Virtual Worlds, Real Risks: Exploring Early Signs of Radicalization in the Metaverse

Author: Tim Verkerk
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT,

As immersive virtual environments become increasingly popular, concerns have started emerging about their potential role in facilitating online radicalization. The real-time interaction combined with anonymity and a high degree of immersion can amplify social influence and emotional engagement within these 3D social platforms. While prior research on online radicalization has mostly been focused on text- and video-based platforms, much is left to learn on how political or polarizing interactions manifest in more immersive settings.

This study aims to explore the presence and nature of politically charged communication in VRChat through five observation sessions in the virtual world "The Black Cat." Using a framework adapted from e-participation and previous research on online radicalization, the study categorizes political expressions from ideological commentary and propaganda to recruitment efforts.

The results reveal a broad spectrum of political communication which came often intertwined with humor and provocation. Some interactions contained disinformation, slurs, or ideological messages while others showed more sincere conversation on issues like abortion, immigration, or religion. The badly moderated and chaotic nature of the environment gave room for both radical conversations as well as more constructive discussions, highlighting the dual potential of VRChat.

The results contribute to research on radicalization in the metaverse by showcasing how the immersive affordances of VR may lower the barriers of extremist communication.

Graduation Committee members:

*Disclaimer, ChatGPT was used for background research and text flow

1. INTRODUCTION

1.1 Background information and context

Over the past years, the boundaries between the physical and digital worlds have started to blur (Goltz, 2011). Immersive virtual environments, have grown into ideal environments for social interaction, entertainment, education and work. Large tech companies like Meta, Microsoft and Bytedance invest billions of dollars in the concept of 'the metaverse' (Financial Times, 2024). Today, the term 'the metaverse' refers to a collection of computer-generated online environments within augmented, mixed and virtual reality (Anderson, Rainie, & Atske, 2022). These spaces can vary from everyday use AR-apps on smartphones to immersive game worlds and digital copies of the real world.

While these developments open new possibilities for connection and creativity, they also bring challenges. Especially regarding harmful content, misinformation and extremist ideology sharing.

Earlier research on online political radicalization had its focus on text and video platforms such as X and Youtube. The new and unique features that 3D virtual environments provide like avatar embodiment, real-time interaction and relative anonymity, intensify social influence in ways that remain underexamined Projects like the RadiGame project and reports by the Global Network on Extremism and Technology suggest that VR chatrooms, virtual meetups in online games and other 3D online social spaces facilitate strong group bonding and rapid formation of an echo chamber where the information and opinions that a user encounter is comparable to their own ideologies with a reinforcing effect. (prif.org; White, 2023)

1.2 Problem Statement

Immersive virtual worlds are evolving rapidly into social meeting spaces that allow millions of people to come together and socialize (Bao Tran, 2025). While offering unique chances for creativity, community forming and making people feel like they have a place where they belong, they can also enable negative thoughts to spread quickly. Most interactions on the platform are social or playful, sensitive topics are also discussed. In some cases, these discussions can contain politically motivated messages or provocative statements. The high degree of immersion, anonymity and real-time interactions can enable extremist content to spread while being hard to notice (Bakshy, Rosen, Marlow, & Adamic, 2012), especially among younger people who are more sensitive to peer pressure and are looking for a feeling of belonging. Up until now, systematic research into concrete signals that indicate the early stages of radicalization is missing. This knowledge gap makes it harder to act in time and to ensure a safe digital environment.

Surveys indicate that younger generations increasingly view virtual platforms as valid spaces for social learning and political discussions (Anderson et al., 2022). However, systematic research on early indicators of radicalization in these immersive settings is largely missing. This gap prevents timely intervention and limits our understanding of security in these emerging digital environments

The aim of this study is to explore the presence and forms of political or polarizing communication in VRChat. Through structured observation, this research hopes to shed light on the new social dynamics in immersive virtual environments.

1.3 Research Question

"In what forms can political or polarizing messages be observed in the VRChat environment The Black Cat?"

To answer this question, this study will use a non-participant observational design, exploring one VRChat environment. The research focuses on the nature and expression of political messages within user interactions in these spaces.

