation to abusive behaviours in metaverse environments. Programmes of immersive or metaverse literacy now could equip stakeholders and the wider public with the information they need to collectively design and advocate for more positive futures for the metaverse.

8 Recommendations for further research

It is worth noting that this paper focusses primarily on metaverse and proto-metaverse contexts as experienced via virtual reality HMD devices. These involve the use of headsets to obfuscate the sights and sounds one's immediate surroundings and introduce psychologically convincing virtual worlds that are most likely unconnected to one's physical surroundings. Artist Ben Joseph Andrews aptly describes this as a "strange occurrence where you're both present somewhere you're not, and absent somewhere that you are." (Andrews, 2024) As technologies continue to evolve, devices are being brought to market with the capability to provide more blended experiences. These augmented, mixed or extended reality experiences preserve something of the user's view of the world around them and introduce virtual elements such as avatars and 3D objects into the user's field of view. The proto-metaverse elements of this are relatively immature, largely limited to video conferencing and co-working paradigms. However, as these more blended virtual environments proliferate, a series of distinct social and behavioural dynamics may emerge that warrant additional consideration. Future research in this area would further support policymakers in addressing issues of harassment and abuse that traverse physical and virtual domains.

Author contributions

VM: Writing–original draft, Writing–review and editing. CA: Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The Council of Europe provided support for an expert report than underpins much of the research for this paper. Additional funds were provided by the University of West of England's Impact, Collaboration and Evidence (ICE) internal QR fund.

Conflict of interest

Author CA is Director of Limina Immersive.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Allen, C. (2021). Digital democracies: unleashing the positive power of the metaverse. *Digit. Democr*. Available at: https://thresholdstudios.tv/unleashing-the-positive-power-of-the-metaverse-catherine-allen/ Accessed July 22, 2024.

Allen, C., and McIntosh, V. (2022). Safeguarding the metaverse. Available at: https://www.theiet.org/media/9836/safeguarding-the-metaverse.pdf Accessed May 06, 2022.

Allen, C., and McIntosh, V. (2023). Child safeguarding and immersive technologies: an outline of the risks. Available at: https://learning.nspcc.org.uk/research-resources/2023/child-safeguarding-immersive-technologies Accessed July 22, 2024.

Almond, E., McIntosh, V., and Allen, C. (2024). *An open letter to Ofcom on the need to urgently review how VR spaces are governed*. London, United Kingdom: The IET. Available at: https://www.theiet.org/media/press-releases/press-releases-2024/press-releases-2024-january-march/3-january-2024-an-open-letter-to-ofcom-on-the-need-to-urgently-review-how-vr-spaces-are-governed Accessed March 09, 2024.

Andrews, B. J. (2024). *Embodied/misembodied*. London, United Kingdom: YouTube. Available at: https://www.youtube.com/watch?v=AoLqaF_Fx94 Accessed March 09, 2024.

Askham, G. (2022). *Metaverse: new documentary exposes racial and sexual abuse*. London, United Kingdom: Glamour. Available at: https://www.glamourmagazine.co.uk/article/metaverse-misogyny Accessed September 21, 2023.

Bailenson, J. (2018). *Experience on demand: what virtual reality is, how it works, and what it can do*. First edit. W.W. Norton and Company.

Botvinick, M., and Cohen, J. (1998). Rubber hands "feel" touch that eyes see. *Nature* 391 (6669), 756. doi:10.1038/35784

Bulger, M., and Davison, P. (2018). The promises, challenges, and futures of media literacy. *J. Media Lit. Educ.* 10, 1–21. doi:10.23860/jmle-2018-10-1-1

Camber, R. (2024). *British police probe VIRTUAL rape in metaverse*. London, United Kingdom: Daily Mail.

Cantor, C., and Price, J. (2007). Traumatic entrapment, appeasement and compl*ex post*-traumatic stress disorder: evolutionary perspectives of hostage reactions, domestic abuse and the Stockholm syndrome. *Aust. N. Z. J. Psychiatry* 41 (5), 377–384. doi:10.1080/00048670701261178

Center for Countering Digital Hate (2021). New research shows Metaverse is not safe for kids. *Cent. Countering Digital Hate (CCDH)*. Available at: https://counterhate.com/blog/new-research-shows-metaverse-is-not-safe-for-kids/ Accessed May 23, 2023.

Cheong, B. C. (2022). Avatars in the metaverse: potential legal issues and remedies. *Int. Cybersecurity Law Rev.* 3 (2), 467–494. doi:10.1365/s43439-022-00056-9

Cortese, M., and Zeller, A. (2019). Designing Safer Social VR Using the ideology of sexual consent to make social VR a better place. Available at: https://immerse.news/designing-safer-social-vr-76f99f0be82e Accessed July 20, 2022.

Desnoyers-Stewart, J., Bergamo Meneghini, M., Stepanova, E. R., and Riecke, B. E. (2024). Real human touch: performer-facilitated touch enhances presence and embodiment in immersive performance. *Front. Virtual Real.* 4. doi:10.3389/frvir.2023.1336581

Eccles, L. (2022). *My journey into the metaverse — already a home to sex predators*. London, United Kingdom: The Sunday Times.

