

**DISCOURSE
SOCIETY
MEDIA**

Patricia Bau,
Tamara Bodden,
Christina Liemann (eds.)

POISONED PLAY

Toxicity
in Online Gaming Discourse

[transcript]

Patricia Bau, Tamara Bodden, Christina Liemann (eds.)
Poisoned Play

Editorial

The series “Discourse – Society – Media” includes linguistic works with an interdisciplinary approach. The studies of the series are concerned with social discourses and the question how language shapes our perception of reality. The focus lies on the role of language in its social contexts and the influence of media conditions.

Patricia Bau, born in 2000, works as a research associate at the Department of German Studies at Universität Koblenz. Her doctoral thesis deals with strategies of tabooing and detabooing in discourses on menopause and midlife crisis. Her research focuses on discourse linguistics, linguistic taboo research, online communication and gender studies.

Tamara Bodden is a junior professor in the Department of German Studies at Universität Koblenz. After her doctorate in discourse linguistics at Universität Kassel in 2022 she worked as a postdoctoral researcher at the Simon Fraser University in Vancouver receiving the Feodor Lynen scholarship of the Alexander von Humboldt foundation. Her research areas include online toxicity, discourse linguistics, AI & gender as well as art communication.

Christina Liemann, born in 1996, works as a visiting scholar at the Waseda University in Tokyo. She wrote her linguistic doctoral thesis on communicative strategies of right-wing influencers at Universität Kassel. Her research is focused on political language, discourse analysis, digital and multimodal communication as well as gender studies.

Patricia Bau, Tamara Bodden, Christina Liemann (eds.)

Poisoned Play

Toxicity in Online Gaming Discourse

[transcript]

This work was funded by the University of Kassel's Open Access Publication Fund.

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available online at <https://dnb.dnb.de>



This work is licensed under the Creative Commons License BY 4.0. For the full license terms, please visit the URL <https://creativecommons.org/licenses/by/4.0/>.

Creative Commons license terms for re-use do not apply to any content (such as graphs, figures, photos, excerpts, etc.) not original to the Open Access publication and further permission may be required from the rights holder. The obligation to research and clear permission lies solely with the party re-using the material.

2026 © Patricia Bau, Tamara Bodden, Christina Liemann (eds.)

transcript Verlag | Hermannstraße 26 | D-33602 Bielefeld | live@transcript-verlag.de

Cover design: Lena Schäfferling

Printing: Elanders Waiblingen GmbH, Waiblingen

<https://doi.org/10.14361/9783839477533>

Print-ISBN: 978-3-8376-8169-7 | PDF-ISBN: 978-3-8394-7753-3

ISSN of series: 3054-4882 | eISSN of series: 3054-4890

Printed on permanent acid-free text paper.

Contents

Introduction

Patricia Bau, Tamara Bodden & Christina Liemann7

“Megan Fox is in a toxic relationship”

Representations of the ‘Toxic’ in Linguistic Community and Media Discourse

Take Hoffmeister & Lars Sörries-Vorberger 21

Tea Bagging, Trolling & Trash Talk

A Multimodal Approach for Toxicity in Twitch Streams

Tamara Bodden & Ashhab Ahmed..... 41

»Next Level Frauenhass«

Tabu in der diskursiven Auseinandersetzung zu *No Mercy*
als toxische Kommunikationsverstrickung

Patricia Bau..... 73

“you woke cultists are a cancer to everything”

Toxicity through Personal References in the Steam Group ‘Woke Content Detector’

Christina Liemann 93

“Dream of a greater Reich”

Fascist Role-Playing, Post-Irony and Alt-Histories
in *Hearts Of Iron IV* Mod Communities

Andreas Niegl 123

Anti-Diversität im Gaming – ein Shortcut in rechtsextreme Gaming-Communitys?	
<i>Julia Reinhold & Mick Prinz</i>	159
 »Crunch as Curriculum«	
Toxic Learning Culture and the Aesthetics of Overwork in Concept Art Education for the Games Industry	
<i>Marion Plank</i>	183
 List of Contributors	201

Introduction

Patricia Bau, Tamara Bodden & Christina Liemann

Insults, harassment, threats, hate speech, aggression, and negativity – while the internet creates new and diverse communication and networking opportunities, more possibilities for the democratization of discourse (Habermas 2006: 554), it can also be a fertile ground for negative interactions. The high levels of toxicity on online platforms are therefore a recurring focus of public discourse and a current research topic in disciplines such as linguistics, sociology, and adjacent fields (e.g. Mall et al. 2020; Fortuna 2020). As it is a complex topic that involves multiple actors, platform affordances, social norms, and semiotic means, “[r]esearch about toxic language, and by consequence also the detection of toxic language, is marked by its interdisciplinarity.” (Gevers et al. 2022: 33)

The gaming market is growing and can no longer be described as a niche phenomenon (e.g. game Verband 2025: 74): It is already more substantial in monetary value than film and music industries combined (Dentsu 2024). This also contributes to an increased social and societal relevance of gaming and the discourses that take place in and around games. Gaming platforms are not only a space for entertaining leisure activities; they can also be significant for finding like-minded communities, forming political opinions, but also can have relevance for establishing new communicative practices – both in a positive and a negative (and potentially toxic) sense.

The public discourse on gaming in Germany initially began with the medium being stigmatized and digital games being associated with arcades and gambling (Montag 2021: 93). Media critics warned about potential negative impacts, especially with regard to young people – as is often the case with a new medium. Going forward, gaming continued to be stigmatized, as especially online gaming and the genre of first-person-shooters were associated with mass shootings in schools, though these accusations have since been refuted in many studies (see among others Garbe 2020; Bareither 2016).

Following this initial period of a negative discourse, gaming related content and communities have recently gained more positive recognition in Germany, while the phenomenon of toxicity – regarding online gaming as well as social media – is surprisingly not as present in the public discourse.

Shedding more light on the phenomenon of toxicity is vital, as 24% of players in Germany are 19 years old and younger (game Verband 2025: 9) and are often unsupervised when using communication in online gaming and related platforms. Experiencing toxicity can contribute to mental health issues and even suicide (Anti-Defamation League 2019: 27). Toxic content can be met with counter speech strategies (see for example the linguistic project Hasskompass). However, it is contested whether counter speech is actually a productive tool in reducing toxicity or if it may in fact contribute to an increase in uncivil behaviour (Sponholz 2024; Liemann in this volume).

Both in public debate and in academic research, concepts such as hate speech – which only partially coincide with the phenomenon of toxicity – have increasingly gained attention (Weitzel, Mündges 2022; Sponholz 2020). The term ‘toxic’ – in its meaning relating to social behaviour – has only been introduced about a decade ago in German discourse, gaining popularity in connection to concepts like toxic masculinity (Hoffmeister 2023; Zifonun 2023). This offers the chance to facilitate an increase in public discourse on toxicity that is grounded in scientific understanding of the phenomenon, and that is therefore better equipped to be more mindful of prejudices against gaming communities and their platforms.

But what advantages does the term ‘toxicity’ offer, apart from being a keyword that can attract attention in public discourse (see Hoffmeister, Sörries-Vorberger in this volume)?

The concept of online toxicity can be useful to help identify toxic actions in the content users are exposed to or produce themselves, though the normative and evaluative character of the term must be critically reflected. The accusation of toxicity, which is repeatedly directed at gaming communities in particular, might overlook the different conceptions of communicative norms in different domains. Certain ritualized forms of supposedly toxic behaviour, like for example, reclaiming negative terms, can also be read positively in terms of the formation of group identity and a sense of community through shared communicative practices (Saleem et al. 2016). By presenting different interdisciplinary approaches that investigate the phenomenon of toxicity, especially in relation to gaming communities, we aim to create awareness about the phenomenon as well as a basis for further research and exchange. Another objec-

tive of this anthology is to counteract the blanket stigmatization of communities and their practices by examining the phenomenon on an empirical basis

In the following, we will give a brief definition of toxicity by contrasting the term with relating concepts like *hate speech*, *incivility*, *flaming*, or *shitstorms* to create a theoretical basis for the upcoming papers as well as further linguistic and interdisciplinary research. Following this introductory definition will be a concise overview of current research on the toxicity of online gaming. This chapter concludes with a brief summary of the individual contributions to this anthology.

Defining Toxicity and Relating Concepts

In German-language research on the metaphorical concept of *toxicity*, the collocations “toxic masculinity” and “toxic relationship” have been highlighted in particular (Hoffmeister 2023: 424). The former is closely linked to toxicity in gaming since “[t]oxic gamer culture is an evolution of geek masculinity” (Adams et al. 2025). Toxicity can be described as “negative behavior” (Dreier, Pirker 2023) that harms others (Hoffmeister 2023). *Toxic communication* – the focus of this volume – refers to the harmfulness of a comment to the conversation and its participants (Demus et al. 2023: 265). However, toxicity is not a static trait but develops dynamically within the communicative context (ibid.). Despite the negative connotations of toxicity, its attribution is productive, for it is only by labeling a practice as toxic that its problematic nature is recognized and exposed (Hoffmeister 2023: 447). While insults and denigration are generally regarded as disruptive factors within a society, they can also contribute to the stabilization and formation of communities (Ellerbrock et al. 2017: 3). Taking action against a shared enemy contributes to the formation of identity within a group and reinforces shared values and norms.

The concept of toxicity can be used as an umbrella term for various phenomena in online discourse that can be described as negative communication practices, such as *online backlashes*, *shitstorms* or *hate speech*. *Online backlashes* occur in the digital sphere and can be described as outbursts of anger (Fischer 2025). A *shitstorm* is characterized by the sudden appearance of a large number of negative comments directed at individuals, institutions, or companies within a short period of time, usually on social media (ibid.). The comments are often insulting and/or aggressive (ibid.). Shitstorms are targeted, public, time-limited, usually follow a clear progression, involve many participants, and are

loosely coordinated (ibid.). Individual comments within a shitstorm are generally toxic and aim to harm the target. A shitstorm is thus a specific form of toxic discourse.

Hate speech and *cyberbullying* are also specific forms of toxic communication. Gevers et al. (2022: 35) “consider toxic language to be an inclusive term, encompassing subfields such as abusive and offensive language, hate speech, and cyberbullying.” What all these phenomena have in common is that they primarily take place online. The anonymity that digital platforms typically enable allows for an increased prevalence of toxic and antisocial comments (ibid.), which is discussed under the term “Online Disinhibition Effect” (Beres et al. 2021: 3). Commenters hide behind the anonymity of the internet and often feel safe from direct (negative) social and/or legal consequences of their actions. This is also reflected in the current criticism of the coarsening of digital discourse cultures (Zollner 2022: 35). A major difference between toxicity and hate speech lies in their distinct legal implications. “While expressions of hate speech imply legal consequences, toxic language also includes utterances that cannot be prosecuted, but can still be harmful to the target.” (Gevers et al. 2022: 33) Hate speech is defined as “communication that attacks a person or a group on the basis of identity factors, such as gender, race, or religion” (Sponholz 2023) and is used consciously or intentionally (ibid.: 149). Sponholz (ibid.: 144) criticizes the overuse and thus often incorrect application of the term *hate speech*, which causes loss in the clarity of the underlying concept: “The roots of hate speech are ideologies of inequality, such as racism, sexism, homophobia, and Islamophobia, not affective action.” (ibid.: 151) While hate speech occurs frequently in the data analyzed in this volume (see for example, the texts by Niegl and Bodden, Ahmed), toxicity can incorporate more linguistic phenomena.

The concept of *invectivity* is also closely linked to toxicity. Ellerbrock et al. (2017: 3) understand invectivity as communicative aspects that have the aim to belittle, hurt, or exclude. These can be realized in various ways, namely “verbally or non-verbally, orally, in writing, through gestures or images” (ibid., translation by the authors). It is important to note that invectivity must also be perceived as such by the participants in the communication and addressed in subsequent communication (ibid.: 8). A speech act does not, therefore, have to be intended to be hurtful in order to be considered invective; rather, it is the attribution that makes it invective (ibid.: 9). Invectivity is thus always performative and must be considered within the context in which it occurs (ibid.: 4).

Antoci et al. (2016: 1) “define online incivility as a manner of offensive interaction that can range from aggressive commenting in threads, incensed discussion and rude critiques, to outrageous claims, hate speech and harassment.” Incivility as well as toxicity are umbrella terms that encompass a wide range of specific discursive practices. Incivility, however, emphasizes the aggressive dimension, whereas toxicity focuses on the harmful effects.

Toxicity can also be employed as a political and/or cultural tool. It can poison discursive spaces, hindering respectful discussion and excluding unwanted perspectives or individuals from the discourse. Ritualized toxic behavior can serve to stabilize a group and invite others to participate in toxic behavior toward certain individuals and positions, allowing participants to become part of the ingroup. This can currently be observed, for example, in anti-diversity movements, as discussed in the articles by Reinhold, Prinz and Liemann, or anti-feminist discussions, which are analyzed by Bau.

After distinguishing between various neighboring terms, we define toxicity as a communicative property of a comment or an act that does not have to be intentional and does not have to be attributed or identified as such by the participants of the situation. Even if not pointed out, toxic comments and acts nonetheless have a harmful effect on the recipient(s). Unlike hate speech toxicity is not limited to (but can include) identity-defining characteristics of communication partners. Toxicity may intentionally or unintentionally exclude individuals or groups from a discourse – either by targeting a person or group directly or by creating a generally negative atmosphere – and thus carries high political and cultural consequences. This can be observed in exclusionary or hostile practices carried out in games as well as in discourses about them.

Toxicity in Online Gaming and Gaming Discourse

Video games – like any other form of art – of course are not inherently or primarily toxic. On the contrary, they can have very positive impacts on players, helping with feelings of isolation or anxiety, and can be used for stress relief, as the global video games report “Power of Play” (Entertainment Software Association 2025) describes. Games can also tell diverse stories and allow players to empathize with identities other than their own through assuming the role of different playable characters or offer the possibility to experiment with the performance of gender (Liemann 2022). Nonetheless, gaming communities in particular are often suspected of toxicity in public discourse (e.g. Banaszczuk

2019). Just like many other communities or social media platforms, they are facing the problem of toxicity occurring in negative actions in (online) game play, live stream communication, or asynchronous gaming discourse on platforms like Steam, Discord, Reddit, Kick, Twitch, or YouTube.

The Covid-19 pandemic led to a strong shift from analogue to online activities, and both positive and negative aspects of online gaming, as well as discourses surrounding games gained more attention (Anti-Defamation League 2020) which innately led to an increase of research on the topic. Platforms and game developers (see for example, Unity 2023) also started focusing more on the phenomenon of toxicity, as communities that are perceived as very toxic will inevitably repel new players and users. Platforms are also concerned with declining revenues that can be the result of users leaving a toxic space or brands fearing negative public perception when being presented in the context of toxic content (Nobata et al. 2016: 151).

Efforts from the gaming industry are therefore united in organizations like the *Thriving in Games Group* that supports gaming platforms like Discord and Twitch, as well as developers with resources to create safer online spaces and gaming experiences. The Anti-Defamation League directly addresses the gaming industry, calling for the implementation of more policies against toxicity, and for regular transparency reports on hate and harassment (Anti-Defamation League 2025). But they also highlight that governments must take a more active role in fighting toxicity in gaming contexts by creating legal frameworks that help victims of online abuse, and they suggest specific gaming safety task forces and stronger funding of research efforts (ibid.).

However, the often limited influence game developers, publishers, and platforms have over their userbases presents a significant challenge in combating toxic behaviour. The countless hours of daily gameplay content, verbal and multimodal interactions, make a constant supervision of potentially toxic interactions impossible and also pose the risk of over-policing discourse. While community guidelines are a common tool of proactive moderation (Seering 2020) and pledges of civil behaviour can have a positive effect on users (Cheng et al. 2017: 1227), safety settings like age restrictions can easily be circumvented, and platforms often create a structure of governmentality that does not necessarily reward the strict persecution of toxic users (Groen 2020, 103). Since polarizing discussions create an increased public interest and more interaction, participation in toxic behaviour can pay off for content creators since they can gain traction, which ElSherief et al. (2018) showed both for victims and aggressors of toxicity on Twitter.

While the usage of block lists that prohibit specific toxic terms is a common practice for many online social media platforms, they are not implemented in all gaming contexts, can be easily circumvented through Algospeak (Calhoun, Fawcett 2023), and the often multimodal communication further complicates automated detection. Even with limited communicative actions, users tend to find ways to be toxic, for example through the use of emotes, toxic playstyles, or sarcastic reinterpretation of positive voice lines (see Bodden, Ahmed in this volume).

A very common tool is the flagging system for toxic players in a game or on platforms, which enables users to classify content or a person as toxic and report this to the platform or publisher. This function can not only be misused for toxic actions (e.g. 'boob policing' on Twitch, see Zolides 2023) but it can also lead to problems such as a lack of transparency in moderation actions or the exploitation of low-paid staff (Seering 2020: 5–6) who must assess toxic interactions based on flagged content within a split second. For community-based actions like flagging, there also needs to be consensus on what is considered toxic, and even when classified as such, toxicity must be considered negative in the first place. As studies (e.g. Shen et al. 2020) show, toxicity is often seen as 'part of the game' and only resilient players stay in toxic gaming communities which reinforces these behaviours: "More toxic players are more likely to become experienced, and to reach higher levels within competitive games." (Lee et al. 2021: 4) Other factors that amplify toxicity in online gaming are the competitiveness of the games (Martens et al. 2015: 6) and the ranking systems that result in higher frustration over losses. To address the problem of toxicity, there are different research approaches that especially focus on the automatic detection of toxic behaviour in the form of in-game actions as well as chat messages (see e.g. Murnion et al. 2018; for an extensive overview of current toxicity research see Anker Nexø 2024).

Of course, toxicity can be found not only within games themselves but also in the discourses surrounding them. Gaming has been instrumentalized as a battleground for right-wing and anti-diversity agendas, often in the context of supposed 'culture wars' (Proctor, Kies 2018: 127). Clear examples of this are the 2014 Gamergate movement (Braithwaite 2016), but also more recently recurrent complaints about supposedly too masculine and unattractive female characters in games, as could be seen in online discussions surrounding for example *Mass Effect: Andromeda* (2017), *Horizon Zero Dawn* (2020), *Baldur's Gate 3* (2023) or *Ghost of Yōtei* (2025) (see Reinhold, Prinz in this volume). In such cases, toxicity lies not only in the often misogynistic, racist, queerphobic, and ableist

ideologies themselves, but also within the semiotic means through which these ideologies are pushed forward and undesired points of view excluded.

Contributions

As described previously, the current German discourse focuses more on attacks on identity-defining traits connected to hate speech, which differs from international approaches that use toxicity as an umbrella term connected to online behaviour on social media platforms. *Toke Hoffmeister* and *Lars Sörries-Vorberger* therefore investigate the concept 'toxicity' with a meta-linguistic approach by conducting a qualitative survey with 434 participants to assess the conceptualization of the term 'toxisch' (toxic) in German speakers. They show that toxicity is primarily connected to people, relationships, and masculinity, contrary to the focus on online toxicity in this volume. The survey is supported by a quantitative discourse analysis of 100 media texts that show similar definitions of toxicity.

As stated before, the online disinhibition effect and competitiveness of online gaming can increase toxic behaviour. Additionally, specific interfaces and platforms contain affordances that shape *how toxic* people tend to be and *how they can perform toxicity*. The classification of different types of toxicity is therefore a focal point of *Tamara Bodden* and *Ashhab Ahmed's* article that investigates multimodal toxicity in Twitch streams. Using a stream sample that can be seen as prototypical for potential toxicity (94,600 tokens), they differentiate between auditive and visual layers of the game, stream, and chat to characterize how different multimodal resources can be used for toxic actions. These findings are supplemented with a discourse analysis of both public meta discourse on toxic streamers and streamers community guidelines.

The public discourse around toxicity in gaming is also the empirical base for *Patricia Bau's* linguistic analysis of the discourse surrounding the game *No Mercy* (2025) that turns the taboos of incest and rape into playable experiences. Bau shows how taboos and toxicity are entwined in the discourse about the game and how toxic practices (like moralizing, whataboutism, and accusations of attention seeking), both from supporters and opponents of the game prevent productive discussion.

The exploration of the boundaries of what can be said is also a topic in a section of three contributions that focus on right-wing discourse connected with gaming.

Christina Liemann investigates toxic discursive practices in the Steam Group ‘Woke Content Detector’, which aims to identify and warn of supposedly woke games. Here, Detectors, who categorize the games, and Counter Speakers, who aim to disrupt the group’s echo chamber, seemingly engage in mutual discussion, yet no conclusion or reconciliation is reached. In her linguistic study, Liemann shows how antagonistic personal references – as a more subtle form of toxicity – create this unproductive, poisoned discursive environment.