2. THEORETICAL INFORMATION

2.1 Defining online radicalization

Radicalization can be defined as: 'the action or process of causing someone to adopt radical positions on political or social issues' (Oxford English Dictionary). It is the slow convincing of individuals to adopt ideas that differ from the normalized way of thinking. *'It is a psychological trajectory that, given the right circumstances, can happen to any person, group, or nation. The trajectory is not right or wrong; it is amoral in the sense that radicalization can occur for causes both good and bad.' (McCaulkey & Moskalenko, 2011, as cited in Thompson, 2012).* Early studies saw radicalization happen in secrecy and face-to-face. For example, Sageman's (2004) investigation of 172 al-Qaeda members which showed that recruitment happened through friendship and family circles and that trust is built by repeatedly meeting in person (Sageman, 2004). Today, many first steps take place on the internet, guided by the algorithms of platforms. A review of 82 studies shows that social-media sites are now the main place where people are introduced to extremist ideas. (Akram & Nasar, 2023).

2.2 Key drivers on social media

Online radicalization is based on four key drivers in social media environments, according to Akram and Nasar (2023):

1. Push factors, Emotional or psychological frustrations push individuals towards extremist content.

2. Pull factors, In the search for identity and belonging, individuals are pulled into online communities that serve as echo chambers.

3. Propaganda, Multimedia formats disguise propaganda in entertaining content, reaching youth covertly (Weimann & Masri, 2020).

4. A lack of moderation, Minimal platform governance can fuel polarization, hostility and misinformation (Del Vicario et al., 2015; Bessi & Quattrociocchi, 2016)

These drivers do not operate in isolation. The interplay between push and pull factors is particularly important in online environments, where platform algorithms and design features accelerate exposure to extremist narratives. According to the United Nations Office on Drugs and Crime (UNODC, 2018), these dynamics often reinforce one another, while extremist content offers belonging and justification. Propaganda and the lack of effective moderation serve as amplifiers, allowing messages to spread quickly and with minimal resistance. In immersive platforms such as VRChat, the potential for radicalization may be even greater. Real-time interaction, spatial presence and anonymity intensify the emotional and psychological impact of social influence. Understanding how

these drivers manifest in such settings is crucial to identifying early signs of political extremism in virtual worlds

2.3 Radicalization in Immersive social virtual reality

Social VR platforms combine social media's accessibility with heightened sensory immersion. Anonymity, the use of pseudonyms and a lack of permanent message logs make it easier for users to discuss extremist ideas unnoticed (Blackwell et al., 2019). Features such as spatial audio, body-tracking and avatar customization increase realism and potentially amplify persuasive influence. This could shape the delivery and reception of political content, as users might adapt their behaviour to match the norm of their avatars. This so-called *Proteus effect* may further encourage conformity to group norms (Yee, Bailenson, & Ducheneaut, 2009).

Additionally, the embodied experience in VR allows for tactics like whispering, cornering, or blocking exits, making aggressive behaviours feel more intense than on more traditional social platforms. Private rooms can offer a secluded space where ideological grooming can occur, blending offline secrecy with online scalability

2.4 Platform power and the illusion of participation

Sherry Arnstein's (1969) "Ladder of Citizen Participation" is a way of evaluating the amount of influence that citizens have over a decision-making process. The ladder identifies 8 steps with the bottom two, manipulation and therapy, being categorized as non-participation. These two can give an illusion of involvement, while the real power of the decision-making process remains in the hands of authorities.

In social VR spaces, the illusion of participation is clearly present. Users can feel like they are being given influence on platform rules and systems through surveys, community guidelines, or pop-up messages that promote respectful behaviour. However, their actual influence remains minimal. The real decisions such as content moderation or enforcement of rules are typically made higher up the chain by the platform owners or automated systems. As argued by Zheng et al. (2022), the responsibility often falls on the users, for example by offering possibilities to mute or block others, rather than solving deeper structural issues.

VR platforms allow users to speak freely and interact in immersive environments that feel very close to reality. Unfortunately, the lack of transparent governance and moderation allow polarizing content to circulate without being monitored. Even though users feel like they have an active role in shaping the platform's culture, the power remains centralized and the ability of users to moderate harmful content is very limited. The lower rungs of Arnstein's (1969) participation ladder connect strongly with the difference between perceived and actual influence within VR worlds which can have consequences for the spread and normalization of polarizing messages.