European Union (2023). The digital services act. Available at: https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package Accessed March 09, 2024.

Europol. (2022). Policing in the metaverse: what law enforcement needs to know, an observatory report from the Europol Innovation Lab. doi:10.2813/81062

Hall, E. T. (1966). *The hidden dimension*. NY Doubleday.

Jones, S., Dawkins, S., and McDougall, J. (2022). A virtual journey towards new literacies in *Understanding virtual reality* (London, United Kingdom: Routledge), 141–156. doi:10.4324/9780367337032-14

Kalyvaji, M. (2023). Navigating the metaverse business and legal challenges: intellectual property, privacy, and jurisdiction. *J. Metaverse* 3 (1), 87–92. doi:10.57019/jmv.1238344

Lee, K. M. (2004). Presence, explicated. *Commun. Theory* 1 (1), 27–50. doi:10.1111/j.1468-2885.2004.tb00302.x

Limina Immersive (2018). Immersive content formats for future audiences. Available at: www.digicatapult.org.uk Accessed September 12, 2023.

Livingstone, S. (2004). Media literacy and the challenge of new information and communication technologies. *Commun. Rev.* 7 (1), 3–14. doi:10.1080/10714420490280152

Local Government Association (2022). *Online safety bill, second reading*. London, United Kingdom: House of Commons. Available at: https://www.local.gov.uk/parliament/briefings-and-responses/online-safety-bill-second-reading-house-commons-19-april-2022 Accessed March 09, 2024.

Madary, M., and Metzinger, T. K. (2016). Recommendations for good scientific practice and the consumers of VR-technology. *Front. Robotics AI* 3. doi:10.3389/frobt.2016.00003

McIntosh, V., and Allen, C. (2023a). Child safeguarding and immersive technologies: key concepts. Available at: https://learning.nspcc.org.uk/research-resources/2023/child-safeguarding-immersive-technologies Accessed September 12, 2023.

McIntosh, V., and Allen, C. (2023b). Risks and opportunities of the metaverse virtual reality. *Springer Sci. Bus. Media Deutschl. GmbH* 26 (2). doi:10.1007/s10055-021-00564-9

Outlaw, J. (2018). Survey of social VR users. *Ext. Mind*. Available at: https://www.extendedmind.io/2018-survey-of-social-vr-users Accessed March 08 2024.

Parisi, T. (2018). The seven rules of the metaverse. A framework for the coming immersive by tony parisi metaverses medium. Available at: https://medium.com/meta-verses/the-seven-rules-of-the-metaverse-7d4e06fa864c Accessed October 03, 2022.

Patel, N. J. (2021). Reality or fiction? *Medium*. Available at: https://medium.com/kabuni/fiction-vs-non-fiction-98aa0098f3b0 Accessed September 21, 2023.

Paul, K. (2022). Meta and other tech giants form metaverse standards body, without Apple. *Reuters*. Available at: https://www.reuters.com/technology/meta-other-tech-giants-form-metaverse-standards-body-without-apple-2022-06-21/#:~:text=Apple%20has%20been%20heavily%20involved,ensure%20it%20supported%20the%20format Accessed June 03, 2024.

Pettifer, S., Barrett, E., Marsh, J., Hill, K., Turner, P., and Flynn, S. (2022). The future of eXtended reality technologies, and implications for online child sexual exploitation and abuse.

Pilacinski, A., Metzler, M., and Klaes, C. (2023). Phantom touch illusion, an unexpected phenomenological effect of tactile gating in the absence of tactile stimulation. *Sci. Rep.* 13 (1), 15453. doi:10.1038/s41598-023-42683-0

Ratan, R. (2012). Self-presence, explicated: body, emotion, and identity extension into the virtual self in *Handbook of research on technoself: identity in a technological society* (IGI Global), 321–335. doi:10.4018/978-1-4666-2211-1.ch018

Rifkind, H. (2022). *The metaverse will be an abuser's paradise*. London, United Kingdom: The Sunday Times.

Section 230 (1934). Crsreports. Available at: https://crsreports.congress.gov.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. Biol. Sci.* 364 (1535), 3549–3557. doi:10.1098/rstb.2009.0138

Steed, A. (2024). Some technical challenges of scaling from social virtual reality to metaverse(s) in *A metaverse for the good*. Editors M. Slater, M. Barngrover, D. Friedman, A. Lopez-Tarruella, O. Niamut, S. Pan, et al. (Barcelona, Spain: Universitat de Barcelona), 126–131.

Taylor, W. (2024). *Police investigate "rape" in metaverse after group of men attack girl in virtual reality room*. London, United Kingdom: LBC News.

XRSI (2023). *The metaverse - X reality safety intelligence (XRSI)*. San Francisco, CA: XRSI. Available at: https://xrsi.org/definition/the-metaverse Accessed March 08, 2024.