In his analysis, *Andreas Niegl* focuses on the discourse inside alt-history communities regarding the game *Hearts Of Iron IV* (2016) on platforms like Steam, Reddit, and Kiwifarms. He discusses how role-playing and post-irony as community-specific practices blur the lines of acceptable speech and open the forums to actual non-ironic support of fascist ideologies. While there is a notable difference in moderated and unmoderated forums, interventions against toxicity can lead to factions in the community and users switching to platforms that tolerate more toxic behaviour.

The article by *Julia Ueberlacker* and *Mick Prinz* from the Amadeu Antonio foundation examines these phenomena from a meta-discursive standpoint and describes why and how gaming communities and platforms can be effective in recruiting new members for far-right movements. Though the authors also stress the diverse nature of gaming communities, they emphasize the importance of understanding far-right codes, mechanisms, and structures to prevent gamers from unintentionally drifting into these spheres. The article additionally discusses various programs, potential for interventions, and the necessity of platform governance.

Lastly, *Marion Plank* provides insight into toxicity in the field of game development where toxic practices are already rooted in the stage of education. She examines the phenomenon of ‘crunching’ as an often idealized but toxic concept in both game development, and game development education, which is masked under the disguise of idealized ‘industry-readiness’. Through guided interviews with both educators and students Plank explores this toxic practice and outlines five principles for a healthier curriculum.

Bibliography

Adams, Phillipa R.; Sommers, Maria; Scholl, Ben; Lusoli, Alberto; Chun, Wendy Hui Kyong (2025): The revenge of ‘tribal’ toxic masculinity: hate raids on

- Twitch. In: *Journal of Visual Culture*. URL: <https://doi.org/10.1177/14704129251327916>.
- Anker Nexø, Louise (2024): Toxic Behaviours in Esport: A Review of Data-Collection Methods Applied in Studying Toxic In-Gaming Behaviours. In: *International Journal of Esports*. 1, 1. URL: <https://www.ijesports.org/article/127/html> [accessed 03.02.26].
- Anti-Defamation League (2019): Free to Play? Hate, Harassment and Positive Social Experiences in Online Games 2019. In: ADL. URL: <https://www.adl.org/resources/report/free-play-hate-harassment-and-positive-social-experiences-online-games> [accessed 24.03.26].
- Anti-Defamation League (2020): Free to Play? Hate, Harassment and Positive Social Experiences in Online Games 2020. In: ADL. URL: <https://www.adl.org/resources/report/free-play-hate-harassment-and-positive-social-experience-online-games-2020> [accessed 24.03.26].
- Anti-Defamation League (2025): Hate and Harassment Present in Almost Half of Online Multiplayer Gaming Sessions Tested, New ADL Research Finds. In: ADL. URL: <https://www.adl.org/resources/press-release/hate-and-harassment-present-almost-half-online-multiplayer-gaming-sessions> [accessed 24.03.26].
- Antoci, Angelo; Delfino, Alexia; Paglieri, Fabio; Panebianco, Fabrizio; Sabatini, Fabio (2016): Civility vs. Incivility in Online Social Interactions: An Evolutionary Approach. In: *PLoS ONE*. 11, 11. P. 1–17. URL: <https://doi.org/10.1371/journal.pone.0164286>.
- Banaszczuk, Yasmina (2019): Toxic Gaming. Rassismus, Sexismus und Hate Speech in der Spieleszene. In: *Aus Politik und Zeitgeschichte*. Ed. by Bundeszentrale für politische Bildung. URL: <https://www.bpb.de/shop/zeitschriften/apuz/294444/toxic-gaming/> [accessed 07.05.26].
- Bareither, Christoph (2016): Gewalt im Computerspiel. Facetten eines Vergnügens. Bielefeld: transcript.
- Beres, Nicole A., Frommel, Julian, Reid, Elizabeth, Mandryk, Regan L.; Klarkowski, Madison (2021): Don't You Know That You're Toxic: Normalization of Toxicity in Online Gaming. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Yokohama, Japan. ACM New York, NY, USA. P. 1–15. URL: <https://doi.org/10.1145/3411764.3445157>.
- Calhoun, Kendra; Fawcett, Alexia (2023): “They Edited Out her Nip Nops”: Linguistic Innovation as Textual Censorship Avoidance on TikTok. In: *Language@Internet*. 21, 1. P. 1–30.

- Cheng, Justin; Bernstein, Michael; Danescu-Niculescu-Mizil, Christian; Leskovec, Jure (2017): "Anyone can become a troll: Causes of trolling behavior in online discussions". In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW '17. New York, NY, USA: Association for Computing Machinery. P. 1217–1230. URL: <https://doi.org/10.1145/2998181.2998213>.
- Demus, Christoph; Schütz, Mina; Probol, Nadine; Pitz, Jonas; Siegel, Melanie; Labudde, Dirk (2022): Hass im Netz – Aggressivität und Toxizität von Hasskommentaren und Postings, Detektion und Analyse. In: Handbuch Cyberkriminologie. Ed. by Thomas-Gabriel Rüdiger and Petra Saskia Bayerl. Springer VS, Wiesbaden. URL: https://doi.org/10.1007/978-3-658-35450-3_13-1.
- Dentsu (2024): 2024: State of Gaming Report. In: Dentsu. URL: https://www.dentsu.com/reports/n2024_state_of_gaming_report [accessed 24.03.26].
- Dreier, Lukas; Priker, Johanna (2023): Toxicity in Twitch Live Stream Chats: Towards Understanding the Impact of Gender, Size of Community and Game Genre. In: 2023 IEEE Conference on Games (CoG). URL: <http://dx.doi.org/10.1109/CoG57401.2023.10333159>.
- Ellerbrock, Dagmar; Koch, Lars; Müller-Mall, Sabine; Münkler, Marina; Scharloth, Joachim; Schrage, Dominik; Schwerhoff, Gerd (2017): Invektivität Perspektiven eines neuen Forschungsprogramms in den Kultur und Sozialwissenschaften. In: Kulturwissenschaftliche Zeitschrift. 2, 1. P. 2–24.
- ElSherief, Mai; Nilizadeh, Shirin; Nguyen, Dana; Vigna, Giovanni; Belding, Elizabeth (2018): Peer to Peer Hate: Hate Speech Instigators and Their Targets. In: Proceedings of the ICWSM 2018, 12, 1. URL: <https://arxiv.org/pdf/1804.04649v1> [accessed 25.04.26].
- Entertainment Software Association (2025): Power of Play 2025 GLOBAL VIDEO GAMES REPORT. URL: <https://www.theesa.com/resources/the-global-power-of-play-report/> [accessed 24.03.26].
- Fischer, Friederike (2025): Shitstorm. In: Diskursmonitor. Glossar zur strategischen Kommunikation in öffentlichen Diskursen. Ed. by Forschungsgruppe Diskursmonitor und Diskursintervention. URL: <https://diskursmonitor.de/glossar/shitstorm> [accessed 24.03.26].
- Fortuna, Paula; Soler, Juan; Wanner, Leo (2020): Toxic, hateful, offensive or abusive? What are we really classifying? An empirical analysis of hate speech datasets. In: Proceedings of the Twelfth Language Resources and Evaluation Conference, Marseille. P. 6786–6794.

- game – Verband der deutschen Games-Branche e.V. (2025): Jahresreport der deutschen Games-Branche 2025. In: game – Verband der deutschen Games-Branche e.V. URL: <https://www.game.de/publikationen/jahresreport-2025/> [accessed 24.03.26].
- Garbe, Andreas (2020): Killerspiele. In: Handbuch Gameskultur. Über die Kulturwelten von Games. Ed. by Olaf Zimmermann and Felix Falk. Berlin: Deutscher Kulturrat e. V. P. 193–197.
- Gevers, Ine; Markov, Ilia; Daelemans, Walter (2022): Linguistic Analysis of Toxic Language on Social Media. In: Computational Linguistics in the Netherlands. 12. P. 33–48. URL: <https://clinjournal.org/clinj/article/view/146> [accessed 24.03.26].
- Groen, Maike (2020): Digital Governmentality: Toxicity in Gaming Streams. In: Games and Ethics. Digitale Kultur und Kommunikation. Theoretical and Empirical Approaches to Ethical Questions in Digital Game Cultures. Ed. by Maite Groen, Nina Kiel, Angela Tillmann, André Weßel. Wiesbaden: Springer VS. URL: https://doi.org/10.1007/978-3-658-28175-5_7.
- Habermas, Jürgen (2006): Ein avantgardistischer Spürsinn für Relevanzen. Was den Intellektuellen auszeichnet. In: Blätter für deutsche und internationale Politik. 5/2006. P. 551–557.
- Hoffmeister, Toke (2023): Von toxischen Substanzen und toxischer Männlichkeit. Eine korpuspragmatische Studie zur Erschließung eines Gegenwartsphänomens. In: Zeitschrift für Germanistische Linguistik. 51, 3. P. 424–458.
- Lee, Matthew; Johnson, Daniel; Tanjitpiyanond, Porntida; Louis, Winnifred R. (2021): It's habit, not toxicity, driving hours spent in DOTA 2. In: Entertainment Computing. 41, 100472. URL: <https://doi.org/10.1016/j.entcom.2021.100472>.
- Liemann, Christina (2022): Gendered Gazes in Games – Die (De-)Konstruktion von Geschlecht durch die First-Person-Kamera am Beispiel von ‚Cyberpunk 2077‘. In: PAIDIA – Zeitschrift für Computerspielforschung. URL: <https://paidia.de/gendered-gazes-in-games-die-de-konstruktion-von-geschlecht-durch-die-first-person-kamera-am-beispiel-von-cyberpunk-2077/> [accessed 07.05.26].
- Mall, Raghvendra; Nagpal, Mridul; Salminen, Joni; Almerexhi, Hind; Jung, Soon-Gyo; Jansen, Bernard J. (2020): Four Types of Toxic People: Characterizing Online Users' Toxicity over Time. New York: Association for Computing Machinery. URL: <https://doi.org/10.1145/3419249.3420142>.

- Martens, Marcus; Shen, Siqi; Iosup, Alexandru; Kuipers, Fernando (2015): Toxicity detection in multiplayer online games. In: 2015 International Workshop on Network and Systems Support for Games (NetGames), P. 1–6. URL: <https://doi.org/10.1109/NetGames.2015.7382991>.
- Montag, Annegret (2021): Zum Einsatz von Computerspielen im Unterricht. Ein aktueller Forschungsstand. In: "Loading..." Game Studies Interdisziplinär. Ed. by Tamara Bodden, Annegret Montag and Marvin Madeheim. Paderborn: Fink. P. 91–106.
- Murnion, Shane; Buchanan, William; Smales, Adrian; Russell, Gordon (2018): Machine learning and semantic analysis of in-game chat for cyberbullying. In: COMPUTERS & SECURITY. 76. P. 197–213. URL: <https://doi.org/10.1016/j.cose.2018.02.016>.
- Nobata, Chikashi; Tetreault, Joel; Thomas, Achint; Mehdad, Yashar; Chang, Yi (2016): Abusive language detection in online user content. In: Proceedings of the 25th International Conference on World Wide Web, ser. WWW '16. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, P. 145–153. URL: <https://doi.org/10.1145/2872427.2883062>.
- Proctor, Willian; Kies, Bridget (2018): Editors' Introduction: On toxic fan practices and the new culture wars. In: Participations. Journal of Audience & Reception studies. 15, 1. P. 127–142.
- Saleem, Haji M.; Dillon, Kelly P.; Benesch, Susan; Ruths, Derek (2016): A web of hate: Tackling hateful speech in online social spaces. Proceedings of Text Analytics for Cybersecurity and Online Safety (TA-COS), LREC Portoroz. P. 1–9.
- Seering, Joseph (2020): Reconsidering community self-moderation: The role of research in supporting community-based models for online content moderation. In: Proceedings of the ACM on Human-Computer Interaction. 4, 107. P. 1–28. <https://doi.org/10.1145/3415178>.
- Shen, Cuihua; Sun, Qiusi; Kim, Taeyoung; Wolff, Grace; Ratan, Rabrindra; Williams, Dmitri (2020): Viral vitriol: Predictors and contagion of online toxicity in World of Tanks. Computers in Human Behavior, 108, P. 1–9. URL: <https://doi.org/10.1016/j.chb.2020.106343>.
- Sponholz, Liriam (2020): Der Begriff „Hate Speech“ in der deutschsprachigen Forschung. Eine empirische Begriffsanalyse. In: SWS-Rundschau. 60. P. 43–65.
- Sponholz, Liriam (2023): Hate speech. In: Challenges and perspectives of hate speech research. Ed. by Christian Strippel, Sünje Paasch-Colberg, Martin

- Emmer and Joachim Trebbe. Berlin: Digital Communication Research. P. 143–163. URL: <https://doi.org/10.48541/dcr.v12.9>.
- Sponholz, Liriam (2024): Hate Speech und Inzivilität: Gegenmaßnahmen jenseits von Gegenrede. *Beyond Countering: Ansätze zur Bekämpfung von Hate Speech und Inzivilität*. In: Bundeszentrale für politische Bildung. URL: <https://www.bpb.de/themen/medien-journalismus/soziale-medien/545938/hate-speech-und-inzivilitaet-gegenmassnahmen-jenseits-von-gegenrede/> [accessed 23.04.26].
- Unity (2023): Toxicity in Multiplayer Games Report. In: Unity. URL: <https://create.unity.com/toxicity-report> [accessed 24.03.26].
- Weitzel, Gerrit; Mündges, Stephan (2022): Hate Speech. Definitionen, Ausprägungen, Lösungen. Wiesbaden: Springer VS. URL: <https://doi.org/10.1007/978-3-658-35658-3>.
- Zifonun, Gisela (2023): Das Toxische und die Männlichkeit. *Sprachreport*, 39, 3, P. 1–6.
- Zolides, Andrew (2023): Toxic community policing: Weaponizing moderation tools on Twitch. In: *Real Life in Real Time: Live Streaming Culture*. Ed. By Johanna Brewer, Bo Ruberg, Amanda L. L. Cullen, Christopher J. Persaud. Cambridge, MA: MIT Press. P. 131–144. URL: <https://doi.org/10.7551/mitpress/14526.003.0013>.
- Zollner, Sebastian (2022): Counter Speech als sprachlich-kommunikative Praktik in digitalen, invektiven Konstellationen. Ein Thema für Linguistik, Medienpädagogik und politische Bildung. In: *merz. medien + erziehung*. 66, 2. P. 35–45.

"Megan Fox is in a toxic relationship"

Representations of the 'Toxic' in Linguistic Community and Media Discourse

Toke Hoffmeister & Lars Sörries-Vorberger¹

Abstract The study investigates the figurative use of the term 'toxic' in contemporary German discourse. Beyond chemical and financial senses, the *Oxford English Dictionary* also records a figurative meaning in which 'toxic' is used in connection with mental states and social relations. A questionnaire study with speakers and the analysis of 100 online texts (69,081 tokens, 2022–2023) examine usage contexts, evaluations, and collocational patterns. The results reveal the semantic extension of 'toxic', highlight associated language attitudes, and demonstrate its role in the construction of social roles and identities.

Keywords Toxic, Attitudes, Usage Data, Discourse Analysis, Corpus Study

Abstract Die Studie untersucht die figurative Verwendung des Ausdrucks ‚toxisch‘ im zeitgenössischen deutschen Diskurs. Neben chemischen und finanztheoretischen Bedeutungen beschreibt das *Oxford English Dictionary* auch ebenjene figurative Bedeutung, nach der ‚toxisch‘ in Verbindung mit mentalen Zuständen und sozialen Beziehungen gebraucht wird. Eine Fragebogen-Studie mit Sprecher*innen und die Analyse von 100 Online-Texten (69.081 Tokens, 2022–2023) beleuchten Verwendungskontexte, Bewertungen und Kollokationen. Die Ergebnisse erschließen die Bedeutungserweiterung von ‚toxisch‘ und zeigen die verbundenen Spracheinstellungen sowie seine Funktion bei der Konstruktion sozialer Rollen und Identitäten auf.

1 Shared first-authorship in alphabetical order.

Schlagworte Toxisch, Haltungen, Gebrauchsdaten, Diskursanalyse, Korpusstudie

1. Toxic: A new attribute to describe social relations

As early as 2003 Britney Spears sang in her song *Toxic*: “With a taste of your lips, I’m on a ride // You’re toxic, I’m slipping under // With a taste of a poison paradise // I’m addicted to you // Don’t you know that you’re toxic” and, almost twenty years later and after becoming widely known via TikTok, the musician BoyWithUke sings in his equally named song: “All my friends are toxic, all ambitionless // So rude and always negative”. It is clear, not only from these examples, that the attribute is intended to convey something negative, a form of harm or threat. This negative sense is already contained in the meaning of the lexeme ‘toxic’: a body or object is adversely affected, endangered, or destroyed through the administration of a toxic influence. The *Oxford English Dictionary* (OED 2025) records four distinct semantic dimensions in total: first, the – one might call it – traditional chemical meaning, namely “poisonous, harmful or dangerous to health or life”; secondly, „caused by or attributable to a toxic substance or substances”; thirdly, a technical, domain-specific meaning from the financial sector, which describes “[d]esignated securities which are based on debt which has a high risk of default, and for which there is not a healthy or functioning market.” Finally, the OED lists a fourth, figuratively marked sense: “[t]hat has a harmful effect on a person’s mental or emotional state, or on the social relations of a group, esp. in a pervasive or insidious way; extremely unhealthy, hostile, or negative; destructive or pernicious.” It is this fourth semantic dimension that is the subject of the present article.

We address the question of how this semantic specification is manifested in discourse and how the adjective ‘toxic’ is used concretely there: on the one hand, in the discourse of language users themselves – where we are interested in metalinguistic reflection – and on the other hand, in the media discourse – where we focus in particular on collocational patterns. The first study includes a questionnaire survey in which 434 speakers are asked extensively about their perception, their use, and generally their associations with the term ‘toxic’. In addition, the participants are asked to rate ‘toxic’ based on five seven-point

semantic differentials² (including correct–incorrect, formal–informal). The second contains a corpus analysis in which 100 online texts (69,081 tokens) from German magazines and online portals (including *Brigitte*, *ELLE*, *Gala*, *Märkische Allgemeine*, *Men's Health*, *Spieletipps.de*, *Stern*, *T-Online*, etc.) from the years 2022 and 2023, which deal with the topic of 'toxicity' in a non-chemical and non-medical context, are quantitatively analyzed. Through this examination of larger quantities of comparable texts from different media, access to the media-mediated discourse can be gained (Baker 2023). This will be flanked by individual qualitative approaches, which appear to be particularly useful with regard to role and identity constructions associated with 'toxic' (Motschenbacher 2018).

Initial contributions (Hoffmeister 2023a; b; Zifonun 2023) address this question from different perspectives. Hoffmeister (2023a) sheds light on the extension of the meaning of 'toxic', elaborates on various domains in which the labeling of something as toxic is found, and analyses the use of 'toxic' in newspaper texts. Zifonun (2023) addresses toxic masculinity in particular, asking how 'toxic masculinity' is used (Zifonun 2023: 3). This article takes these two studies as its starting point and asks: How is 'toxic' used in general in the German language community and media discourse? These papers mark some of the earliest scholarly attempts to address a phenomenon that has, to date, received little systematic attention within German-language research. Consequently, toxicity as a disciplinary concept has so far occupied only a marginal position in the German-speaking academic discourse.

The paper is structured as follows: In section 2, we briefly outline the theoretical background of our argumentation – here we pay special attention to the notion of 'representations' of the toxic and locate our work within socio-pragmatic discourse linguistics. We then present the results of two studies (Section 3): firstly, the usage- and attitude-study within the speech community (Section 3.1), secondly, the corpus study (Section 3.2). Section 4 summarizes and discusses the findings and offers an outlook.

2 Semantic differentials (Osgood et al. 1957) are psychometric measurement instruments that use bipolar adjective pairs as endpoints of rating scales. Respondents are asked to indicate the perceived position of a given concept along the continuum between the opposing terms (e.g., helpful vs. useless).

2. What we mean by ‘representations’ of the toxic

The concept of ‘representation’ is multifaceted and used across linguistic and non-linguistic disciplines, often with an attributive qualifier such as linguistic or mental. In this study, we adopt a socio-pragmatic discourse-linguistic perspective, where the concept has received relatively little attention (Spieß 2025: 115). Its central contribution lies in linking socio-constructivist, collective perspectives with mental, cognitive, individual-level perspectives. Linguistic phenomena – including the adjective ‘toxic’ – are treated as actualizations or consequences of mental representations, which are not directly accessible but must be inferred from usage (ibid.: 116–117).