2.5 Research gap

Although social media radicalization has been widely studied, research on its translation into 3D immersive environments is still limited. As immersive virtual worlds become more mainstream, understanding how known drivers manifest in these new settings becomes increasingly urgent (Anderson et al., 2022)

3 Methodology

3.1 Research Design

This study uses a qualitative, exploratory, non-participant observation approach to examine political and polarizing discourse in VRChat. Given the novelty of the topic and a limited amount of existing data, a qualitative design is most suitable.

Two VRChat worlds will be selected and entered by the researcher as a silent observer. The goal is to record naturally occurring interactions without interfering, although minimal introduction may occasionally be required.

3.2 Data collection

Observations will be conducted over a one-week period (June 7 – June 14), with one observation per day, each lasting 20-40 minutes. This frequency allows for a manageable dataset while maintaining richness throughout said dataset. Contextual details- date, world name, login time and number of active users- will be noted upon entering. Politically charged or polarizing messages will be transcribed in real time. Notes will include tone, reactions and group dynamics. Recordings will be made using field notes only (no audio or video) to maintain participant anonymity. After each session, a short reflection (5-10 sentences) will summarize the impressions and notable interactions.

3.3 Classifying Political Communication

Messages will be categorized using a structured framework of political expression adapted from e-participation literature (Arnstein, 1969) and research on digital activism (Akram & Nasar, 2023)   Each category represents a different kind of political engagement or messaging that may reflect risks of radicalization. For clarity, the categories are summarized in table 1 below

Category	Description	Level of Participation
Incitement to violence	Calls to violence or glorification of violent acts	Manipulation (non-Participation)
Political promotion	Advocacy for political parties, leaders, or ideologies	Informing (Tokenism)
Recruitment for political purposes	Attempts to recruit users into movements or groups	Consultation (Tokenism)
Manipulation or disinformation	Spreading of false or misleading political content	Manipulation (non-participation)
Propaganda (Visual/Audio)	Use of images, slogans, or sounds to communicate ideologies	Therapy or Informing
Commentary on social issues	Discussion of societal topics like inequality or discrimination	Placation (Tokenism)
Other Relevant Expressions	Content that is political but does not fit in one of the afore mentioned categories	Varies depending on content

'Table 1'

3.4 Sampling

To find a chatroom that would be fitting for this study, the final choice fell on 'The Black Cat'. The Black Cat is a chatroom that has the looks of an old and cozy café/bar. The choice was based on user activity, accessibility and the reputation that this chatroom holds as a place for informal social interaction. The world is known for its diverse user base and its spontaneous conversation which makes it a suitable place for an observing study.

Sampling was done while keeping certain things in mind. The sessions were held at different times of the day to have a higher chance of observing users from different time zones and because of a possible difference in activity. The room was entered using a standard avatar and stayed away from any participation to ensure minimal influence on the interactions between users.

Other VRChat worlds were considered but rejected due to either low user activity, language barriers or a setting that does not support casual conversation that well. By focusing on The Black Cat, the study hopes to gain a sample of political messaging in a social and unfiltered virtual environment.

3.5 Data analysis

The collected data will be analysed through a manual coding process with a focus on identifying recurring patterns and notable themes in the observed interactions. After each session, the notes and transcriptions will be reviewed to highlight statements or behaviours that are related to political or polarizing messages.

Each highlighted message will be assigned one of the predefined categories listed in the classification table. The analysis will then involve grouping similar examples and summarizing what they say about user behaviour and political expression in VRChat.

This approach allows for transparency and flexibility while following the structure of the category framework. Short summaries per room will be used to reflect on context and interaction dynamics.

4 Results

4.1 General Observations

The analysis is based on five non-participant observation sessions conducted in the VRChat world The Black Cat. This world was chosen for having a consistent high number of active users and has a culture where spontaneous conversations are stimulated.

The general interaction style in The Black Cat is fast-paced, chaotic and often driven by humor. Users strive for attention or social validation by making provocative or exaggerated remarks. This can lead to the normalization of controversial language. Despite this, several sessions revealed meaningful discussions on a wide range of topics.

4.2 The different sessions

The following section consists of short summaries of each of the observation sessions. The summaries show in what way political and polarizing discussions took place in The Black Cat. They highlight the variation in tone, content and user interaction across the different sessions.

4.2.1 Session 1, The Black Cat, 11-06-2025, 21:15-22:00

This session introduced the general culture of The Black Cat: competitive, chaotic and prestation driven. Users often spoke over one another and conversations rapidly shifted. Political content was low, though at some moments cultural and ideological topics were mentioned. Someone shouted "Allahu Akhbar" as a joke and a short discussion about Sea Shepherds framed them to be like modern-day pirates. The overall tone of this session was unserious yet clearly showed how normalized boundary pushing language is in the social metaverse culture.

4.2.2 Session 2, The Black Cat, 12-06-2025, 14:00-14:50

In this session more political content was mentioned, still often in provocative or sarcastic tones. Comments included anti-ICE statements and support for the riots in Chicago. Some participants said ICE was to blame for deporting U.S. citizens, others joined in expressing disdain for the government. A conversation struck up about George Floyd, with one user saying "he deserved what happened" this statement was met with silence and disbelieve from the other users. Racial and homophobic slurs occurred frequently, most of these were uttered without intent to harm, suggesting that the culture is one of desensitized expression.

4.2.3 Session 3, The Black Cat, 13-6-2025, 18:00-18:45

This session was the most political. A debate on abortion unfolded between one anti-abortion user claiming the fetus can feel pain and others defending the right to abortion as one of bodily autonomy. The conversation was heated but structured and the setting here was less chaotic compared to the previous sessions. The discussion shifted into broader political topics, about immigration and a discussion on Trump vs Kamala Harris. Pro-Trump comments were mostly about closing the borders and reducing crime, while the other side raised ethical concerns and referenced controversial quotes made by Trump and unverified claims. This showed that The Black Cat, though mostly relaxed and casual, can also function as a space for more nuanced discussions.

4.2.4 Session 4, The Black Cat, 13-06-2025, 22:00-22:40

This session had a more religious tone. One user, who sounded like a young adult, delivered something of a sermon on Christian salvation. He declared things like "even Jeffrey Dahmer can be forgiven", "Jesus looks at who we are, not what we do" and "Keep reading the Bible!" with conviction. While this can be interpreted as a good-natured expression of faith, the framing of forgiveness for extreme crimes can be seen as moral normalization. The room was largely silent or indifferent, listening but not paying too much attention to it. The session showed how religious ideology can manifest in persuasive ways in VR spaces.

4.2.5 Session 5, The Black Cat, 14-6-2025, 14:00-15:00

This session stood out for its calm and thoughtful tone. The participants talked about what their grandparents or great grandparents experienced during the Second World War. One person shared an anecdote on their great grandfather being in the Norwegian army during the war. From here, the discussion slowly started shifting towards a discussion about the necessity of nuclear weapons. The group broadly agreed that the invention of nuclear arms has shown to be a tragic one, the weapons now serve as a deterrent for greater violence. This topic is mostly geopolitical, however the discussion remained respectful and reflective which showed that serious nuanced conversations are also supported in The Black Cat

Combined, these sessions show the diversity of discourse in immersive social VR. Political and polarizing content emerges spontaneously, often intertwined with humor, provocation, or

personal storytelling. While not all expressions show signs of radicalization, they reveal a space where ideology, identity and community are continuously shaped through conversations.

4.3 Categorization of Political Communication

To better understand the nature of political and polarizing messages in VRChat, observed statements were organized into seven categories. Each category reflects a specific function or intention behind the message which can range from emotionally charged opinion sharing to deliberate persuasion.

One of the most common categories observed was commentary on social issues. This category includes discussions on controversial societal topics. Users express personal opinions on themes such as race, gender, police violence and abortion. Examples included statements like "abortion is murder" or "Do you think the Chicago Riots are a good thing? – Yes", showing that social debates can take place in immersive environments. Users can share opinions like these freely because of the lack of moderation within VRChat.

Political promotion was also an often-occurring category, particularly about U.S. politics and immigration. Users expressed their admiration for people like Donald Trump or criticized government policies or agencies such as ICE. Statements like "ICE is the problem" and "We don't let everyone in. That's what I like about Trump" show that the political exchanges were often highly personal.