Following Spieß (2025), such representations are:

1. a ‘work of representation’ (Hall 1997),
2. non-intentional and non-reflective,
3. based on a common ground and thus constructive,
4. dynamic processes of knowledge negotiation, and
5. discursive signifying practices embedded in and generating contexts.

Representations are social constructs crucial to meaning-making through social practices (Spieß 2025: 117). They require active interaction to negotiate knowledge and meaning (1, 4), are neither intentional nor reflective (2), as meaning emerges from socio-constructivist rather than individual or object-intrinsic processes. Meaning-making is a cooperative endeavor, dependent on a common ground (3) for coordinating semantic components. Processes (1)–(5) constitute signifying practices that both shape and are shaped by diverse contexts (‘Lebenswelten’ (lifeworlds), usage contexts, co-texts). When we investigate ‘representations’ of the toxic, we are mainly interested in usage and evaluation data.

3. What people and media mean by labeling something as ‘toxic’

The following section discusses two studies that aim to provide a multidimensional reconstruction of the meaning of the attribute ‘toxic’ and, in doing so, examines its representation (in the sense outlined in Section 2).

3.1 Usage and evaluation in the linguistic community (study 1)

In our first study, we investigate the following six research questions, considering the areas perception, meaning, production and evaluation:

- 1. Where have people heard someone using the label ‘toxic’? (perception)
- 2. How do people understand the label ‘toxic’? (perception)
- 3. What do people mean by labeling something as ‘toxic’? (meaning)
- 4. Which different entities do they label as ‘toxic’? (production)
- 5. What associations are connected to the attribute? (production)
- 6. How do people evaluate the labeling of something/someone as ‘toxic’? (evaluation)

For this purpose, we conducted an online questionnaire study between September 2023 and December 2023, comprising a total of 10 questions and completed by n = 434 German speaking participants from all over the German speaking area. The questionnaire was distributed via social media and email lists. Among these 434 participants, 298 identified as female (68.66 %), 120 as male (27.65 %), 11 as non-binary or diverse (2.53 %), and five provided no information (1.15 %). The distribution of participants’ educational background is as follows:

Tab. 1: Information on the highest educational qualification

Highest Educational Qualification	Absolute Frequency	Relative Frequency
University degree	196	45.16 %
General higher education entrance qualification ('Abitur')	115	26.50 %
Doctorate (PhD)	48	11.06 %
Secondary school diploma ('Mittlere Reife')	30	6.91 %
Vocational training	9	2.07 %
Postdoctoral qualification ('Habilitation')	8	1.84 %

Highest Educational Qualification	Absolute Frequency	Relative Frequency
No response	28	6.45 %
Total	434	

Overall, the data reveal an education bias, which can be traced to the study’s distribution channels. This should be taken into account when interpreting the results.

The age distribution is presented as shown in table 2. The median year of birth is 1992.

Tab. 2: Age distribution

Decade of Birth	Absolute number (relative number)
1930s	2 (0.46 %)
1940s	5 (1.15 %)
1950s	12 (2.76 %)
1960s	35 (8.06 %)
1970s	40 (9.22 %)
1980s	81 (18.66 %)
1990s	153 (35.25 %)
2000s	100 (23.04 %)
No entry	6 (1.38 %)
Total	434

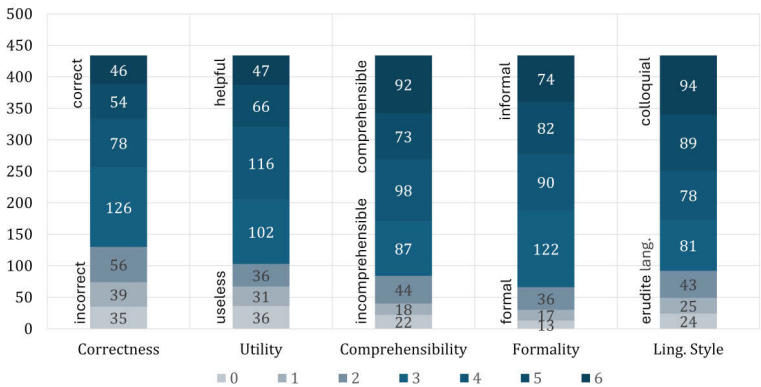
Accordingly, the participant group can be characterized as predominantly female, relatively young, and highly educated. All answers were inductively encoded, following a Qualitative Data Analysis (QDA) paradigm (Kuckartz 2014; Mayring 2014; 2015). The responses were systematically collected and coded according to semantic criteria. These coded responses were subsequently consolidated into broader clusters, representing groups of thematically related answers resulting in a structure of meta- and subcodes that

together provide a comprehensive picture of production, perception, and evaluation. The categories were subsequently checked and adjusted, if necessary, after assigning the answers. In summary, the study uses an approach between inductive and deductive methods.

When asked whether they were familiar with the use of the term ‘toxic’ to describe something –that is, whether they had encountered it before – 98 % of participants responded affirmatively. But only 27.87 % are using the attribute themselves quite regularly or regularly. On the other hand, 53.92 % are using the attribute never or very rarely and 18.20 % are using it sometimes.

The evaluation of this attribute (fig. 1) was conducted using seven-point semantic differentials (0–6) across the following dimensions: correctness (correct–incorrect), utility (helpful–useless), comprehensibility (comprehensible–incomprehensible), formality (informal–formal), and linguistic style (colloquial–erudite language). Overall, participants rated the attribute ‘toxic’ as somewhat correct (mean = 3.2), rather helpful (3.4), neither particularly comprehensible nor incomprehensible (3.0), somewhat informal (3.8), and quite clearly tending toward colloquial (3.9).

Fig. 1: Overview of the evaluation based on semantic differentials



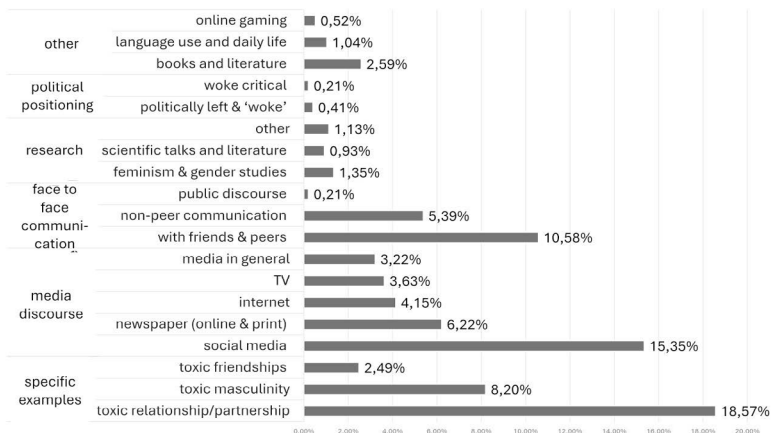
Responses to the first question ‘Where have you encountered the term ‘toxic?’ revealed a wide range of contexts (cf. fig. 2). The term is frequently linked with specific examples such as ‘toxic relationships’, ‘masculinity’, or ‘friendships’, and appears predominantly in social media discourse. Occurrences in newspapers, on the internet more generally, or in television and other

media are reported much less frequently. Furthermore, the data indicate that the term is primarily heard in peer-to-peer communication (cf. fig. 2) and – especially relevant for this volume – some, albeit very few, people connect ‘toxic’ with online gaming (0.52 %). 3.41 % of the people have heard of ‘toxic’ in research, especially in feminism & gender studies (1.35 %) and scientific talks and literature in general (0.93 %). For a small group of people ‘toxic’ is linked to a specific politic ideology, namely a rather left, referred to as ‘woke’ ideology.

The specific answers (originally in German, translated by TH/LSV) show the focus on media on the one hand (cf. also our second study in section 3.2) and the focus on public communication and the linkage to specific examples such as ‘toxic masculinity’ on the other:

1. “In the media, for example in articles in Süddeutsche or ZEIT, there is often talked of ‘toxic’ relationships or even ‘toxic’ masculinity. Just recently, an article in Süddeutsche (I think) was titled: ‘And what if I’m toxic myself?’ Or something like that”
2. “Mainly in the context of ‘toxic masculinity/relationships’. Occurs verbally and in writing; public communication rather than private conversations.”
3. “You hear/read it very often in the news, newspapers, etc. Toxic masculinity is mentioned particularly often. I can’t think of any specific individual situations in which I’ve noticed it.”

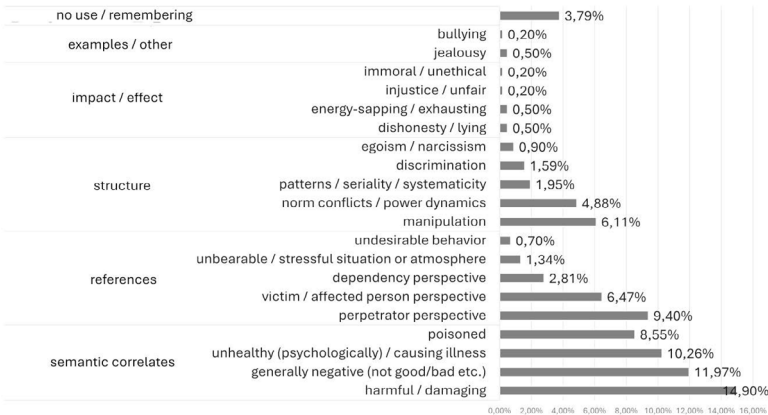
Fig. 2: Contexts in which participants have encountered the label ‘toxic’



The second question was: ‘How have you understood the utterance? What do you think was intended to be expressed by this (cf. fig. 3)?’

Most respondents mention semantic correlates (45.68 % in total) such as harmful/damaging (14.9 %), generally negative features (such as not good/bad, 11.97 %), unhealthy in a psychological manner or causing illness (10.26 %), and poisoned (8.55 %). Multiple responses naturally occurred in the open-text answers. In addition, a specific perspectivization is also taken via references (20.72 % in total), for instance by emphasizing the perpetrator’s (9.4 %) or the victim’s perspective (6.47 %). We also find perspectivizations that highlight dependency between people (2.81 %) or an unbearable/stressful situation or atmosphere (1.34 %). A third major thematic block concerns structural characteristics (15.43 % in total): here, the focus is primarily on manipulation (6.11 %) and norm conflicts or power dynamics (4.88 %) underlying the attribute ‘toxic’. Further structural features include patterns, seriality and repetition, discrimination, as well as egoism and narcissism (cf. fig. 3).

Fig. 3: Perception and interpretation of ‘toxic’



In more detail, participants in the questionnaire reported the following on how they understood the attribute:

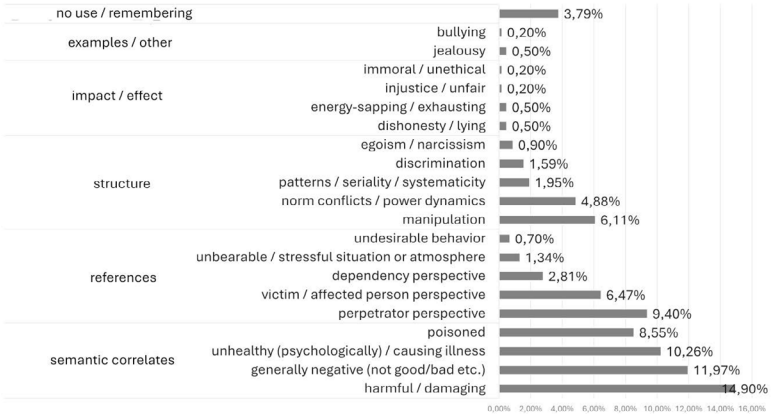
- 4. “‘Toxic masculinity’ refers to behavior that is ultimately based on conservative and mostly heteronormative, socially constructed ideas of gender re-

lations: men/people read as masculine are expected, for example, to provide for the family, protect the woman, not show (have?) any feelings, be ‘strong’. This can result, for example, in men who want to meet these supposed (normative) expectations of their own behavior, acting possessive, domineering, dominant towards women, or not allowing themselves to be affected by emotions such as sadness, despair or similar. [...]“

5. “The situation is unbearable, strong negativity emanates from a person or a group, destructive attitude, poor work organization, no leadership, hurtful remarks, bullying”

Thirdly, we asked what people themselves mean by labeling something as ‘toxic’ (cf. fig. 4). Most respondents, who use ‘toxic’, refer to negative and threatening behavior or behavior patterns (14.93 %). Moreover, they mean an unhealthy or hurting (7.81 %), destructive, dysfunctional or damaging (7.64 %) process, they state negative consequences that come with a certain kind of behavior that is labeled as ‘toxic’ (7.29 %) or simply mean something not good or bad (6.94 %). Several other semantic features <6.0 % can be retraced in fig. 4.

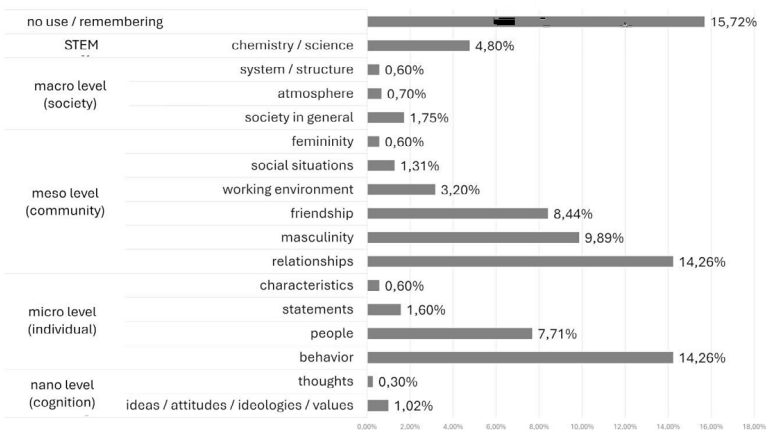
Fig. 4: Meaning attributed to the label ‘toxic’



Concrete examples of people's reasons to use 'toxic' are:

- 6. "Because it's a great word :-)) as I said, I wanted to express that something is not 'healthy', but somehow poisoned. That it's the way it shouldn't be or the way you don't imagine it to be in an ideal state, perhaps"
- 7. "The connotation of the use of 'toxic' in comparison to, for example, 'destructive', 'harmful' would be the reference to a 'modern' discourse within which a great sensitivity to power imbalances plays a central role."
- 8. "I wanted to specify the balance of power and make it clear that there is a situation that cannot be fruitful for good relationships. Sometimes to warn people, sometimes to make myself aware of it. However, I only use the term when many things come together and the situation is really extreme."

Fig. 5: Entities labeled as 'toxic'



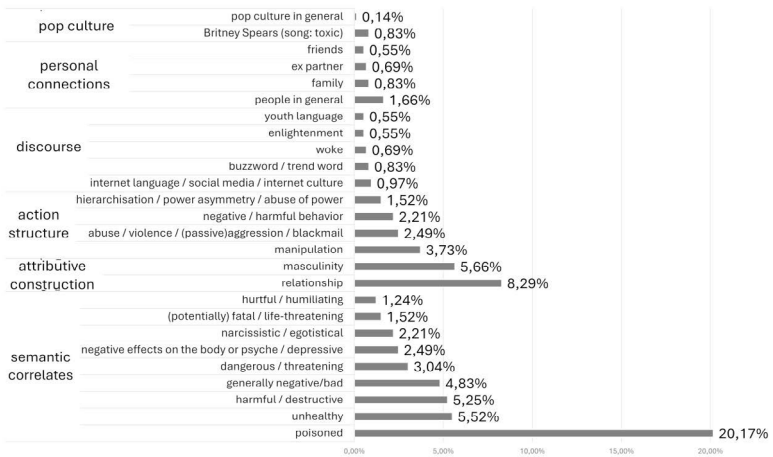
It has already become evident that masculinity and relationships in particular play a significant role in relation to toxicity and are frequently attributed as such. However, it remains unclear what the quantitative proportions of these designations are and which other entities may also be conceived of as being described as toxic. Fig. 5 lists those entities that the respondents identified as toxic. While 15.72 % reported never having used the term in this way themselves, relationships (14.26 %) – together with behavior – represent the entities most frequently attributed as toxic. Masculinity ranks third with 9.89 %. Fur-

thermore, friendships (8.44 %) and people in general (7.71 %) also play a substantial role. In addition, there are further entities mentioned by fewer than 5.0 % of respondents, including objects from chemistry and the natural sciences. We can thus conclude that ‘toxic’ has become established in everyday language as an attribute used to describe social relations.

Finally, we asked participants which cognitive associations they have with the expression ‘toxic’ (fig. 6). One fifth of the respondents indicated the meaning ‘poisonous’ as their association. Relationships were mentioned by 8.29 % of participants, making it the second most frequent association. Further associations include masculinity (5.66 %), unhealthy (5.52 %), harmful/destructive (5.25 %), and generally negative or bad (4.83 %). These associations once again illustrate the semantic shift of ‘toxic’ from a technical term originating in the domains of chemistry and medicine (Hoffmeister 2023a) toward a social attribute.

In addition, a small number of respondents (0.83 %) – as already noted in the introduction to this paper – also referred to Britney Spears and her song *Toxic*.

Fig. 6: Cognitive associations connected to ‘toxic’



Overall, the findings demonstrate that ‘toxic’ has undergone a notable semantic and pragmatic shift from a technical term within the natural sciences to a socially and discursively embedded evaluative attribute. Participants pre-

dominantly encounter the term in everyday and media discourse – especially in connection with relationships, masculinity, and behavior – and tend to interpret it as denoting harmful, unhealthy, or destructive patterns within social interaction. The associations and self-reported uses further underscore the term's function as a moral and relational qualifier, reflecting contemporary sensitivities to power dynamics, psychological well-being, and norm conflicts. This extension from its original chemical and medical meaning to the social sphere illustrates an ongoing process of metaphorical transfer and semantic broadening characteristic of modern public discourse.

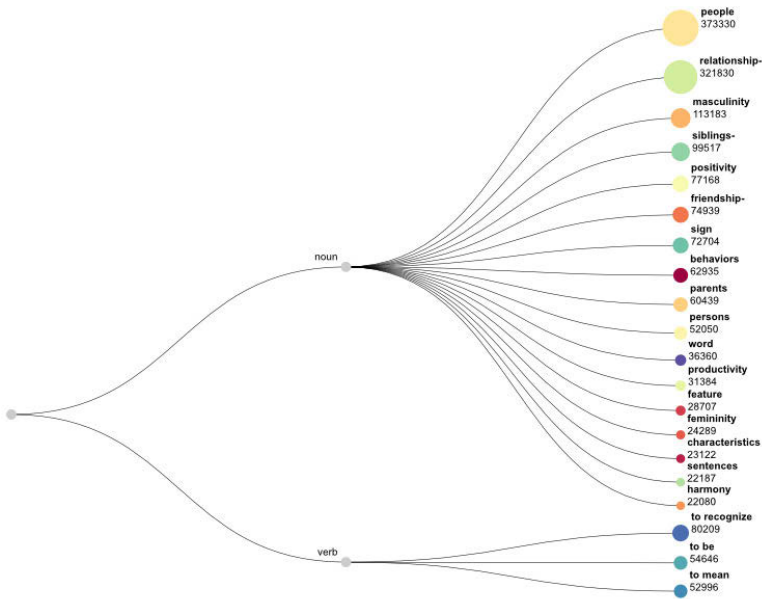
3.2 The media discourse – a corpus-based approach (study 2)

The second analysis is conducted as a corpus-based discourse analysis (e.g. Kennedy 2014; Lemnitzer, Zinsmeister 2015, for corpus linguistics, and e.g. Schiffrin et al. 2005; Busse, Teubert 1994; Spitzmüller, Warnke 2011, for discourse analyses). It is based on a corpus comprised of 100 texts that are publicly available online. To compile the corpus, a search was conducted on Google News using the keywords 'Toxizität' and 'toxisch' on July 16, 2023. This approach ensured that the selected media articles were both current and, above all, pertinent. The first 100 articles addressing social toxicity were collected and processed. The corpus represents therefore an excerpt from the media discourse as a network of statements about social toxicity. The selected texts were published between September 30, 2022, and July 9, 2023, and originated from various media outlets (including *Brigitte*, *Business Insider*, *Glamour*, *Mein MMO*, *Myself – das Magazin für starke Frauen*, *Neue Zürcher Zeitung*, *Stern*, *Stuttgarter Nachrichten*, and *WMN – Das Lifestyle-Magazin*, $n > 2$), totaling 69,081 tokens. Within this corpus, forms of the word 'toxic' (such as 'toxisch', 'toxische', 'toxischer') appear 1,014 times.