Another much occurring category is manipulation or disinformation. Unverified accusations, such as "There is proof that Trump and Epstein raped a 13-year-old." And "Obama used to be against opening borders and now he is for it." were confidently said without any factual evidence. On one hand, the real-time conversations within VRChat make the spread of misinformation easier, on the other, the online setting of VRChat allowed users to quickly browse the internet and look up statements that they do not immediately believe. Even though misinformation is easily spread in the social metaverse, there are also tools that users can use to check facts and statements made by others.

Propaganda was seen through promoting belief systems, most notably in religious aspects. One user preached about Christian salvation and forgiveness and even went as far as saying "Even Jeffrey Dahmer can be saved by Jesus". These messages were mostly delivered in a preaching way where one user spoke and clearly reflected their own ideology.

Although less frequent, some things that can fall into the recruitment category were also seen. In the sessions in The Black Cat, the only statements for this category were about religion. Things like "Keep reading the Bible!" and "If you truly give in to Christ, you will be forgiven" were said in a direct but non-aggressive tone. Although these messages were not overly radical, they can, over time, still influence others.

Incitements to violence were rare. No comments were heard that really activated someone to do something violent. However, certain remarks were said that justified or normalized harm. For example, a comment about George Floyd suggested that what happened to him was deserved. This does not call for violence, but it can contribute to the normalization of harmful situations.

Finally, many statements did not fit neatly in one of the predefined categories but were still valuable. These were grouped under other relevant expressions. This category mostly consists of jokes, sarcastic remarks or exaggerations. In The Black Cat, these statements were common and they reveal that the group dynamic is often one where language that tests the boundaries is frequently used. Besides political sarcasm and ironic nationalistic views, racist and homophobic slurs were also often used. Though most of the time, these remarks were meant as jokes and not in a harmful way, they can contribute to a culture where discrimination is normalized.

5 Discussion

5.1 Interpretation of Results

VRChat has shown to facilitate a wide range of different types of conversation in a, often, chaotic social setting. Anonymity, avatar embodiment and spatial presence lower social barriers which can lower the threshold of expressing polarizing or sensitive opinions.

Though most conversations started in a casual way and stayed that way, some conversations made a gradual shift into deeper discussions. Humor and provocative statements are two things that were used to gain attention or to test the reaction of others. This dynamic allows users to experiment with ideas that are mostly harmless but could have a negative impact on the normalization of extreme or controversial statements.

Some sessions showed that users can take strong positions and engage in structured discussions. In other cases, such as religious preaching, the immersive nature of VRChat highlighted how the platform can be used to spread one's ideology. Meanwhile, calm casual conversations alternated with chaotic yelling and swearing, showing that VRChat can, one second be a calm and respectful environment and the next, a place where no one can be heard because of all the chaos.

The contrast between the different types of dialogue found in VRChat shows how VRChat can be a platform for constructive and meaningful conversation, while also being a space where social boundaries become blurred and are easily experimented with or crossed.

5.2 Theoretical Implications

The results of this study confirm and extend existing frameworks on online radicalization. Akram & Nasar's (2023) identification of push and pull factors is supported by the dynamics observed in VRChat. Users showed feelings of mistrust toward institutions like ICE, while bonding with other users through shared complaints or jokes. This aligns closely with how the radicalization process is described in earlier online settings, but the immersive nature of social VR intensifies this.

Arnstein's (1969) ladder of Citizen Participation gives another strong viewpoint. Many of the observed interactions fell into the lower rungs like manipulation, therapy, or tokenism. However, the illusion of participation was strong. Users took political stands, vented frustrations and even attempted to persuade others.

The immersive nature of VR enhances the performative and emotional aspects of political conversations. The Proteus Effect (Yee & Bailenson, 2007) may explain how users' behavior and tone can shift depending on the group context, avatar appearance, or peer pressure. Because the line between role-play and real belief is thin, it is easy for users to test radical or controversial ideas without any consequence. This blurred boundary can lower the barriers to extremist engagement.