The analysis employed AntConc (Anthony 2022) to identify collocations within a five-word window to the left and right of the word forms of 'toxic', which were then summarized (e.g. singular and plural forms) and sorted by their likelihood (i.e., their statistical probability of collocation), with those exceeding a likelihood threshold of 22,000 being selected for detailed examination. Complementing this quantitative approach, a qualitative approach was chosen, in which the texts were subjected to close reading and the evidence for individual collocations was examined more closely (e.g. Motschenbacher 2018). To maintain analytical clarity, and due to the research interest, contexts involving 'toxic' combined with pronouns, articles, adverbs, or adjectives were

excluded. The results can be seen in fig. 7 (created with RawGraphs; the hyphen indicates that word forms, mostly declensions, have been combined).

Fig. 7: Collocations with ‘toxic-‘ (sorted by word class and likelihood)



The most frequent context – in the sense of probability of collocation – for ‘toxic’ describes it as an attribute of people in general (likelihood=373,330, frequency=258; ‘Mensch’, ‘Menschen’ mostly in plural) – in both predicative and attributive position. These contexts often involve broad statements about toxic people or general ways to deal with them, such as “Toxic people are jealous and envious” (DP 05/23), or “Recognizing toxic people: Common signs at a glance” (NB 11/22).³ The second major context pertains to relationships, particularly ro-

3 Here and in subsequent sections, the cited examples are original texts translated by the two authors. Due to the size of the corpus, not all sources can be listed in detail. For reference, the medium and date are provided. The abbreviations are decoded as follows: B = Brigitte, DP = Die Presse, G = Gala, MH = Men’s Health, M = Myself, NB = Nordbayern.de, N = NZZ, P = plan.de, R = RTL News, S = Stern, SN = Stuttgarter Nachrichten, W = WMN.

mantic or couple-related (likelihood=321,830, frequency=245), discussing the definition of toxic relationships (e.g. "What is a toxic relationship?", M 02/23), specific instances like „Megan Fox is in a toxic relationship" (N 06/23), and behaviors related to toxic relationships, including questions about the possibility of fixing such relationships (e.g. "Can a toxic relationship still be saved?", R 07/22).

The third most frequent context involves masculinity (likelihood=113,183, frequency=52), interestingly with no evidence of 'toxic' men in the corpus. These references include explanations of 'toxic masculinity' (e.g. "Toxic masculinity refers to problematic behaviors, attitudes, or thought patterns exhibited by men toward women, but also toward marginalized people [...]", MH 04/23), discussions about their (negative) consequences (e.g. "What negative effects does toxic masculinity have for him?", MH 04/23), and reflections on the term on a meta-level (e.g., "Toxic masculinity is currently a popular term, especially on social media" P 07/23). 'Toxic femininity' is also addressed in the texts, but the references are comparatively less frequent (likelihood=24,298, frequency=11). It is mostly mentioned in comparison to 'toxic masculinity', emphasizing differences between both (e.g. "This is a key difference from toxic masculinity.", SN 01/23). In addition, 'toxic femininity' is generally addressed (e.g. "Is there such a thing as toxic femininity?", MH 04/23), or even questioned, such as "Toxic femininity sounds like something invented by men's rights activists who want to counter accusations of toxic masculinity" (SN 01/23).

Additional frequent references of being toxic concern other individuals and social groups, including siblings (likelihood=99,517, frequency=33), parents (likelihood=60,439, frequency=58), and people in general (likelihood=52,050, frequency=31). Furthermore, another kind of relation appears often in the texts, namely friendships (likelihood=74,939, frequency=32). Moreover, there is a connection with 'signs' (likelihood=72,704, frequency=47) and 'words' (likelihood=36,360, frequency=18). Signs of toxic phenomena (see 'toxic harmony' below) are regularly mentioned in the discourse. The collocation with 'word' can be traced back to the metalinguistic commentary on 'toxic' (e.g. "The word toxic is used almost inflationary today.", G 06/23, see also below).

Behaviors (likelihood=62,935, frequency=82), features (likelihood=28,707, frequency=13), and characteristics (likelihood=23,122, frequency=17) are frequently described as 'toxic'. They are often not specified in detail but outlined at a general level (e.g. "Find out here how to quickly identify toxic characteristics and what the signs are.", NB 11/22). In some cases, examples of behaviors

and characteristics are then given in the text, but these are not explicitly and directly attributed as 'toxic'.

Initially surprising is, that 'toxic' even occurs in rather positive contexts. Specifically classified as toxic are positivity (likelihood=77,168, frequency=38), productivity (likelihood=31,384, frequency=15), and harmony (likelihood=22,080, frequency=10). 'Toxic positivity' is often defined in general and distinguished from optimism (e.g., "The difference between optimism and toxic positivity [...]“, S 07/22), and illustrated by concrete examples (e.g., "‘Just look at it positively.’ This sentence is toxic positivity in its purest form.“, S 07/22). The last evidence can also explain the collocation with 'sentences' (likelihood=22,187, frequency=17), as statements (in the texts: 'sentences') are often given as examples of toxic phenomena. Similarly, 'toxic productivity' is often basically explained (e.g., "Essentially, toxic productivity can be described as an endeavor to be productive at all times", W 02/23), supplemented sometimes by practical suggestions on how to overcome it (cf. "Toxic productivity: How to escape the vicious cycle", W 02/23). References to 'toxic harmony' often center on signs observable in relationships (e.g. "3 Signs of Toxic Harmony – Toxic harmony exists in many relationships [...]“, B 02/23). Notably, the evidence "Toxic harmony – now what on earth is that?" indicates a current change in the usage of the term 'toxic', reflecting new phenomena being labeled as 'toxic' in discourse.

In addition to different people, relationships, traits, and masculinity, primarily phenomena traditionally deemed positive, such as positivity or harmony, are labeled as 'toxic' in the discourse – highlighting one tendency in this discourse to label exaggerated manifestations of beneficial qualities, without any reflection, as harmful (in the sense of toxic).

Beyond the collocation with 'word' (see above), the meta-linguistic reflections on the term 'toxic' can be observed in collocations with forms of (to) 'mean' (word forms of 'toxic' as subject, e.g. 'toxic means'), indicating a linguistic shift, in the sense of a new emerging usage respective meaning of 'toxic' that still seems to require explanations and discussions. This reflects precisely the change that was already evident in the data from the first study (see above).

Various constructions with forms of (to) 'be' are also commonly used to describe toxic phenomena at a meta-level, explaining 'toxic' in its new meaning and in new contexts, as in "What is a toxic person?" (SN 06/22). Another connection with a verb can be observed between 'toxic' and (to) 'recognize'. Thus, the discourse is often about 'recognizing toxicity'. This recognition frequently

appears as an explicit goal or advice, such as "Find out here how to quickly recognize toxic characteristics and what the signs are." (NB 11/22), indented as a first step in dealing with toxicity.

Overall, the analyzed texts reveal that it is predominantly people – with varying degrees of specificity – who are described as 'toxic'. This should be understood within the context of the research subject's constitution, viewing 'toxic' as a social phenomenon. The connection between 'toxic' and masculinity, which has already been addressed in previous research (cf. Zifonun 2023) and was also evident in the first study, remains clear in the discourse excerpt examined here. The numerous general, nonspecific mentions and the metalinguistic references reflect the ongoing change in the meaning and usage of 'toxic'. A potentially new and specific context of usage identified here is the attribution of inherently positive phenomena as 'toxic', such as harmony or positivity. However, when these phenomena are exaggerated, they are constructed as toxic in discourse, often accompanied by references to recognition of toxicity and cautionary warnings.

4. What we can learn from the two studies: Conclusion, Discussion and Outlook

The findings of the two studies collectively demonstrate that the attribute 'toxic' has undergone a substantial semantic and pragmatic shift from its original technical meaning in chemistry and medicine to a socially and discursively embedded evaluative term. Study 1 shows that the term is widely recognized among German-speaking participants, yet self-reported use remains moderate, suggesting that 'toxic' functions more as a socially mediated evaluative label than as an integral part of everyday vocabulary. Participants predominantly encounter the term in peer-to-peer communication and on social media, particularly in contexts relating to relationships, behavior, and masculinity. The semantic associations provided – ranging from harmful, destructive, and unhealthy to norm conflicts and power dynamics – highlight that 'toxic' has acquired a morally and relationally evaluative dimension, reflecting contemporary sensitivities to interpersonal dynamics and psychological well-being.

Study 2 corroborates these findings within media discourse, revealing that 'toxic' is most frequently used to describe people and relationships, particularly in discussions of toxic masculinity. Notably, the corpus analysis also identifies

emerging contexts in which traditionally positive phenomena – such as positivity, productivity, or harmony – are framed as toxic when exaggerated or dysfunctional. This indicates a further broadening of the term's evaluative scope, which extends beyond its initial association with negative traits or behaviors to include cautionary evaluations of social or personal excesses. The frequent meta-linguistic usage observed, including constructions with 'to mean' and 'to recognize', underscores an ongoing negotiation of the term's meaning and the necessity of interpretive guidance in contemporary discourse.

Together, the two studies suggest that 'toxic' functions as a flexible, context-dependent label with a dual role: it identifies harmful social phenomena while simultaneously participating in broader public discussions about norms, relational ethics, and power dynamics. The convergence of survey and corpus data indicates that semantic broadening and metaphorical transfer characterize its current trajectory in both everyday language and media discourse.

In conclusion, 'toxic' exemplifies a term in transition, reflecting a shift from a technical descriptor to a socially evaluative label. Its frequent application to relationships, behavior, and masculinity highlights both its diagnostic and normative functions, while emerging uses in describing exaggeratedly positive phenomena signal ongoing semantic innovation. The combination of perception, production, and evaluation captured in Study 1, along with the corpus-based discourse insights from Study 2, underscores the multidimensional nature of the term and the importance of contextual sensitivity in understanding its contemporary usage.

Bibliography

- Anthony, Laurence (2022): AntConc (Version 4.1.4) [Computer Software]. URL: <https://www.antlab.sci.waseda.ac.jp/> [accessed 11.11.25].
- Baker, Paul (2023): *Using Corpora in Discourse Analysis*. London: Bloomsbury.
- BoyWithUke (2021): Toxic. Self-released. In: YouTube. URL: <https://www.youtube.com/watch?v=Mvaosumc4hU> [accessed 07.01.26].
- Busse, Dietrich; Teubert, Wolfgang (1994): Ist Diskurs ein sprachwissenschaftliches Objekt? Zur Methodenfrage der historischen Semantik. In: *Begriffsgeschichte und Diskursgeschichte. Methodenfragen und Forschungsergebnisse der historischen Semantik*. Ed. by Dietrich Busse, Fritz Hermanns and Wolfgang Teubert. Opladen: Westdeutscher Verlag. P. 10–28.

- Hall, Stuart (1997): *Representation. Cultural Representations and Signifying Practices*. London: Sage.
- Hoffmeister, Toke (2023a): Von toxischen Substanzen und toxischer Männlichkeit. Eine korpuspragmatische Studie zur Erschließung eines Gegenwartsphänomens. In: *Zeitschrift für germanistische Linguistik (ZGL)*. 51, 3. P. 424–458.
- Hoffmeister, Toke (2023b): Toxizität / das Toxische. In: *Diskursmonitor. Glossar zur strategischen Kommunikation in öffentlichen Diskursen*. Ed. by Forschungsgruppe Diskursmonitor und Diskursintervention. Published December 20, 2023. URL: <https://diskursmonitor.de/glossar/toxizitaet-das-toxische/> [accessed 18.11.25].
- Kennedy, Graeme (2014): *An Introduction to Corpus Linguistics*. London: Routledge.
- Kuckartz, Udo (2014): *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Weinheim, Basel: Beltz Juventa.
- Lemnitzer, Lothar; Zinsmeister, Heike (2015): *Korpuslinguistik. Eine Einführung*. 3rd, revised and expanded edition. Tübingen: Narr Francke Attempto.
- Mayring, Philipp (2014): *Qualitative Content Analysis. Theoretical Background and Procedures*. In: *Approaches to Qualitative Research in Mathematics Education. Examples of Methodology and Methods*. Ed. by Angelika Bikner Ahsbals, Christine Knipping and Norma Presmeg. Dordrecht, Heidelberg, New York, London: Springer. P. 365–380.
- Mayring, Philipp (2015): *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim, Basel: Beltz.
- Motschenbacher, Heiko (2018): Corpus linguistics in language and sexuality studies. Taking stock and future directions. In: *Journal of Language and Sexuality*. 7, 2. P. 145–174.
- Osgood, Charles E.; Suci, George J.; Tannenbaum, Percy H. (1957): *The Measurement of Meaning*. Urbana, Chicago, London: University of Illinois Press.
- Oxford English Dictionary (2025): "toxic (adj. & n.)," June 2025. URL: <https://doi.org/10.1093/OED/4084147255>.
- Schiffrin, Deborah; Tannen, Deborah; Hamilton, Heidi E. (2005): *The Handbook of Discourse Analysis*. Malden, Oxford: Blackwell Publishers.
- Spears, Britney (2004): *Toxic. Album In the Zone*. Jive Records.
- Spieß, Constanze (2025): Die soziopragmatische Perspektive auf den Repräsentationsbegriff – Zur Repräsentation von Krankheit und Gesundheit in

- bioethischen Diskursen. In: Repräsentationen aus linguistischer und interdisziplinärer Perspektive. Special Issue of Germanistische Linguistik. 56, 2. P. 113–158.
- Spitzmüller, Jürgen; Warnke, Ingo (2011): Diskurslinguistik. Eine Einführung in Theorien und Methoden der transtextuellen Sprachanalyse. Berlin, Boston: De Gruyter.
- Zifonun, Gisela (2023): Das Toxische und die Männlichkeit. In: Sprachreport. 39, 3. P. 1–6.

Tea Bagging, Trolling & Trash Talk

A Multimodal Approach for Toxicity in Twitch Streams

Tamara Bodden & Ashhab Ahmed

Abstract Toxicity research already investigates chat and gaming communication. Live streams on platforms like Twitch, however, are a challenging research object due to their multimodal affordances, which enable toxic actions alongside verbal transgressions. Existing approaches that focus only on easily extractable chat communication cannot capture the complex interplay of toxicity among multiple participants in a livestream (players, streamers, viewers in chat) across different multimodal layers. This paper uses a stream (over 94,600 Tokens) of the game *DotA 2* as a basis to describe the different types of multimodal toxicity that can be found in stream and meta discourse, as well as the interplay between the different actors and layers of the stream. The analysis shows that while there is awareness about certain topics that overstep norms (for example, racism), there is a high percentage of actions that can be considered mildly toxic, for example, in using swear words or attacking the streamer in a bantering fashion.

Keywords Toxicity, Stream, Online Gaming, DotA 2, Twitch

Abstract Die Forschung zum Thema Toxizität befasst sich bereits mit Chat- und Gaming-Kommunikation. Live-Streams auf Plattformen wie Twitch stellen jedoch aufgrund ihrer multimodalen Möglichkeiten toxische Handlungen zulassen, die über rein verbale Überschreitungen hinausgehen, ein anspruchsvolles Forschungsobjekt dar. Bestehende Ansätze, die sich ausschließlich auf leicht extrahierbare Chat-Kommunikation konzentrieren, können daher nicht das komplexe Zusammenspiel von Toxizität zwischen mehreren Teilnehmer*innen eines Live-Streams (Spieler*in, Streamer*in, Zuschauer*in im Chat) auf den verschiedenen multimodalen Ebenen erfassen. Dieser Beitrag nutzt einen umfangreichen Stream (über 94.600 Tokens) des

Spiels *DotA 2* als Grundlage, um verschiedene Arten multimodaler Toxizität zu beschreiben, die im Stream- und Metadiskurs zu finden sind, sowie das Zusammenspiel zwischen den verschiedenen Akteur*innen und Ebenen des Streams. In der Analyse wird deutlich, dass zwar grundsätzlich ein hohes Bewusstsein zum Thema Toxizität besteht (und z.B. bestimmte rassistische Äußerungen vermieden werden) aber gleichzeitig zahlreiche Formen von ‚milder‘ von Toxizität z.B. in der Nutzung von Kraftausdrücken oder in einer Praxis des Beleidigens in Form von ‚bantering‘ zu finden sind.

Schlagworte Toxizität, Stream, Online Gaming, *DotA 2*, Twitch

1. Introduction

Twitch is a digital platform primarily used by gamers that reaches over 2 million average concurrent users (Twitchtracker) and attracts an international audience across different age groups. With their interactive mix of ‘hot’ (live videos) and ‘cold’ (chat functions) media content (Sjöblom, Hamari 2016: 4), live streams offer a unique communication structure in which meaningful communities emerge and form specific practices (Diwanji et al. 2020: 7). Online multiplayer gaming has always faced high levels of toxicity, i.e. anti-social behaviour in interpersonal communication (e.g. Kwak, Blackburn 2014; Martens et al. 2015), which can also influence streams as an extension of gaming culture. Twitch streams, therefore, can give insights in how to understand (and potentially prevent) toxic behaviour. But they also make a re-evaluation of the classification of toxic actions and speech investigated by prior research necessary that often classifies toxicity exclusively as a negative phenomenon since specific toxic actions seem to be common and accepted practices in many streams. Toxicity is therefore conceptualized following Wittgenstein (1967: PU 23) as *Sprachspiel* (language game) that adheres to specific rules and is embedded in a context (Mall et al. 2020: 2; 1222; Nobata et al. 2016: 152), requiring a pragma-linguistic approach that enables a more adequate evaluation of toxicity. This study, therefore, uses a qualitative sample, laying the groundwork for further investigation of the phenomenon.

Knowledge about the investigated communities is not only necessary to understand and recognize specific linguistic patterns but also helps to differentiate societal norms from a potential self-classification of toxic language by users: Zhou et al., for example, note that mention of minority identities is

often classified as toxic by current automated systems, while some of these words can express emotionality and closeness between the communicators (2021: 3143–3145). The same expression can be considered acceptable or toxic in different groups (Nobata et al. 2016: 146). Filters or lists are thus problematic if they don't distinguish the different assessments of toxicity in diverse communities (Mall et al. 2020: 2) and lead to false positives (discussed a.o. in Thompson et al. 2017). Multimodal toxicity is another challenge of Twitch streams, and especially image-based toxicity can stay undetected with text-based approaches that currently dominate toxicity research. To address these challenges of group-specific toxicity and multimodal data, the paper focuses on how toxicity arises across the different multimodal layers of the stream.

As a basis for the study, a stream sample was chosen that meets criteria favouring toxicity (massive stream size, male streamer, competitive online gameplay; Dreier, Pirker 2023) to identify various types of toxic actions and linguistic patterns. These patterns will be presented in examples, quantitative overviews and a tabular form to distinguish between targets, types, topics and linguistic/visual characteristics.

After giving a brief overview of current toxicity research in the context of online gaming and Twitch (2.) we will describe the data acquisition process and the chosen sample (3.). The discourse on the term 'toxic' in public discussions around streamers will be analysed (4.), setting the basis for the toxicity analysis (5.) that can help identify multimodal toxicity in streams. This analysis will focus on toxicity triggers (5.1), potential toxicity (5.2) and meta-discourse about toxicity (5.3).

2. Toxicity Research on Gaming and Twitch

In linguistic research, many automated approaches for toxicity detection have been developed either based on public discourse, with a focus on chat communication (Stoop et al. 2019 provide an early overview) or on communities that self-define as practising hate speech (Saleem et al. 2016). Piot et al. (2024), for example, combine over 1.2 million social media posts from prior studies to create the database MetaHate to further joint efforts in toxicity research. Though these quantitative approaches are constantly evolving and are already being implemented for automated toxicity detection on platforms, they also come with negatives like the aforementioned false flagging of minorities (Zhou et al. 2021), undetectable ambiguous meanings (e.g. Stoop et al. 2019: 20) and a lack

of context in annotations that can result in false detection or overlooked toxicity. Sheth et al. (2022) therefore emphasise the value of qualitative analysis that can be combined with quantitative approaches.