5.3 Practical Implications

The results show that immersive platforms challenge conventional approaches to content moderation and political communication. Unlike traditional social media, where text and images are often monitored by filters and not synchronized with live interactions, VRChat operates in real time with voice,

embodiment and spatial dynamics. This creates blind spots for moderation and makes it difficult to intervene in the spread of harmful or polarizing content or disinformation.

Developers and platform administrators may need to think of new moderation strategies in these environments. Automated moderation tools are not able to detect sarcasm, jokes, coded language, or persuasion when it's delivered through live voice chat. A hybrid approach where human moderators, community-based reports and context sensitive AI work together to identify patterns of harmful communication may prove more successful.

Additionally, programs could be adapted to virtual worlds that equip users with tools to recognize manipulative speech, propaganda and peer pressure. Prevention strategies can include building inclusive social spaces where empathy, collaboration and cross-cultural dialogue are rewarded rather than provocation and performance.

The platform can also add a purpose where it can be used as a place for civic engagement. The ability of users to engage in political debates, even in unstructured settings, shows that virtual spaces have an untapped potential for education, deliberation and conversation. Studying how these environments can be used to educate and empower, may find value in the future.

5.4 Limitations

This study encountered several limitations that affect the generalizability and depth of its results. First, the research is based on five observation sessions within one virtual world (The Black Cat) over the course of four days. This world was chosen for its social nature and activity; however, the sample remains small.

Second, the non-participatory nature of the research limited deeper engagement with users. By choosing to observe rather than interact, the study ensured the authenticity of the conversations, but sacrificed the chance for follow up questions or intent clarification. Many statements, especially the sarcastic or humoristic ones, were open to multiple interpretations.

Finally, due to ethical and privacy considerations, no audio or video recordings were used. Although the field notes proved valuable, they are limited by the observer's perspective and memory. This may have led to incomplete capture of fast-paced, overlapping conversations.

In hindsight, a more participatory approach where conversations were actively joined and potential interviewees were sought out may have provided deeper insights. Users who spend a significant amount of time in the metaverse are more likely to have a more nuanced understanding of recurring ideological patterns and the dynamics of the platform. Future research should aim to explore these spaces more extensively to build a richer, context sensitive understanding of the political culture in immersive environments.

5.5 Recommendations

Based on the results and limitations of this study, a couple of recommendations can be made for future research and for those involved in the moderation or design of virtual environments.

First, a more participatory approach could be useful. The non-participant approach of this study allowed for a natural view of user behavior and made sure that the conversations that were observed occurred naturally, it also created limitations in understanding the user's intent. As the research time was also limited, a participatory approach could help with gaining a

deeper insight into long-term patterns and social norms by interviewing users who spend more time in virtual environments.

Additionally, the focus on a single world limited generalizability. Broadening the view to include more varied VRChat worlds would help find out if what was observed in The Black Cat is consistent with the norm or if other worlds contain completely different views.

The study also shows that moderation is very limited. Many polarizing or provocative statements were made with an undertone of humor or sarcasm. Future moderation should try to find a way to better combine voice recognition, real-time flagging and human moderation

People in The Black Cat often had little to no reaction to polarizing, discriminatory or racist statements. Spreading awareness around the normalization of such messaging is another recommendation. Providing tools to encourage reflection and reporting, like a short survey when leaving a world that asks if something uncomfortable happened, could help with spreading this awareness.

Finally, the study highlights that VR platforms can also be used for civic engagement. Even though the environments are mostly chaotic and unserious, users did occasionally engage in serious conversations about politics, ethics and society. It suggests that platforms like VRChat have an opportunity to become a multifaceted space that does not only serve as a hub for social interaction and entertainment but can also function as a place for education and other civic engagement.

5.6 Conclusion

The study set out to explore the dynamics of political communication in immersive virtual environments, with VRChat as its primary case. Through five non-participant observation sessions in The Black Cat, a spectrum of political conversation was identified. The findings reveal that social VR platforms could function both as incubators for radical or polarizing ideas and as spaces for genuine dialogue and civic expression.

While the observed conversations were often unstructured, moments of ideological persuasion and group polarization appeared naturally. The categories of manipulation, political expression and ideological messaging helped structure the analysis and showed a range of behaviors that mirror broader online radicalization patterns.

Ultimately, the study underscores the risks and opportunities of political conversation in the metaverse. Platforms like VRChat are not inherently political but they can facilitate environments where political arguments can be formed, challenged and negotiated. As immersive technologies continue to evolve, gaining a better understanding of their role in shaping political conversation becomes not only relevant but necessary.

6 Acknowledgements

References

Goltz, N. (2010). *ESRB warning: Use of virtual worlds by children may result in addiction and blurring of borders – The advisable regulations in light of foreseeable damages. Pittsburgh Journal of Technology Law and Policy, 11*. https://doi.org/10.5195/tlp.2011.57

Peace Research Institute Frankfurt (PRIF). (n.d.). *RadiGaMe project*. https://www.prif.org/en/research/radigame

Financial Times. (2024, May 24). *Meta's $80bn gamble on virtual reality yet to pay off*. https://www.ft.com/content/c513949e-3fc1-43a2-9358-363dff823bc1

Tran, B. (2025). *Metaverse Adoption Rate: How Many Users Are Joining?* https://patentpc.com/blog/metaverse-adoption-rates-how-many-users-are-joining

Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. A. (2012). The role of social networks in information diffusion. *In Proceedings of the 21st international conference on World Wide Web (WWW '12)* (pp. 519–528). ACM. https://doi.org/10.1145/2187836.2187907

Lamphere-Englund, G., & White, J. (2023, May). *The online gaming ecosystem: Assessing socialisation, digital harms and extremism mitigation efforts*. Global Network on Extremism and Technology (GNET). https://doi.org/10.18742/pub01-133

Oxford University Press. (n.d.). *Radicalization*. In *Oxford English Dictionary*. https://www.oed.com/

Thompson, R. L. (2012). *Friction: How radicalization happens to them and us. By Clark McCauley and Sophia Moskalenko (New York: Oxford University Press, 2011)* [Review of the book *Friction: How radicalization happens to them and us*, by C. McCauley & S. Moskalenko]. *Journal of Strategic Security, 4*(4), 195–196. https://doi.org/10.5038/1944-0472.4.4.10.

Sageman, M. (2004). *Understanding terror networks*. University of Pennsylvania Press.

Akram, S., & Nasar, M. (2023). Online radicalization: A systematic review of literature. *Journal of Policing, Intelligence and Counter Terrorism*, *18*(1), 1–22. https://doi.org/10.21121/eab.1166627

Weimann, G & Masri, N. (2020). *Research Note: Spreading Hate on TikTok. Studies in Conflict & Terrorism. 46. 1-14*. https://doi.org/10.1080/1057610X.2020.1780027

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2015). *Echo chambers in the age of misinformation*. arXiv. https://arxiv.org/abs/1509.00189

Bessi, A., & Quattrociocchi, W. (2016). *Echo chambers: Emotional contagion and group polarization on Facebook. arXiv*. https://arxiv.org/abs/1607.01032

Blackwell, L., Ellison, N. B., Elliott-Deflo, M., & Schwartz, R. (2019). *Harassment in social virtual reality: Challenges for platform governance. Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–25. https://doi.org/10.1145/3359202

Yee, N., Bailenson, J. N., & Ducheneaut, N. (2009). *The Proteus Effect: Implications of transformed digital self-representation on online and offline behavior. Communication Research, 36*(2), 285–312. **https://doi.org/10.1177/0093650208330254**

Arnstein, S. R. (1969). *A ladder of citizen participation. Journal of the American Institute of Planners*, *35*(4), 216–224. https://doi.org/10.1080/01944366908977225

Anderson, J., Rainie, L., & Atske, S. (2022). *The metaverse in 2040. Pew Research Center.* https://www.pewresearch.org/internet/2022/06/30/the-metaverse-in-2040/

United Nations Office on Drugs and Crime. (2018). *Counter-terrorism module 2: Key issues – Drivers of violent extremism. United Nations.* https://www.unodc.org/e4j/en/terrorism/module-2/key-issues/drivers-of-violent-extremism.html

Zheng, Qingxiao & Ngoc, Do & Lingqing, Wang & Yun, Huang. (2022). *Facing the Illusion and Reality of Safety in Social VR.* https://doi.org/10.48550/arXiv.2204.07121