Both Twitch and Online Gaming seem to be abundant sources for material in toxicity research due to the competitiveness of online gaming (Blackburn, Kwak 2014) and factors such as the online disinhibition effect (Beres et al. 2021). A common approach to research in these areas is the analysis of chat logs or automated log data in gaming (e.g. Martens et al. 2015; Thompson et al. 2017; Stoop et al. 2019). In a meta-study of 54 research papers on toxicity in online gaming, Anker Nexø (2024) identifies a current lack of multimodal approaches, an insight parallel to the review by Harpstead et al. (2019) of 46 studies of the platform Twitch. While there are already many ‘technically supported’ studies (which, for example, use metadata of the platform or chats) and ‘technically agnostic’ studies (such as surveys, for example, Cai, Wohn 2019), both Anker Nexø and Harpstead et al. describe how ‘technically challenging’ studies working with multimodal data are not yet as present but promise valuable insights.

Since toxicity is characterized by its property to spread in communication (Seering et al. 2017: 112), it is vital to see all toxic actions in their context, which, on Twitch, includes the broadcasted gameplay, interactions in the video/audio of the streamer and lastly the chat with its multimodal messages. Especially the emotes (images and GIFs used in chat by typing their names) differentiate the Twitch chat from many other platforms investigated in prior toxicity research, such as X (ElSherief et al. 2018) and Reddit (Mall et al. 2020). The reluctance for multimodal analysis of Twitch data on an extended scale can be explained by the difficulties in acquiring and transcribing the data (hence Harpstead et al.’s coining as “technically challenging studies” 2019: 6), which will be described in the following section.

3. Data acquisition and description of the sample

The aim of this paper is to describe different triggers (events or terms that can instigate or lead to toxicity) and multimodal ways to express toxicity. The objective for a data sample therefore was not to find a prototypical Twitch stream but to use a stream likely to be very toxic as a basis for the analysis. To find a suitable stream, the first step was to create a small sample (13 texts, 17,959 tokens) of meta discourse on toxic streamers by searching for the (potential) headline “Most toxic Twitch Streamers” (on the search engine Ecosia 04.07.25),

checking for articles that mention or list specific streamers and their supposedly toxic behaviour.¹ The thirteen collected articles/blogs/forums were published between 2018 and 2024 (in the following called corpus B), and they list 90 different streamers, though only seven have had four or more mentions. Of these seven streamers (xQc, Ice Poseidon, Amouranth, Destiny, Forsen, JiDion, Pokimane), three were not active on Twitch (anymore) because they primarily produced content on YouTube or were banned from Twitch for violating community guidelines.

For all remaining four streamers, the allegations of toxic behaviour were checked in detail, and it became apparent that for the female streamers, their gender², and popularity on the platform might have contributed to the allegations. Thus, following the definition of toxicity in this paper (see 4.), neither female streamer met the criteria of interest and therefore did not seem promising for classifying different toxic actions. However, for the streamer xQc (mentioned in seven texts) and his community of ‘Juicers’ as well as Forsen’s ‘Bajs’ (four texts), the articles often describe both streamers and their communities as extremely toxic. According to Seering et al. (2017), streamers have a great impact as role models on their communities, so it is likely that toxicity will be a shared trait between streamers and their communities. It seems that “the group is defined by the speech and the speech comes to define the group” (Saleem et al. 2016: 2).

Dreier and Pirker (2023) identify gender, stream size and game genre as factors influencing toxicity levels in Twitch streams. In this regard, both male streamers with large audiences³ are more likely to experience toxic incidents during streams. DotA 2 (Defense of the Ancients), a competitive online game considered very toxic in gaming communities and in research (Lee et al. 2021; Martens et al. 2015), was chosen for a stream sample since both streamers occasionally play the game. Based on a first multimodal impression of current streams from Forsen and xQc (e.g. use of swearwords, toxic messages in chat),

-
- 1 Articles that only focused on specific games and their communities were excluded as well as videos that were suggested by the search engine.
 - 2 In a misogynist view of women as ‘bad gamers’ female live streamers allegedly use sexuality to attract audiences (and consequentially draw them away from serious streamers) while producing lower quality content. Community policing (for example flagging) or hostile raiding can therefore be used to exclude female streamers (see Adams et al. 2025: 8; Zolides 2023: 134).
 - 3 At the time the data was collected (22.08.25), xQc was ranked 8th globally among Twitch streamers, and Forsen was ranked 251th on twitchtracker.com.

we decided on an xQc stream as a suitable sample, since he displayed more obvious toxicity on stream and seemed to have made toxicity a (potentially partially performative) part of his streaming style.⁴ Because of the density of the material of a massive stream and the explorative character of this study, only one stream was analyzed, but the results will be discussed in future work in the context of nine more streams that are currently being collected.

Chosen as a sample for this study (corpus A) was a 140-minute stream of *Dota* gameplay from summer 2025,⁵ which happened at the end of a 12-hour stream with various content. The late streaming time, as well as the two losses in the game, were assumed to contribute to increased toxic utterances and actions, which impacted the decision on this sample.

For the analysis, a transcript of the visual and auditive elements of the game, stream and chat was generated, utilizing Recktenwald's (2017) transcription scheme for Twitch. The chat messages, including time stamps from the Twitch stream, were copied to a txt-file and anonymized (using a Python script and following data protection guidelines). The chatters were assigned consistent identifiers (User+Number) to better understand shifts of behaviour over time; only the streamer and Bots (Schnozebot and Fossabot) kept their names. Moderators were marked as such since they can have a greater impact on both positive and negative community behaviour (Seering et al. 2017: 113). To simplify the annotation of the emotes, they were reduced to a text layer and were visibly marked through duplication of the emote's name (for example, LOLLOL). To better understand their content, all 890 emotes were additionally mapped in a CSV file to aid annotation and pre-sort potentially toxic emotes (see 5.2).

For the audio transcript, the automated transcript available on YouTube (of the reupload of the stream by xQc) was used as a basis, and the text was separated into timestamps (in seconds), a column for the streamers' (and other people's) utterances on stream, as well as the gameplay voices (characters, announcers, co-players). All swearwords that had been automatically censored

4 "It seems that being banned multiple times and cultivating a reputation as one of the most toxic streamers on Twitch is actually good for xQc, as he has almost 10 million followers on Twitch at the time of writing." (Dungca 2021)

5 Note on research ethics: Working with public data in the context of toxicity research can come with the stigmatisation of users, therefore the exact date of the chosen stream is not shared in this paper in order to not expose any viewers to potential negative consequences. For a further discussion on this topic see Legewie, Nassauer (2018).

by YouTube were uncensored, and mistakes of the transcript (like misspelling of in-game names, game-specific vocabulary or overlapping speech) were corrected manually following modified orthography transcript principles (eye-dialect, e.g. Edwards 1993: 11–20). A Python script was used to align timestamps and lines of text in a CSV file. In the next step, the titles of music played on the stream were manually added (using the app Shazam to identify unknown songs) as were visual descriptions of both the gameplay and the streamer. This resulted in a dense transcript of all three layers (game, stream, chat) as illustrated with a small sample in Table 1.

Tab. 1: Excerpt of the combined stream transcript.

Time	Game		Stream			Chat
	Visual	Audio	Visual	Audio		
				Voice	Music	
10:20:09	Wraith King (XQC) kills Ezalor (Enemy)				Culture Beat: Mr. Vain	User537: EZEZ, User463: LOLLOL, User23: EZEZ
[...]						
10:20:11				fuck-ing loser	CB: MV	User238: wow, User538: EZEZ, User28: AYYYY, User66: EZEZ NICE, User82: nice
[...]						
10:20:13		WK: It was kind of you to grow that skeleton to maturity for me.			CB: MV	User142: AlienPlsAlien-Pls EDMEDM Alien-PlsAlienPls EDMEDM AlienPlsAlienPls EDMEDM AlienPlsAlienPls EDMEDM AlienPlsAlienPls EDMEDM, User541: PogPog, User71: EZEZ, User89: LoserLoser, User98: EZEZ

Assigning every second of the sample one row in the CSV sheet (6,057 lines), the transcript contains the artist and title of identified music in the background, 18,200 chat messages (79,129 tokens), 855 lines of audio commentary from the streamer (5,916 tokens), as well as 411 descriptions of game visuals (3,279 tokens) and 389 of stream visuals (1,936 tokens).

4. Definition of the Term Toxic – Meta discourse

Toxicity can include all “actions of a user or users that create a negative atmosphere for the other users of social media” (Mall et al. 2020: 1). While this broad definition contains different multimodal forms of toxicity, it is also inevitably tied to a very subjective assessment, because the negative atmosphere must either be defined by the victims of a toxic action and thereby participants of a specific situation or through societal standards. The first viewpoint doesn’t allow for judgement by external viewers, but societal standards, on the other hand, might not be appropriate or applicable to every specific community. Before annotating toxicity, therefore, insights from research as well as the public meta discourse on toxicity were evaluated, tying up to the three layers of the stream (game, streamer and chat):

Anker Nexø (2024) presents a list of 60 toxic actions discussed in analyzed research papers. These toxic actions range from unspecified toxicity, negative attitudes, to verbal actions (a.o. blaming, commands, complaints, cursing, flaming, hate speech, insults, leaking of information, malicious jokes, profanity, threats, name-calling or trash-talking), game specific behaviour (a.o. tea bagging, cheating, camping, intentional feeding, AFK-ing) or actions around the gameplay (a.o. lobby dodging, flagging, playing off-meta or smurfing). While the overview by Anker Nexø helps identify toxic game actions, it does not attempt to bring order to the often overlapping and synonymous descriptions of toxicity. Additionally, the streaming setting of the analyzed video brings another layer of potential toxic (non-verbal) actions with it that were, according to Nexø, not as present in prior research (2024: 7). Many studies on toxicity have a predefined concept of what is considered (and therefore annotated) as toxic instead of a data-driven approach that is followed here. For a less subjective classification of streaming-specific toxic behaviour, two sets of discursive data were analyzed to comprise a comparable list of typical toxic actions on Twitch. These negative actions were annotated using the

software MAXQDA, following principles of the grounded methodology in an open annotation process.

While Twitch already implies certain undesirable actions in the Community Guidelines, acceptable or forbidden behaviour is often further specified by individual channel rules. In a preceding study⁶ a corpus (corpus C) of over 50 channel rules with 3,517 tokens was annotated to find commonalities in explicitly undesired behaviour. The 490 annotations led to 44 common codes in eight primary and 36 subcodes. While many channels often use positive guidelines reminding of mindfulness/kindness, respect for streamers, viewers and moderators or having fun, toxicity and rudeness in general were often mentioned as prohibited actions. This includes specific practices like posting of (potentially dangerous) links, (self) promoting, mentioning of other streamers, doxxing, trauma dumping⁷ or chat actions like spamming or the use of all-caps. Racism and sexism were usually general categories that were forbidden in any form; especially female streamers pointed out (sexual or any) comments on their bodies as reasons for bans. For practices more exclusively relevant to live-streaming, some streamers also explicitly prohibited 'backseating' or mentioning the stream's viewer count as personal preferences.

The discourse on toxic streamers (corpus B, described in section 3 containing 13 texts and 17,959 tokens) was also analyzed in MaxQDA by annotating and classifying all mentioned actions that could be considered toxic. 480 annotations distinguished between the topics of counter-actions against toxicity (5), results of toxicity (152) and toxic actions on Twitch in general (163), specifically by streamers (143) or viewers (17). Toxicity was described mostly in regard to discrimination or offensive speech, but there were also extreme examples discussing cases of pedophilia, suicide threats, sexual harassment or bullying. Often mentioned concerning streamers' behaviour on camera were animal abuse, aggressive actions, substance abuse, sexually explicit content, shaming/lashing out at viewers, but also the stirring of drama (to gain views) or hostile raiding. Especially actions related to finances, such as Ponzi schemes, promotion of cryptocurrencies, pressuring people into donating to the streamer

-
- 6 Presented by Bodden at GAL 2024 with the title "'No toxicity in chat' – Net-Etiquette im third space Twitch". The analyzed stream categories on Twitch were IRL, Minecraft, Fortnite, Music and Just Chatting. Especially Fortnite as the only clear competitive online gaming had the least number of guidelines in the checked accounts.
 - 7 Meaning the (over)sharing of personal problems (e.g. relating to mental health) that brings a negative atmosphere into streams.

or gambling, were often mentioned as toxic practices that led to the banning of streamers from Twitch. Other toxic practices were attributed exclusively to viewers (e.g. swatting streamers, simping, spamming or making inappropriate requests), and it was discussed that toxic streamers attract dangerous communities that can show more extreme behaviour. A few of the analyzed articles discussed ways to prevent or counter toxicity, such as disengaging or directly calling out toxic actions. Most of them also mentioned the potential punishment for toxic actions in the form of banning, but also mental health problems, fines, the ending of careers or loss of monetary incentives were discussed.

The categories derived from the annotations of corpus B and C, alongside the reviewed literature, served as the basis for the following analysis of corpus A by sensitizing for different phenomena in online gaming and Twitch that can be understood as toxic.

5. Analysis of Triggers, Potential Toxicity and Meta Discourse

While the annotation was guided by the categories in the analyzed literature, guidelines and meta discourse, again an open coding process was followed where labels were newly generated, adjusted or merged over time, but followed the three guiding categories *Triggers* (5.1), *Potential Toxicity* (5.2) (here also differentiating between the layers game, stream and chat) and *Meta Discourse* (5.3). Additionally, further information, such as explanations of specific terms or practices, was added in memos (virtual notepads attached to a specific section of the transcript).

To better understand where toxic actions occur across the layers of the stream, the columns of the joined transcript were separated again into distinct documents that included the timestamp and text of each section, which already clearly indicate the density of potential toxicity (8,641 annotations, see 5.2) relative to tokens in each layer.

Tab. 2: Annotations and Potential Toxicity in the different layers of the transcript.

	Game		Stream		Chat	Total
	Visual	Audio	Visual	Audio (Voice only)		
Tokens	3,279	4,347	1,936	5,916	79,129	94,607
Total Annotations	249	226	4	382	8,695	9,556
Potential Toxicity	66	51	4	348	8,172	8,641

As shown in Table 2 in the game layer, most annotations were trigger events (5.1), and there were only a few toxic occurrences (usually in the form of chat messages by the co-players or taunts). The visual layer of the stream barely has any annotations, while the voice/audio part of the stream shows many annotations of toxic speech. The chat has both the highest number of tokens and with over 8 thousand annotations of toxicity around every 5th message contained toxicity.

The transcribed music cannot be discussed in detail in this paper and therefore does not appear in the overview. However, it seems fruitful to include this in a transcript because the communication often picks up the background music.⁸ The levels of toxicity by the streamer and chat seem to loosely coincide with the tone/style of music in the background (e.g. more cursing with rap music than pop), which could be just a coincidence of the small sample and therefore has to be analyzed in detail in future work.

8 For example, someone insults the players as “99 NOOBS” (Chat, Pos. 838) tying up to the song ‘99 Luftballons’ in the background. Especially the song ‘Karma is my boyfriend’ (or as someone rephrases, “Karma is losing Dota” Chat, Pos. 1072) by Taylor Swift sparks controversy in chat: When the streamer sings along parts of the lyrics, he is repeatedly labelled as a soyboy (a man who is showing more feminine traits than certain concepts of masculinity) and ridiculed.

5.1 Toxicity Triggers

Almerekhi et al. (2019: 291) describe Toxicity Triggers “in the context of online discussions as comments that incur direct toxic responses. [...] [T]hese triggers of toxicity may differ by the community and by topic due to different norms and uses of language”. With a corpus of more than 104 million Reddit comments they identify that terms like “muslim, israel, kill, europe, support, tax, vote, law, president, woman” (ibid.: 292; also see TOXTRIG by Zhou et al. 2021: 3144) can lead to a toxic shift in comments. While political and religious topics are almost non-existent in the investigated sample,⁹ other triggers, like for example ‘president’, were mentioned and triggered toxic reactions, but not in relation to real-world politics.¹⁰ Other terms like ‘kill’, however, are very present in the gaming discourse of a MOBA (Multiplayer Online Battle Arena) but are not as likely to spark toxic reactions. A lexicon-based approach tied to these lists, therefore, does not seem promising.

Since research on toxicity in multiplayer online games agrees that it is usually in-game actions that form the basis for toxic behaviour, these were the focal points of the analysis. 374 potentially triggering game events were annotated across the game’s layers (visual: 183; audio: 177) and stream (audio: 14). It is assumed that all events that are very negative in the game can lead to frustration in players, which increases the chance of toxic behaviour. Certain events (see table 2) are communicated in some form of visual or auditive ping or announcements, which were transcribed and annotated. This can help see which elements of the coded actions actually have an impact, or how, for example, the combination of a series of low-impact triggers can add up to a response that can be classified as potential toxicity (in 5.2). The impact of these events on the negative mood of the streamer can also be seen in the annotations of trigger events (here in the form of meta discourse).

9 Occasionally some viewers attempt to start a political discussion, but they seem to be lost in the masses of chat messages and usually do not serve as toxicity triggers “Is this the appropriate venue to share my Strongly held views on Israel?” (Chat, Pos. 2522). As Sjöblom & Hamari (2016) found out in their study about motivations for Twitch users, many people use streams for relaxation and don’t want to encounter stressful topics. Consequentially, many streamers in the guideline corpus (C) explicitly forbid politics as a topic on their stream.

10 The streamer makes the proposition that he as a president would punish players that use taunts by turning of their utilities.

Tab. 3: Toxicity Triggers in the 3 layers of the stream.

Game visual	Game Audio	Stream Audio (meta discourse)
Lobby / Preparation Seeing a loosing streak Seeing bad matchmaking rank Queue resets because other player decline Seeing score board & match statistics Gold being lost by team for failure to pick a hero Chosen hero gets banned In-game Team members disconnect (& reconnect at a bad time) Player gets disconnected Mocking taunts by co-players Death of player / team members Courier being killed Towers being lost Enemies archive milestones Enemies' spells prevent player's actions Chat messages by co-players Announcements of enemies' success (double kill, triple kill, monster kill etc. by enemy team)	Lobby / Preparation Sound of gold being lost Countdowns for team actions (5 seconds remaining) Sound of hero being banned In-game Voiceline of character being under attack Announcements of enemies' success (double kill, triple kill, monster kill, killing Roshan) Announcements of team's losses (tower being attacked and falling etc.) Announcement of enemies' team win Voice taunts by co-players	Lobby / Preparation "this day I had a really bad loss streak" (Stream (Audio), Pos. 43)¹¹ "three in a row and I can't I I can't win" (Stream (Audio), Pos. 77) "oh my god we double picked him" (Stream (Audio), Pos. 240) In-game "is this guy gonna connect or what" (Stream (Audio), Pos. 475) "oh he reconnected (grunt) he reconnected at the last second that's absolute diabolical" (Stream (Audio), Pos. 589–594) "I'm lagging" (Stream (Audio), Pos. 3291)

It is assumed that all toxic actions can also function as toxicity triggers for further toxic behaviour tied to the broken window hypothesis (Cheng et al. 2017: 1219): especially when toxic actions go uncommented, this can be seen as an invitation for further toxic behaviour.

¹¹ All quotes from the transcript will contain the layer (e.g. Chat / Stream (Visual)) and line (Pos.) in the transcript, spelling mistakes will not be corrected.

5.2 Potential Toxicity

The description of “Potential Toxicity” is based on the rejection of a definite attribution of ‘toxicity’ as a normative term that does not do justice to the various communities and their linguistic norms. Since the goal of this paper is not to judge certain behaviours but to annotate anything that can fall into a wide classification as ‘toxic’, there will be a tendency to over-annotate phenomena that can be mildly toxic or not considered problematic at all. For example, backseating (e.g. unwanted tips for the streamer’s gameplay), which is explicitly forbidden in many individual community guidelines, is a common practice in the analyzed stream. Likewise, the use of swearwords (or abbreviations like wtf) is normalized within gaming communities. All these occurrences will still be considered as potentially toxic for this annotation. Other community-specific practices that seem to reside in a grey zone of toxicity are bantering and trolling, which the following example illustrates: People in chat start pointing out a loss in MMR (matchmaking rank) by stating numbers like “-50”, indicating the streamer is not skilled, which can be toxic. This turns into a creative game where more and more absurd objects that are being lost are named. The commentators, therefore, make fun of the streamer (for being a bad player) as well as the chatters (who constantly need to point that out):

User1182: 100 PLATES PogPog, User1608: -100 Airplanes / User786: xqcMxqcM 50 stones (Chat, Pos. 3523–3524) / User601: -100 seashells (Chat, Pos. 3527) / User8: POGGERSPOGGERS 500 clams (Chat, Pos. 3553) / User786: xqcMxqcM 50 granules of sand (Chat, Pos. 3554) / User309: 100 BABY TURTLES OOOOOOOO (Chat, Pos. 3556) / User1422: -38261 chicken nuggets LULWLULW (Chat, Pos. 3569)

While the streamer is usually ignoring the chat altogether and also does not seem to be bothered by negative messages when he sees them, the same kind of fun at another person’s expense could be highly toxic in a different context. Displaying vulnerability online, especially as a streamer, is not always possible, so we cannot be sure of the impact of bantering and mockery. Normalizing these actions could negatively influence interactions in different communities as well. To investigate this, future research needs to engage with Twitch actors to better understand their attitudes and interpretation of potentially toxic actions, but insights from one community cannot be generalized easily.

5.2.1 Game

In addition to tagging (and classifying) toxicity, the annotation also includes the targets of a toxic action. These can be *directed* at a specific person (streamer, his girlfriend, teammates, enemies, other streamers, chat) but also *generalized* (for example, using ableist language) (Waseem et al. 2017). For this classification, the context and communication channels play an important role in identifying who is being addressed, for example, through pronouns.

In the game layer, players can use in-game chat, which is visible to either the team or the whole group of players or they can use taunts in both chat and game actions. Even positive or neutral taunts like “That was unreal” (Game (Audio), Pos. 2104) were used ironically to comment on a bad play by the streamer making them potentially toxic. The voice chat in contrast, is only audible to team members. While most of the communication channels can be blocked if a player is considered toxic, certain game actions (like throwing an item called ‘Cloak of Intelligence’ in front of another character to imply someone is stupid), pinging or drawing on the mini map are additional possible communication channels. Toxic players still find a workaround to be toxic with the communicative means available, though the means of communication for extremely toxic players (flagged by former co-players) will be restricted.

For *Dota*, characters’ voice lines appear to be randomized. These voice lines usually are reactions to triggers and are often harmless bragging or taunting, but can sometimes contain curse words, like the character Wraith King saying, “Death is my bitch” (Game (Audio), Pos. 481), which is often picked up in the chat. Some voice lines also directly target opponents and can be classified as toxic: “You won’t win the war by dying for your Ancient. You’ll win by making the other dumb, unfed bastard die for theirs!” (Game (Audio), Pos. 3346)

5.2.2 Stream

The streamer has the most channels for communicating and the highest control of the communication situation, which increases the possible ways to be toxic. He can directly be toxic towards his teammates, enemies, viewers and people who are physically in his stream, in addition to talking about people in a toxic fashion. With the shared view of the game and the video of the streamer, other toxic non-verbal behaviour is also visible, for example, when he is baiting his teammates into a fight by asking for help and teleporting away, commenting on his actions in voice chat with “peace out” (Stream (Audio), Pos. 2350). In the stream, toxic utterances are usually directed at the teammates (with the primary topics skill and intelligence), or the game itself (“this is so goddamn

fucking boring dude” (Stream (Audio), Pos. 5206)), and a lot of swear words are used (362 occurrences of swearwords without targeted toxicity in all layers).

The negative mentality towards team members is expressed, for example, when he comments “I don’t give a fuck if my team dies” (Stream (Audio), Pos. 2372) or directly says in voice chat “you guys are awful by the way [...] you’re awful” (Stream (Audio), Pos. 2929–2931). As can be seen later on in the listing in Fig. 1, there are far more attacks on team members than enemies, which aligns with the findings by Martens et al. (2015: 5) that there is more toxicity towards team members, especially in losing games.

As we progress through the video, the frustration of the streamer becomes apparent in the visual layer, through fidgeting behaviour (resulting in the viewers spamming emotes like ‘ADHD’ displaying the streamer) or in aggressive gestures that sometimes emphasize toxic speech. Here, pitch, intensity of speech and prosodic markers of irony are impactful to classify toxicity and were partially noted in the transcript but would require further investigation. In terms of the visual layers of the stream (that barely received any annotations), especially the behaviour of the streamer AikoBliss, xQc’s girlfriend, who is partially on camera, is interesting: One of the most viewed clips¹² on Twitch of the stream shows her in the background, exaggeratingly imitating his angry gesticulations to mock him while he complains about his co-players. Something similar happens in chat after xQc repeatedly urges his co-player (Shadow Shaman) to buy back¹³ into the game to help the team (Stream (Audio), Pos. 2871) with higher and higher pitched voice and intensity that seems performative:

Shaman buy back / Shaman buy back / Shaman buy back please man / Shaman please man buy back man / Shaman please buy back man / Shaman hehe Shaman buy back please man please man / Shaman buy back please man / Shaman please buy back man (Stream (Audio), Pos. 2871–2893)

It is assumed that toxic conduct by a streamer is not necessarily their authentic behaviour but can also be a form of performative toxicity for an audience.¹⁴ The

12 If the clip function is enabled by streamers, any viewer can clip short sequences of a stream as highlights that can be shown on the streamers page but are also often circulated on other social media.

13 When characters die, players can pay in-game gold to make them respawn faster.

14 Anker Nexø points out that while Twitch streams give access to important audio-visual data that accompanies gameplay it also “risks biased data with incomplete, edited, se-

chat picks up this ‘performance’ by creating an echo that either repeats his pleas for support or distorts the utterance in a mocking fashion, as AikoBliss did with her gestures:

User1303: shamaaan, User1150: BABAK, User697: LOLO, User357: DENTGE, User507: SHAMANBUYBACKMAN / User326: LOLLOL, [...] User66: sha man, User82: omEomE, User8: wtf are u sayin LULWLULW [...] User1205: BAKBAK MAN, User686: sher man / User246: LOLLOL, User697: hehe, User1252: sha man, User150: OMEGALULOMEGALUL, [...] User813: PLEASEPLEASE, User768: LOLLOL [...] User1194: wtf are you saying LOLLOL, User199: hehe, [...] User1390: LOLLOL, User786: OMEGALULOMEGALUL, User1142: hehe, User571: DentgeDentge, User865: LOLO bebek, User28: LULLUL / User758: SHA MaNMaN, User11: LOLLOL hehe, [...] User288: hehe / User357: LOLLOL, User663: LOLLOL, User304: sha MaNMaN, User1362: BAHBEHMEN, User45: LOLLOL / User813: hehe, User1391: SHA MAN, User49: LOLO, User659: Shamman, User276: he cant do anything without u, User1303: shamaaan (Chat, Pos. 2884–2891)

While the mocking in both instances could be potentially toxic, it can be assumed that this is bantering in the specific stream and community. The distortion effect of this bantering can however sometimes show ableist characteristics when impacted speech is associated with a lack of intelligence. This is condensed in the following example from chat where the insults “OkayegOkayeg beg reerrd playa” (Chat, Pos. 93) (directed at xQc) stands for ‘big retard player’ but also mimics an impairment in the spelling that is combined with an ableist emote.

5.2.3 Chat

With over 8,172 annotations of Potential Toxicity in the 18,200 chat messages by 2,506 different users, the stream’s chat is the most important space for toxic behaviours. While both in the game and stream, skill and intelligence are predominantly the topics of insults, the themes become more diverse in the chat layer. Here, there are only two potential modalities available to express toxicity: written chat or emotes that can appear both as GIFs and images.

lected, or over-accentuation of behaviours. Hence, some video data may be produced with an audience in mind, which risks threatening the naturalistic emergence of behaviours and interactions displayed.” (2024: 9)

As discussed before, emotes are a specific feature of Twitch compared to many other chat-based social media networks. Emotes can derive from three different types of sources: there are global Twitch emotes (like ‘Kappa’ or ‘PogChamp’), streamer-specific emotes that are usually an incentive to subscribe to a channel by paying a monthly fee and lastly third-party extensions (for example BetterTTV, FrankerFazeZ or 7TV).¹⁵ In the analyzed section of the stream, there were 890 different emotes used, which were manually annotated in a mapping document. By looking up origins, meanings and context of the emotes and their usage in chat, they were classified as potentially toxic when applicable. This was used as a basis to automatically annotate thousands of emotes throughout the document. For example, the emote ‘GYATT’ shows the streamer with an exaggerated expression, looking at something. While *gyatt* as a term can be an expression of being impressed in contemporary youth language, its origin refers to admiring female behinds, so the term is intertwined with highly misogynistic concepts of male gazing. This reading of the emote can be confirmed in the sample where ‘GYATT’ is posted mostly in a situation when a player’s profile in the game (that is customized) shows a selection of the hovering cleavages of several female DotA characters.

Some of the emotes have multiple topics and targets that were annotated additionally. For example, the ableist emote ‘Dentge’ (a frog and a variation of the Pepe-meme with an indented head implying a physical and mental disability) is generally toxic by targeting disabled people, but simultaneously can be used to express, for example, a lack of skill in a co-player.

The emotes on Twitch are considered memes, since they are communicative elements that are based on a creative use of intertextuality (specifically meme patterns) and virally distributed by numerous users (Shifman 2014: 44). Just like classic meme templates, they can fulfill important functions of mimic and gestures in the written conversation in the chat since “even the snapshot depictions of embodied features—including facial expressions, gestures, (changes in) postures—however much reduced they may be compared to full face-to-face interaction, contribute much to the emergence of meaning, especially stance meanings” (Dancygier, Vandelanotte 2025: 13). Most of

15 The first type of emote can be used and be seen by everyone, community emotes are curated by the streamers and are only available to subscribers but can be seen by everyone and lastly the third party extensions can be generated by every user but they are not being displayed for anyone who does not use the extensions so they can appear as (sometimes incomprehensible) text.

the memes, therefore, show actual people, animals or cartoon characters to express some form of emotion.

But this also means that any emote (though it might seem toxic at first glance) can be used in very different ways depending on its context, which can be described as a “memetic grammar” (ibid.: 4). For example, the combination of the frog-emote ‘okayeg’ and a clock emoji can be read as ‘retard time’ insulting the players of the prior game actions, but ‘okayeg’ can also be used to convey doubt in another context. The comment “MOST INVOKERS 5Head5Head ThinkingThinking DentgeDentge Thinking2Thinking2” (Chat, Pos. 3315) uses four combined emotes to draw a head next to a thinking bubble showing the physically impaired ‘Dentge’ frog. The difficulty with the emote ‘5Head’ on Twitch is that it can both stand for someone very smart or very stupid. The chat messages seem to use this ambiguity to express a message like ‘Most players of the character Invoker think they are very smart, but they are actually very dumb (/disabled)’. The knowledge of the game, the community, and the context is, therefore, necessary to correctly estimate the intended meaning for this kind of message, which underscores the approach of classifying everything that could be toxic as potentially toxic. Emotes can also share positive connotations, imply irony or change the meaning of a verbal sentence, which was already considered for emojis in an automated classifying approach by Thompson et al. (2017: 151).

While it is easy to spot clearly racist, sexist or ableist emotes in the analyzed material, there are two problems for efforts to restrict this form of toxicity: First of all, many users might not know the toxic background of a specific emote (for example emotes featuring topics like ‘soy boys’ or ‘clown world’ relating to AltRight movements¹⁶) or they do not care about their origin and hurtful connotations. This becomes very clear when looking at the list of hosted emotes by the streamer.¹⁷ But also apart from emotes, the chat can contain visual elements comprised of graphemes by using ASCII art. This was enabled by the

16 See Liemann (2023) for a deeper discussion of alt-right memes relating to gender.

17 For an overview see: <https://www.xqcupdates.com/emotes>. It is noteworthy that many of the curated emotes show the streamer himself tagged with derogatory terms like Loser, Jackass, Cucklord or RetardJam.

streamer through the automated Schnozebot, for example, when a user writes the command “!lewd” in chat, the bot draws an Ahgaho-face.¹⁸

Additionally to the visual elements in chat, of course, most toxicity was expressed verbally. For example, different kinds of curse words appear in chat unfiltered but often abbreviated. In addition to these slur words, derogatory terms (for example, the streamer is often called bub, good boy, lil gup, lil buddy) or dehumanizing metaphors (e.g. animal, doggie, roach, monkey or bot) are used to either attack people in the game or the streamer. On the level of sentence structures, toxicity can also be expressed, for example, by the usage of repetition and distortion, which was already visible in the ‘buy-back-Shaman’ example.

Complex forms of intertextuality are another resource that is used to create new forms of creative insults. When the streamer, for example, is not playing the game according to the chat’s standards, there are repeated comments about a POV (point of view, a term from digital storytelling culture), meaning he only shows the game from a non-player perspective or commenting on him only walking instead of fighting, indicating his lack of skill:

User1358: Cameraman POV (Chat, Pos. 5120) / User1925: sakura pov LOLLOL (Chat, Pos. 5307) / User1932: COURIOUR POV LOLLOL (Chat, Pos. 5312) / User2074: you out there doing cardio prePfftttprePffttt (Chat, Pos. 5321) / User1861: walking sim (Chat, Pos. 5125) / User2015: walking ward (Chat, Pos. 5120)

Many chat messages in massive Twitch streams are highly intertextual and might not seem to make sense at first glance. For the disambiguation of both emotes and written text in chat, Ford et al.’s conceptions of ‘crowdspeak’ and the three practices that create coherence (2017: 863) in streams can therefore be helpful. They describe that “massive chats are filled with non-sequiturs, verbatim repetition, variations on prior messages, blocks of text copied from elsewhere, and tiny messages that include only a few words or emotes” (ibid.: 859). ‘Shorthand’ (the contraction of text into a smaller space¹⁹) and ‘bricolage’

18 A special term for (usually female) faces expressing pleasure in Japanese animated porn. Ahgaho-faces are explicitly forbidden in Twitch emotes which, in this case, is circumvented by the use of the bot.

19 For example, the message “ur cm afk om” (Chat, Pos. 760) stands for ‘your Crystal Maiden (a character in *Dota*) is away from their keyboard’ and om is likely standing either for OMG (oh my god) or ‘OmE’ (short for Omegalul) used as an intensifier to ex-

(the recombining of elements from a small repertoire²⁰) can be seen, for example, in the described practice of distortion and are guiding principles for both text messages and meme culture. But especially the concept of ‘voice taking’ can be helpful for toxicity annotation: ‘Voice taking’ describes the adoption of shared viewpoints and mannerisms by people in chat, which means that not as many new thoughts are being presented in chat, but they are just repeated in variations. Massive chats can therefore “maintain[] coherence by reducing the total volume of meaningful content participants produce and process” (Ford et al. 2017: 860). For toxicity annotation, this means that messages close to one another are likely to relate to the same subject and might therefore have comparable levels of toxicity.

For example, the following chat messages all contain the same meaning (ridiculing xQc’s hair, appearing in about 320 comments in total) using emotes, sarcasm, comparisons or metaphors to convey a toxic message:

User578: LOLLOL HAIR (Chat, Pos. 5831) / User1313: xqcTreexqcTree (Chat, Pos. 5832) / User786: bird hair (Chat, Pos. 5833) / User8: NAHHNAHH WTF IS THAT HAIR (Chat, Pos. 5834) / User1470: holy hair OMEGALULOMEGALUL (Chat, Pos. 5835) / User2261: forsenLaughingAtYouforsenLaughingAtYou HAIR (Chat, Pos. 5835) / User278: LOOK AT THAT BANAN PEEL (Chat, Pos. 5837) / User1473: brother cut that shit (Chat, Pos. 5836) / User64: omEomE nice hair (Chat, Pos. 5838) / User2301: WutFaceWutFace look at him (Chat, Pos. 5841) / User2303: (palmtree emoji) ClapClap (Chat, Pos. 5844) / User2274: MOPHEAD ANDY LOLLOL (Chat, Pos. 5857) / User681: @xQc brother that hair makes you look like a mushroom (Chat, Pos. 5873)

5.2.4 Quantitative overview of potential toxicity in all layers

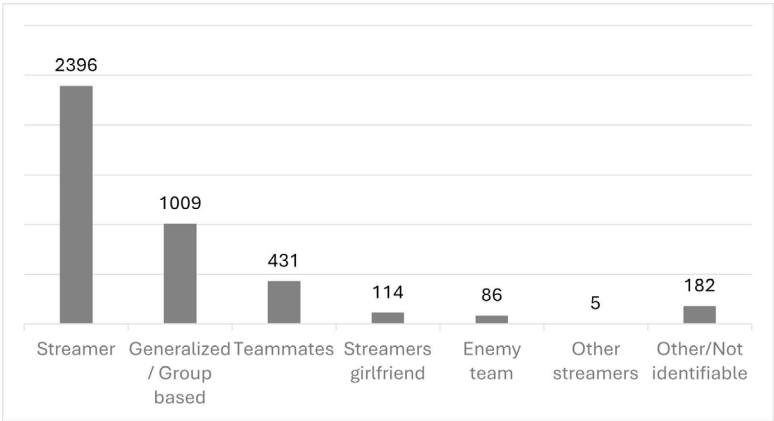
As visible in Fig 1. the streamer is the most popular target of toxicity since he is the center of attention for anything happening in chat. It is noteworthy that this is often the result of a sort of toxic echo: Whenever he brags, down-

press excitement or Schadenfreude. While it could be a literal interpretation of the in-game action, the term afk is often used to say someone is so bad it is equal to them not playing at all. In another comment someone writes „alt+f4“ (Chat, Pos. 2953) (a command on the pc to shut the window you are currently using) to say ‘stop playing the game’.

20 An example for this would be the variations of LOL, LUL, OMEGALUL or omE (all expressing surprise or enjoyment) that can be found both in emotes or texts.

talks something, or attacks co-players, the chat will either join in on the attack or will react in a counterattack. There is also a high degree of group-based or generalized toxicity, for example, in hate speech regarding racism, sexism and ableism. The streamer’s girlfriend AikoBliss is only partially visible during the stream but receives a lot of attention, both with positive and negative messages. While there are some comments that can be read as sexist, most of them request her to do something (for example, strike a pose with the emote ‘PotFriend’).

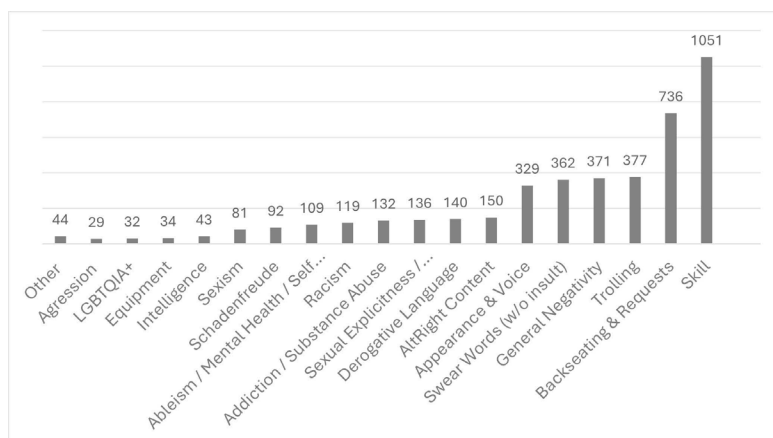
Fig. 1: Targets of potential toxicity



Some of the presented examples in this text might seem comparably harmless but there are also very toxic comments, for example, in the form of explicit sexual harassment, especially towards the male streamer, which often only become apparent through context. For example, when he says about the game “I am actually so solid I am going to bust everybody” (Stream (Audio), Pos. 1349), which already has sexual connotations, a few people in chat answer “me next” (Chat, Pos. 1361) or post more explicit content.

Through the open annotation process, a mix of categories was annotated that can provide a better insight into the different content and their quantity:

Fig. 2: Annotated themes of potential toxicity.



As visible in Fig. 2, most of the potential toxicity relates to the game by criticizing skill or backseating²¹ but also trolling and swear words as mild forms of toxicity are very present. Under the term ‘General Negativity’, different types of toxic actions that can spread a negative mood (for example, complaining about the chosen game) were annotated, but also attacks that did not fit any of the other categories and seemed too minor to start a new section. Throughout the annotation process it became apparent that the categories have very different content because they described both types of actions, like trolling or backseating and topics such as racism. This led to annotations sometimes overlapping different categories, which affected the clarity. For further research with a qualitative approach, it therefore could make sense to differentiate not only between targets but also types and topics of a toxic action. Furthermore, a classification can also include a description of the linguistic or visual means that were used to convey the toxicity, as discussed in the examples before:

21 For this, only sentences with explicit imperatives were counted and polite suggestions or single words (like the name of an item) were excluded and would have resulted in a much higher number.

Tab. 4: Multimodal Annotation Grid

Target	Type	Topic	Linguistic / Visual Means
Generalized Directed unknown, streamer, chat, co-players (enemy & team), ad- ditional streamers, people on stream	spam, backseat- ing, requests, aggressive behaviour, inappropri- ate language, (ad-hominem) attacks, cri- tique, doxxing, Schadenfreude, down talking something	racism, sex- ism, ableism, LGBTQIA+, poli- tics, religion, ap- pearance, ageism, mental health, skill, equipment, intelligence	<u>Linguistic</u> swear words, derogative terms, raised voice, rhetor- ical questions, anaphor, 'easy language', sarcasm, verbal irony, homophony <u>Both</u> distortion, intertextual- ity, repetition, metaphors, ASCII art <u>Visual</u> mimic, gestures, move- ments, images, all caps

5.3 Meta Discourse

To better understand the perceptions and mechanisms of toxicity on Twitch, meta-discursive discussions related to toxicity were also annotated. This includes comments that call out toxic behaviour, counter speech, comments that instigate more toxic behaviour or reflections on moderation or community norms.

Some people in chat tried to counter toxicity by making it explicit, which could enable a discussion of appropriate behaviour. In the annotated material, four different types of meta-communication about the toxicity of the streamer were identified:

Tab. 5: Reflections on toxic behaviour in chat

Irony	Statement	Perspective of victim	Requests
LOLLOL toxic much (Chat, Pos. 2990) nothing better thn a bit of toxicity in dota (Chat, Pos. 5802) W TOXIC (Chat, Pos. 2696)	rude to your team WeirdChamp- WeirdChamp (Chat, Pos. 2044) D:D: so mean (Chat, Pos. 2841) ur soo toxic (Chat, Pos. 3782) minus behavior score (Chat, Pos. 5626)	True but poor him (crying cat emote) xqc (Chat, Pos. 5620) Buddy comes home from work, he finds a dota 2 match and instantly gets into a match with his favourite streamer. He picks support to support his streamer. Streamer blames him. Buddy becomes sad. Buddy no longer wants to play. (Chat, Pos. 1053)	be nice (Chat, Pos. 5508) BEAN ICE SAJSAJ (Chat, Pos. 2053) SAJSAJ apologize (Chat, Pos. 5533) D:D: dont be mean (Chat, Pos. 2848) @xQc Dont be rude mofo (Chat, Pos. 5524)

While one user asks xQc to watch his language specifically for a younger audience²² sometimes his use of milder curse words is also commented upon since his audience expects a certain kind of language.²³ Especially the second example that takes the perspective of the victim of an attack in table 5 was posted repeatedly in chat and creates a narrative about his co-player suffering from the toxicity (Chat, Pos. 1053). While the streamer does not pay attention to the chat for most parts of the stream and will not likely be influenced by this attempt of counter speech, it can show people in chat that not everyone accepts toxic behaviour, which can have an impact on regulating the community.

But the streamer himself also partially reflects on the limits of what is sayable: While he queues for a game on a South American server he comments: “South American Dota by the way is– I can’t I don’t know if I can say what I want to say– like it’s just it’s just it’s just special” (Stream (Audio), Pos. 175–180).

- 22 “WeirdCatWeirdCat stop swearing all the time my kids are watching” (Chat, Pos. 5207). The stream was not marked as +18 and therefore should not contain swearwords in accordance with Twitch’s community guidelines.
- 23 One chat message points out that xQc’s way of speaking sounds like the community of the competitive online game Valorant (“HE HAS THE VAL ACCENT LMFAOLMFAOO” (Chat, Pos. 5613)), which is connected to the attacking of co-players (Çakır 2023).

While this sparks countless of racist comments towards South American players, it also opens a meta discourse where people urge him to clarify (“UNIQUE OR SPECIAL?” (Chat, Pos. 183), “wdym SusSus” (Chat, Pos. 186), “say it” (Chat, Pos. 185)) or give him the free pass to be clearly racist (“i won’t tell anyone if you say it” (Chat, Pos. 182)). The chat is also flooded with the ‘IFISPEAK’ emote,²⁴ and xQc receives the advice to “tread carefully” (Chat, Pos. 209) due to the potential backlash. This is made more explicit by a comment about a fellow streamer who faced consequences after talking negatively about Brazilians: “bro u gonna get hated like Zombs for no reason (skull emoji)” (Chat, Pos. 203). It is not clear, however, if viewers actually disagree with potentially racist utterances or just fear for their streamer to get punished (e.g. with a shitstorm) for his behaviour.

This also becomes apparent when a user writes “XQC you should say ‘OSUJDAYU’ when he says such a things on stream” (Chat, Pos. 5429), which stands for the Russian word осуждаю and means something like ‘I judge’ and is commonly used as a defence for offensive language so streamers don’t get banned for the behaviour of their community or co-players. This practice of claiming to judge someone for toxicity seems to be a ritualized defence mechanism without the necessity to believe in the statement.

There is an ongoing interaction in chat between users and moderators about accepted behaviour and forms of punishment, for example, when viewers request or question decisions for banning:

User 266: @User1121 mods GTABGTAB (Chat, Pos. 1713) / User266: @User1022 MODSMODS BANDBAND BANDBAND (Chat, Pos. 1793) / User2398: @xQc MOD HIM (Chat, Pos. 5934) / User73: @User323 game complainer MODSMODS (Chat, Pos. 795) / User438: holy shit he said that and didn’t get banned i said the word and got 5 year ban (Chat, Pos. 5964)

Especially the emotes ‘GTAB’ (get that ass banned), ‘Band’ (showing a band but using homophony) and ‘MODS’ (to inform moderators) are used in this context, but this practice is also seen as negative actions of a “snitch @User73 :(“ (Chat, Pos. 814).

The collected material (though only discussed here with examples) shows there is an ongoing awareness and debate even in a seemingly toxic community

24 It features the football coach Jose Mourinho during a post-game interview. He was asked a controversial question and said, “If I speak I am in big trouble”.

about what is sayable both in chat and for the streamer. Moderation actions like banning can help regulate behaviours in chat (Seering 2020) but are often not visible for the rest of the community. While the different types of counter speech towards the streamer don't seem to have a great impact on his behaviour they defend values of civilized speech. These actions can be impactful for the community, as challenging toxic practices makes user's disagreement visible and might create awareness in more viewers to question them. But as an article from the corpus B (Shrivastava 2022) reports, xQc has in the past banned users who call him toxic from his stream and labeled them 'Social Justice Warriors' (a term that ridicules people who stand up for more progressive causes), which is another form of regulating what is sayable in this community.

6. Conclusion

The article gave a first impression of the many challenges and the great potential to investigate toxicity in multimodal streams on Twitch. The paper uses insights from toxicity research both on Twitch and in gaming as a basis to expand it with a multimodal approach. For this, a methodology to transcribe Twitch streams was developed, and a suitable sample (corpus A) was found through discourse analysis. In order to be mindful of the different conceptions of what can be considered as toxic, research articles, Twitch communities (corpus C) and an excerpt of the public discourse (corpus B) were analysed. This led to the decision to annotate all cases of *potential* toxicity and to conceptualize toxicity tied to Wittgenstein as a *Sprachspiel* that can be a communicative practice.

In the analysis, the three focal points of Triggers, Potential Toxicity and Meta Discourse were discussed using examples from the corpus and a quantitative overview. A main distinction between existing research on triggers in the context of gaming is that multimodal data shows all actions that can elicit negative feelings in players, allowing the impact of smaller triggers to be studied in more detail.

For annotating Potential Toxicity, the analysis was broken down into the three layers of game, streamer and chat, but still reflected their interconnect-edness. In the game layer, multiple actors were able to use voice chat, written chat or in-game actions like taunts to attack each other. It was mostly team-mates being toxic to one another rather than enemy players, and the streamer was the only person who was directly toxic in voice chat.

For the stream layer, it was shown that both game actions and the verbal layer were used for toxicity, but also generalized toxicity, like the usage of swear words, was very common. For the visual layer, the video of the streamer illustrated the shift to a negative mood and aggressive gestures intensified toxic speech. The style of bantering in the interaction of the community also became apparent, for example, in the streamers' girlfriend mocking his complaints about co-players with exaggerated gestures.

Regarding the chat, emotes (as a form of meme culture) and 'crowd speak' were discussed to understand the interaction dynamics between viewers. Principles like repetition, distortion and voice taking helped to classify messages that can be highly toxic and would not be easy to detect with current approaches. The chat had diverse topics, but the primary focus of all toxic attacks was the streamer himself, who was criticized for his gameplay, appearance, taste in music or behaviour. Whenever he attacked co-players, it resulted in a form of echo that would partially target his co-players but mostly fire back at him, which could be interpreted as some form of punitive mechanism for toxic behaviour. This also becomes apparent in the analyzed meta-discourse, where the streamer either gets warned of consequences, there are attempts to push him to overstep by being explicitly racist or forms of counter-speech in chat.

Due to the sheer amount of data, many phenomena could not be discussed here but will be the object of future research, which will, for example, demonstrate the interconnectedness of the three layers of streams in a map visualisation (showing all trigger events and toxicity in the three layers according to the timeline). Another step is also to compare this manual annotation process and its results with automated approaches and to investigate the functions of the observed behaviours for community building.

Bibliography

- Adams, Philippa R.; Sommers, Maria; Scholl, Ben; Lusoli, Alberto; Chun, Wendy H. K. (2025): The revenge of 'tribal' toxic masculinity: hate raids on Twitch. In: *Journal of Visual Culture*. 24, 1. URL: <https://doi.org/10.1177/14704129251327916>.
- Almerexhi, Hind; Kwak, Haewoon; Salminen, Joni; Jansen, Bernard J. (2019): Detecting Toxicity Triggers in Online Discussions. In: *Proceedings of the*

- 30th ACM Conference on Hypertext and Social Media (HT'19). Hof, Germany. P. 291–292. URL: <https://doi.org/10.1145/3342220.3344933>.
- Anker Nexø, Louise (2024): Toxic Behaviours in Esport: A Review of Data-Collection Methods Applied in Studying Toxic In-Gaming Behaviours. In: *International Journal of Esports*. 1, 1. URL: <https://www.ijesports.org/article/127/html> [accessed 03.02.26].
- Beres, Nicole A.; Frommel, Julian; Reid, Elizabeth; Mandryk, Regan L.; Klarkowski, Madison (2021): Don't You Know That You're Toxic: Normalization of Toxicity in Online Gaming. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Yokohama, Japan. ACM New York, NY, USA. P. 1–15.
- Blackburn, Jeremy; Kwak, Haewoon (2014): STFU NOOB! Predicting Crowdsourced Decisions on Toxic Behavior in Online Games. In: *WWW '14: Proceedings of the 23rd international conference on World wide web*. P. 877–887. URL: <https://doi.org/10.1145/2566486.2567987>.
- Cai, Jie; Wohn, Donghee Y. (2019): What are Effective Strategies of Handling Harassment on Twitch? Users' Perspectives. In: *Conference companion publication of the 2019 on computer supported cooperative work and social computing*. P. 166–170.
- Çakır, Gökhan (2023): What is the VALORANT accent? Explained. “Toxicity of our city”. In: *Dotesport.com*. URL: <https://dotesports.com/valorant/news/what-is-the-valorant-accent-explained> [accessed 08.05.26].
- Cheng, Justin; Bernstein, Michael; Danescu-Niculescu-Mizil, Christian; Leskovec, Jure (2017): “Anyone can become a troll: Causes of trolling behavior in online discussions”. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW '17*. New York, NY, USA: Association for Computing Machinery. P. 1217–1230. URL: <https://doi.org/10.1145/2998181.2998213>.
- Dancygier, Barbara; Vandelandotte, Lieven (2025): Embodiment and simulated interaction in online stance expression. In: *Front. Psychol.* 15, 1479825. URL: <https://doi.org/10.3389/fpsyg.2024.1479825>.
- Dreier, Lukas; Pirker, Johanna (2023): Toxicity in Twitch Live Stream Chats: Towards Understanding the Impact of Gender, Size of Community and Game Genre. In: *2023 IEEE Conference on Games (CoG)*, 21.08.2023, 1. URL: <http://dx.doi.org/10.1109/CoG57401.2023.10333159>.
- Dungca, Gio (2021): 5 Most Infamous Gaming Streamers of All Time In: *Clutchpoints*. URL: <https://clutchpoints.com/gaming/5-most-infamous-gaming-streamers-of-all-time> [16.05.26].

- Diwanji, Vaibhav; Reed, Abigail; Ferchaud, Arienne; Seibert, Jonmichael; Weinbrecht, Victoria; Sellers, Nicholas (2020): Don't just watch, join in: Exploring information behavior and copresence on Twitch. In: *Computers in Human Behavior*. 105, 106221.
- Edwards, Jane (1993): Principles and Contrasting Systems of Discourse Transcription. In: *Talking Data. Transcription and Coding in Discourse Research*. Ed. by Jane A. Edwards & Martin D. Lampert. Hillsdale: Lawrence Erlbaum Associates Inc. P. 3–31.
- ElSherief, Mai; Nilizadeh, Shirin; Nguyen, Dana; Vigna, Giovanni; Belding, Elizabeth (2018): Peer to Peer Hate: Hate Speech Instigators and Their Targets. In: *Proceedings of the ICWSM 2018*. 12, 1. URL: <https://arxiv.org/pdf/1804.04649v1> [accessed 25.04.26].
- Ford, Colin; Gardner, Dan; Horgan, Leah Elaine; Liu, Calvin; Tsaasan, A. M.; Nardi, Bonnie; Rickman, Jordan (2017): Chat Speed OP PogChamp: Practices of Coherence in Massive Twitch Chat. In: *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems – CHI EA '17*. P. 858–871. URL: <https://doi.org/10.1145/3027063.3052765>.
- Harpstead, Erik; Sebastian Rios, Juan; Seering, Joseph; Hammer, Jessica (2019): Toward a Twitch Research Toolkit: A Systematic Review of Approaches to Research on Game Streaming. In: *Proceedings of the Annual Symposium on Computer- Human Interaction in Play*. P. 111–119.
- Kwak, Haewoon; Blackburn, Jeremy (2014): Linguistic analysis of toxic behavior in an online video game. In: *Social Informatics. SocInfo 2014*. Ed. Luca Aiello, Daniel McFarland. Springer, Cham. URL: https://doi.org/10.1007/978-3-319-15168-7_26.
- Legewie, Nicolas; Nassauer, Anne (2018): YouTube, Google, Facebook: 21st Century Online Video Research and Research Ethics. In: *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*. 19, 3 URL: <https://doi.org/10.17169/FQS-19.3.3130>.
- Liemann, Christina (2023): Hysterische Feministinnen, Soy Boys und SJWs. Geschlecht als multimodale Argumentationsstrategie in rechts-alternativen YouTube-Videos. In: *Aptum. Zeitschrift für Sprachkritik und Sprachkultur*. 19, 1. P. 2–17.
- Lee, Matthew; Johnson, Daniel; Tanjitpiyanond, Porntida; Louis, Winnifred R. (2021): It's habit, not toxicity, driving hours spent in DOTA 2. In: *Entertainment Computing*, 100472. URL: <https://doi.org/10.1016/j.entcom.2021.100472>.

- Mall, Raghvendra; Nagpal, Mridul; Salminen, Joni; Almerexhi, Hind; Jung, Soon-Gyo; Jansen, Bernard J. (2020): Four Types of Toxic People: Characterizing Online Users' Toxicity over Time. In: Association for Computing Machinery. URL: <https://doi.org/10.1145/3419249.3420142>.
- Martens, Marcus; Shen, Siqi; Iosup, Alexandru; Kuipers, Fernando (2015): Toxicity detection in multiplayer online games. In: 2015 International Workshop on Network and Systems Support for Games (NetGames). P. 1–6. URL: <https://doi.org/10.1109/NetGames.2015.7382991>.
- Nobata, Chikashi; Tetreault, Joel; Thomas, Achint; Mehdad, Yashar; Chang, Yi (2016): Abusive language detection in online user content. In: Proceedings of the 25th International Conference on World Wide Web, ser. WWW '16. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee. P. 145–153. URL: <https://doi.org/10.1145/2872427.2883062>.
- Piot, Paloma; Martín-Rodilla, Patricia; Parapar, Javier (2024): MetaHate: A Dataset for Unifying Efforts on Hate Speech Detection. In: Proceedings of the International AAAI Conference on Web and Social Media. 18, 1. P. 2025–2039. URL: <https://doi.org/10.1609/icwsm.v18i1.31445>.
- Recktenwald, Daniel (2017): Toward a Transcription and Analysis of Live Streaming on Twitch. In: Journal of Pragmatics. 115. P. 68–81.
- Saleem, Haji M.; Dillon, Kelly P.; Benesch, Susan; Ruths, Derek (2016): A web of hate: Tackling hateful speech in online social spaces. Proceedings of Text Analytics for Cybersecurity and Online Safety (TA-COS), LREC Portoroz. P. 1–9.
- Seering, Joseph (2020): Reconsidering community self-moderation: The role of research in supporting community-based models for online content moderation. In: Proceedings of the ACM on Human-Computer Interaction. 4, 107. <https://doi.org/10.1145/3415178>.
- Seering, Joseph; Kraut, Robert; Dabbish, Laura (2017): Shaping pro and anti-social behavior on twitch through moderation and example-setting. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW '17. New York, NY, USA: Association for Computing Machinery. P. 111–125. URL: <https://doi.org/10.1145/2998181.2998277>.
- Shrivastava, Aarnesh (2022): Twitch streamers who absolutely lost it in front of their viewers during a livestream. In: sportskeeda. URL: <https://www.sportskeeda.com/esports/5-twitch-streamers-absolutely-lost-viewers-live-stream> [16.05.26].

- Sjöblom, Max; Hamari, Juho (2016): Why do people watch others play video games? An empirical study on the motivations of twitch users. In: *Computers in Human Behavior*. 75. P. 985–996. URL: <http://dx.doi.org/10.2139/ssrn.2779543>.
- Sheth, Amit; Shalin, Valerie, L.; Kursuncu, Ugur (2022): “Defining and detecting toxicity on social media: context and knowledge are key”. In: *Neurocomputing*. 490. P. 312–318. URL: <https://doi.org/10.1016/j.neucom.2021.11.095>.
- Shifman, Limor (2014): *Memes in Digital Culture*. Cambridge, Massachusetts: The MIT Press.
- Stoop, Wessel; Kunneman, Florian; van den Bosch, Antal; Miller, Ben (2019): Detecting harassment in real-time as conversations develop. In: *Proceedings of the Third Workshop on Abusive Language*. P. 19–24. URL: <https://doi.org/10.18653/v1/W19-3503>.
- Thompson, Joseph J.; Leung, Betty H. M.; Blair, Mark R.; Taboada, Maite (2017): Sentiment analysis of player chat messaging in the video game StarCraft 2: Extending a lexicon-based model. In: *Knowledge-Based Systems*. 137. P. 149–162. URL: <https://doi.org/10.1016/j.knosys.2017.09.022>.
- Twitchtracker: Twitch Statistics & Charts. URL: <https://twitchtracker.com/statistics> [15.05.26].
- Waseem, Zeerak; Davidson, Thomas; Warmusley, Dana; Weber, Ingmar (2017): Understanding abuse: A typology of abusive language detection subtasks. In: *Proceedings of the First Workshop on Abusive Language Online*, Vancouver. P. 78–84.
- Wittgenstein, Ludwig (1967): *Philosophische Untersuchungen*. Frankfurt a.M. 1967.
- Zhou, Xuhui; Sap, Maarten; Swayamdipta, Swabha; Choi, Yejin; Smith, Noah (2021): Challenges in automated debiasing for toxic language detection. In: *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics*. P. 3143–3155. URL: <https://doi.org/10.18653/v1/2021.eacl-main.274>.
- Zolides, Andrew (2023): Toxic community policing: Weaponizing moderation tools on Twitch. In: *Real Life in Real Time: Live Streaming Culture*. Ed. by Johanna Brewer, Bo Ruberg, Amanda L. L. Cullen, Christopher J. Persaud. Cambridge, MA: MIT Press. P. 131–144. URL: <https://doi.org/10.7551/mitpress/14526.003.0013>.

»Next Level Frauenhass«

Tabu in der diskursiven Auseinandersetzung zu *No Mercy* als toxische Kommunikationsverstrickung

Patricia Bau

Abstract This article examines the discourse surrounding the game *No Mercy* (2025) by *Zerat Games* in terms of its entanglement with taboo-breaking and toxicity. The game was heavily criticized in April 2025 because of its objective to rape female characters, including, according to media reports, the player character's own mother. This highly taboo-breaking content triggered worldwide campaigns against the game and led to media reporting. The paper explores how taboo-breaking and toxicity intertwine and influence each other. Therefore, the studio's statement published on Steam, articles about the game and comments under these articles are analyzed, and the communicative practices are classified from a taboo and toxicity theory perspective. Moralizing, whataboutism, and accusations of attention seeking are particularly relevant here. It is also interesting to note that both critics and supporters of the game use similar (toxic) discourse practices.

Keywords taboo, taboo-breaking, toxic discourse practices

Abstract In diesem Artikel wird der Diskurs über das Videospiel *No Mercy* (2025) des Studios *Zerat Games* hinsichtlich der Verstrickung von Tabubrüchen und Toxizität untersucht. Das Spiel wurde im April 2025 stark kritisiert, da es dessen Ziel ist, weibliche Figuren – darunter laut Medienberichten auch die Mutter des Protagonisten – zu vergewaltigen. Dieser in hohem Maße tabubrechende Inhalt löste weltweit Kampagnen gegen das Spiel aus und führte zu einer medialen Berichterstattung. Wie Tabubrüche und Toxizität ineinandergreifen und sich bedingen, wird in diesem Artikel herausgearbeitet. Dazu werden das Statement des Studios auf Steam sowie Artikel

über das Game und Kommentare unter diesen Artikeln analysiert und die kommunikativen Praktiken aus tabu- und toxizitätstheoretischer Perspektive eingeordnet. Dabei sind insbesondere Moralisierung, Whataboutismus und Vorwürfe der Aufmerksamkeitsgenerierung relevant. Interessant ist auch, dass sowohl Kritiker*innen als auch Befürworter*innen des Games ähnliche (toxische) Diskurspraktiken nutzen.

Schlagworte Tabu, Tabubrüche, toxische Diskurspraktiken

1. Einleitung

Im Spiel *No Mercy* beobachtet die männliche, spielbare Figur die eigene Mutter beim Fremdgehen. Begründet durch diese Erfahrung ist das Spiel darauf ausgelegt, so viele weibliche NPCs wie möglich zu vergewaltigen. Das Hauptziel ist die Vergewaltigung der Mutter der Spielfigur. Durch diesen Inhalt erhielt das Spiel sehr viel Kritik und wurde von der Kauf-Plattform Steam entfernt.

Dass das Spiel so stark kritisiert wurde, liegt an der in hohem Maße tabubrechenden Spielhandlung, denn sowohl nicht-einvernehmlicher Geschlechtsverkehr als auch Inzest bzw. deren Darstellung sind starke Verstöße gegen geltende gesellschaftliche Normen und lassen sich toxischer Männlichkeit zuordnen.

»Toxic geek masculinity« (Adams et al. 2025) und »toxic gamer culture« (Fox, Tang 2017: 1291) gelten als weit verbreitet in Gaming-Räumen. Adams et al. (2025) gehen davon aus, dass sich die toxische Gamer-Kultur aus der »geek masculinity« (ebd.) heraus entwickelt hat, womit sie Toxizität elementar auf Männlichkeit zurückführen. Die toxischen Elemente von Männlichkeit führen sie folgendermaßen aus:

While the elements of what constitutes the ideal figure of masculinity are not static, in a US context, hypermasculine values have remained relatively fixed since the mid-20th century and have tended to champion dominance and control, aggression, self-reliance, physical and sexual prowess, violence [...]. (ebd.)

Ebendiese Aspekte toxischer Männlichkeiten finden sich auch in *No Mercy* wieder und manifestieren sich verdichtet in (inzestuöser) Vergewaltigung. Der

Diskurs um das Game ist somit auch ein Teildiskurs über toxische Maskulinität.

Dieser Beitrag beschäftigt sich mit den Verbindungen von Tabu und Toxizität auf theoretischer Ebene und analysiert diese Beziehungen am Beispiel des (deutschsprachigen) Diskurses um *No Mercy*. Dabei wird herausgearbeitet, welche als toxisch wahrnehmbaren Kommunikationsformen über das Game im Diskurs vorzufinden sind und welche Reaktionen darauf existieren. Die komplexe Beziehung von Tabubrüchen im Game, Reaktionen auf diese sowie Kommentare zur Diskussion und deren toxisches Potential werden in diesem Beitrag aus linguistischer Perspektive erörtert. Dazu werden das Statement des Entwickler*innen-Teams¹ auf Steam, Artikel über das Spiel und Kommentare unter diesen Artikeln analysiert.

Zunächst werden tabutheoretische Grundlagen (2.) dargestellt und anschließend mit Überlegungen zur Toxizitätsforschung verknüpft (3.). Nachfolgend findet eine tabutheoretische Einordnung der diskursiven Auseinandersetzung zu *No Mercy* (4.) statt, wobei Moralisierung (4.1), Whataboutismus (4.2) sowie Vorwürfe der Aufmerksamkeitsgenerierung (4.3) untersucht werden. Abschließend werden die Ergebnisse im Fazit (5.) zusammengefasst und der Zusammenhang von Tabubrüchen im Game und toxischen Diskurspraktiken aufgezeigt.

2. Tabutheorie

2.1 Definitorische Einordnung

Tabus wurden nach ihrer ersten Beschreibung durch James Cook Ende des 18. Jahrhunderts zunächst als ethnologisches und theologisches Fremdwort verstanden, das mit den religiösen Bräuchen auf polynesischen Inseln verbunden wurde (Benthien, Gutjahr 2008: 7; Kuck 2021: 229).² Heute gelten Tabus hingegen als »kulturelle Sagbarkeits-, Zeigbarkeits- und Machbarkeitsgrenzen, die

- 1 Es ist unklar, wer das Spiel entwickelt hat, da die Webseite des Studios lediglich zum Game führt und keinerlei weitere Informationen aufweist (Zerat Games Website).
- 2 Dabei gilt zu beachten, dass der Tabubegriff einer eurozentristischen ethnologischen Forschung entstammt und als »Aneignung eines fremdkulturellen Phänomens« (Egert 2002: 16) verstanden werden kann. Er diente zur Fremdbezeichnung von Gesellschaften, die aus ebendieser Perspektive als »primitiv« und »unkultiviert« galten und wurde zur Erforschung der »Anderen« genutzt (Frietsch 2008: 11).

das soziale Zusammenleben ordnen« (Kuck 2021: 227). Dabei befindet sich ihr Geltungsraum häufig in sozialen und moralischen Kontexten (Braun 2007: 7).

Die Ebenen, auf welchen Tabus wirken, sind allumfassend, sodass Tabus Begrenzungen von Handlungen, Kommunikation, Sprache, Abbildungen bzw. Darstellungen, Wahrnehmungen und Gefühlen umfassen (Kuck 2021: 230; Schröder 2008: 58).

2.2 Soziale Wirkung

Obwohl »Tabus [im Alltagsverständnis] als etwas Negatives und Irrationales« (Mildenberger, Schröder 2012) gelten, besitzen sie wichtige soziale Funktionen. »Man kann Tabus mit Normen vergleichen, die uns sagen, was wir unter bestimmten Bedingungen machen sollen, müssen, dürfen, oder eben nicht.« (Kuhn 1987: 23) Tabus sind somit implizite Gesetze, die gruppen-, kultur- und zeitabhängig sind (Lautenschläger 2021: 217). Eine Missachtung dieser führt zu einer sozialen Sanktion für diejenige Person, die den Tabubruch ausführt (Kuck 2021: 227; Rada 2001: 32). Tabus regulieren somit soziales Verhalten »jenseits des Gesetzes« (Oetl 2019: 28). Gemeinsame geteilte Werte sowie festgelegte Handlungsräume sind konstitutiv für Gesellschaften, weshalb Tabus gemeinschaftsbildend wirken, denn die Mitglieder einer Gruppe teilen das gemeinsame Wissen darüber, was in ebendieser als tabuisiert gilt. »Gesellschaftliches Zusammenleben ist ohne Tabus undenkbar, weil der Einzelne ansonsten ohne implizites Wissen über Verhaltensgebote und ohne Schutz vor Verletzung seiner persönlichen Rechte leben würde.« (Gutjahr 2008: 20) Des Weiteren sind gemeinschaftliche Werte, die in Tabus kondensiert sind, identitätsstiftend.

Dennoch sind Tabus als Grenzen des gesellschaftlich Akzeptierten überschreitbar.

Tabus [betreffen] meistens etwas [...], was nicht getan, gesagt, gedacht, gefühlt, auch nicht gewusst und berührt werden darf – dennoch aber machbar, sagbar, denkbar, fühlbar und erkennbar sowie berührbar ist; ansonsten brauchte es ja nicht durch ein Tabu geschützt zu werden. (Schröder 2008: 53)

Ein Tabubruch resultiert in sozialen Konsequenzen und löst i.d.R. Überraschung und/oder Verwirrung aus (Birk, Kaunzner 2009: 402). Die beiden Verhaltensweisen, die nach dem Brechen gesellschaftlicher Normen beobachtet werden können, sind Angriffs- und Fluchtverhalten (ebd.). Tabubrecher*in-

nen können gemieden werden oder durch Empörung sowie verbale oder körperliche Anfeindungen attackiert werden (ebd.). Wer Tabus gebrochen hat, kann diese sozialen Sanktionen durch konventionelle Reparaturmechanismen wie Entschuldigen und Erklären vermeiden bzw. minimieren (ebd.). Während Tabus auf verschiedenen Ebenen gebrochen werden können – durch Handlungen, Abbildungen, Themen, konkrete Sprachhandlungen – müssen Tabubrecher*innen gewöhnlich sprachlich auf ihre Normverletzung reagieren und Sprechakte anwenden.

Tabubrüche können bewusst eingesetzt werden, um zu provozieren und damit Aufmerksamkeit zu generieren. Dies kann einen Versuch der Enttabuisierung und damit einhergehend des Anstoßes zu gesellschaftlichem Wandel darstellen (Schröder 2008: 62).

3. Tabu und Toxizität

Das Verhältnis von Tabu und Toxizität ist multidimensional: Beide Konzepte können sich gegenseitig bedingen und ihre Beziehung ist wechselseitig. Toxisches Verhalten kann als »negative behaviour« (Dreier, Pirker 2023) betrachtet werden und meint im Bereich der Kommunikation die »Schädlichkeit/ Giftigkeit eines Kommentars für die Konversation und damit auch für andere Konversationsteilnehmer« (Demus et al. 2023: 264). Somit bricht toxisches Verhalten häufig gesellschaftliche Normen und damit Tabus. Relevant dabei ist, dass Toxizität – ebenso wie Tabus – gruppenspezifisch ist (ebd.: 265). Unterschiedliche Personen(gruppen) nehmen verschiedenes Verhalten als toxisch und tabubrechend wahr, wodurch unterschiedliche Meinungen und Reaktionen auf von einigen Personen als toxisch wahrgenommene Handlungen erfolgen können, die von sozialer Sanktion bis hin zu Verteidigung reichen können. Grundsätzlich müssen »die Umstände, unter denen etwas für jemanden *toxisch*, d.h. schädlich ist, [...] aus der diskursiven Situation heraus erschlossen werden.« (Hoffmeister 2023a; Herv. i. O.)

Wenn toxisches Verhalten von Personen, die dieses als solches bewerten, sozial sanktioniert wird – etwa durch das öffentliche Kundmachen des eigenen Unmuts oder Aufforderungen zum Boykott, wenn es sich um Waren und Dienstleistungen handelt – treten typische Reaktionsketten aus der Kommunikation um Tabubrüche auf. Die Reaktionen auf als toxisch Wahrgenommenes kann dabei jedoch selbst toxische Elemente aufweisen. »[O]nline toxicity has a contagious nature, and thus, one toxic comment has the potential to at-

tract more toxic comments quickly.« (Almerekhi et al. 2019; 291) Auch Demus et al. (2023: 273) halten fest, dass toxische Kommentare die Entstehung von Hass erhöhen. Dabei wird sowohl allgemeiner Hass als auch Hass gegen konkrete Personen gefördert (ebd.). »Eine hohe Toxizität geht häufig mit einem negativen Sentiment (Stimmung, Gefühlslage) des Kommentarautors einher.« (Demus et al. 2023: 265) Findet ein Tabubruch statt, so werden dementsprechend negative Emotionen ausgelöst, welche wiederum das Aufkommen von Reaktionen mit toxischen Elementen fördern, womit beide Phänomene durch negative Gefühle miteinander verbunden sind und einander dadurch bedingen bzw. verstärken.

Toxische Kommunikation führt häufig zu Stereotypisierung und Depersonalisierung von Outgroup-Mitgliedern und ist somit ein Mechanismus des Otherings³ und damit einhergehend der Identitätsstiftung (Fox, Tang 2017: 1292). »It can manifest as aggression, harassment, and bullying, disrupting online communities.« (Dreier, Pirker 2023) Aus linguistischer Perspektive zählen zu toxischen Elementen in der Kommunikation etwa Beleidigungen, der Rückgriff auf negative Stereotype und in hohem Maße negative Bewertungen bzw. Verurteilungen. Relevant für die linguistische Betrachtung toxischer Kommunikation ist außerdem die Kommunikationssituation: Zu beachten gilt – neben den sprachlichen Elementen, die Toxizität evozieren – wer von wem womit angegriffen wird.

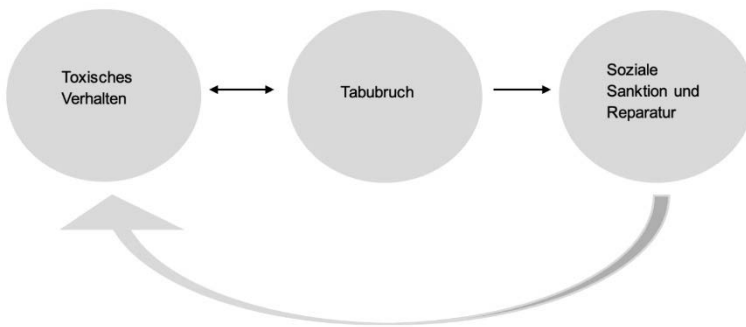
Auch die Anonymität des Internets ist ein relevanter Faktor für Tabubrüche und toxisches Verhalten (Kuck 2021: 232; Fox, Tang 2017: 1294). Online-Anonymität wirkt als Schutzmechanismus vor Sanktionen des eigenen (kommunikativen) Verhaltens im realen Leben. Sowohl das Begehen von Tabubrüchen als auch Toxizität in der Interaktion werden also durch die Bedingungen von Online-Kommunikation ermöglicht und deren Aufkommen kann somit erhöht sein. Da Gaming-Diskurse primär online stattfinden, sind diese Diskurse anfällig für toxisches und tabubrechendes Verhalten.

Tabu und Toxizität sind zusammenfassend Phänomene, die häufig simultan stattfinden, sich gegenseitig verstärken und durch Anonymität vermehrt auftreten. Toxisches Verhalten bricht i.d.R. Tabus, sodass auf Toxizität meist Reaktionen aus der Kommunikation über Tabubrüche folgen: soziale Sanktionen, die durch Reparaturmechanismen bewältigt werden können. Toxizität

3 Grundsätzlich gilt: »[D]as Fremde ist toxisch« (Hoffmeister 2023b: 455) und wird Personen zugeschrieben, von denen man sich distanzieren möchte, was auch für das Konzept von Tabus gilt, denn auch hier sind die Tabubrechenden die Nicht-Zugehörigen.

und Tabubrüche sind zudem Auslöser von negativen Emotionen und ausgelöst durch Negativität. Sanktionen und Reparaturen können wiederum selbst toxische Elemente aufweisen und eine erneute Reaktionskette von Toxizität und Tabubruch auslösen. Diese Verkettung wird in der folgenden Übersicht veranschaulicht:

Abb. 1: Verkettung von Tabu und Toxizität



4. Tabu und Toxizität in der diskursiven Auseinandersetzung zu *No Mercy*

In den nachfolgenden Unterkapiteln wird der Diskurs über *No Mercy* in Bezug auf die Verschränkung von Tabu und Toxizität betrachtet und eingeordnet. Dafür wurde eine beispielhafte Auswahl an Texten getroffen, anhand derer die zuvor präsentierten theoretischen Überlegungen überprüft und ausgeweitet sowie spezifische sprachliche Muster herausgearbeitet werden. Das dafür verwendete Korpus umfasst die englischsprachige Erklärung des Entwickler*innen-Teams *Zerat Games* (1.029 Token), welche auf Steam veröffentlicht wurde, sowie deutschsprachige Artikel, die sich mit dem Game und dessen Entfernung von Steam auseinandersetzen, von größeren Online-Journalen (*Blick*, *Der Standard*, *Kurier*, *taz*, *Watson* (in der deutschen und der schweizerischen Ausgabe), *WinFuture*) und deren Kommentare⁴ (sieben Artikel mit ins-

4 Kommentare werden im Original ohne Anpassung der Rechtschreibung übernommen.

gesamt 496 Kommentaren und 27.327 Token). Die Texte wurden im September 2025 für die Untersuchung ausgewählt, wurden jedoch alle im Zeitraum zwischen dem 11.04.2025 und dem 24.04.2025 veröffentlicht – in zeitlicher Nähe zur Entfernung des Spiels von Steam. Die Artikel bilden stichprobenartig die deutschsprachige Diskurslandschaft ab, da sie von größeren und stark rezipierten Journalen stammen. Diese Texte repräsentieren die Entwicklungen, die auf einen Tabubruch folgen und spiegeln die gesamtgesellschaftliche Auseinandersetzung mit dem Thema wider.

4.1 Moralisierung als toxische Diskurspraxis

Im Diskurs über *No Mercy* finden sich – sowohl in den Artikeln über das Spiel und den Kommentaren dazu als auch im Statement von *Zerat Games* – zahlreiche Moralisierungen, die als toxische Diskurspraxis verstanden werden können. Da eine moralisierende Einordnung auf grundlegenden, als allgemeingültig geltenden Normen der Gesellschaft – und damit auf Tabus – aufbaut, wird eine weitere Auseinandersetzung zu unterbinden versucht und ein produktiver Austausch wird verhindert. Ein sachlicher Austausch, der dazu führt, dass verschiedene Perspektiven aufeinandertreffen und gesellschaftliche Aushandlung ermöglicht, kann durch Moralisierungen nicht stattfinden. Nach einer kurzen theoretischen Einleitung werden Beispiele aus dem Korpus aufgeführt und analysiert.

Als moralisierende Sprachhandlungen gelten »diskursstrategische Verfahren [...], in denen die Beschreibung von Streitfragen und erforderlichen Handlungen mit moralischen Begriffen enggeführt werden.« (Becker et al. 2024: 123) Dabei wird auf moralische Werte verweisende Lexik genutzt, um Forderungen durchzusetzen, wobei keine weitere Begründung erfordert wird (ebd.). Hinter dem Referieren auf moralische Werte steckt somit die »Absicht, eine Aussage als unstrittig gültig darzustellen« (ebd.: 124). Diese Diskurspraxis kann als toxisch betrachtet werden, da sie keine Diskussion mehr ermöglicht und die Unantastbarkeit der Moralexik einen sachlichen Austausch verhindert und damit Gegenstimmen nicht akzeptiert. »Das Muster lautet: X gilt ohne jede weitere diskursive Aushandlung, weil damit das Konzept Y (ein sogenanntes Hochwertkonzept) verbunden ist.« (ebd.: 125) Über diese Hochwertkonzepte herrscht gesellschaftlicher Konsens, wodurch ein Überschreiten oder Ablehnen dieser einen Tabubruch darstellt. Bei der Moralisierung kann auch auf Lexik zurückgegriffen werden, die Einstellungen meint, die allgemeingültig abzulehnen sind (ebd.: 126). Diese »negativen Moralvokabeln« (Felder, Müller

2022: 250) werden als »Delimitationswörter« (ebd.) bezeichnet. »Das sind solche Wörter, die die Grenzen des moralischen Fundaments der Gemeinschaft von außen markieren.« (Ebd.) Auch hier lässt sich eine klare Verbindung zu Tabus feststellen: Tabus markieren Grenzen des gesellschaftlich Akzeptierten und Delimitationswörter verdeutlichen Überschreitungen dieser Grenzen auf sprachlicher Ebene.

In der Erklärung des Entwickler*innen-Teams werden Vorwürfe angebracht, die durch Delimitationslexik geprägt sind. So wird behauptet, es würden Fehlinformationen über das Game verbreitet werden (Zerat Games o. D.). Fehlinformationen sind ein gefährlicher und manipulierender Sachverhalt, der in Demokratien dringend unterbunden werden sollte. Durch diesen Vorwurf wird den Kritiker*innen somit ein abzulehnendes und demokratiegefährdendes Verhalten zugesprochen. Trotz dieses absoluten Arguments folgt eine kurze Ausführung des Themas. Dabei enden die Autor*innen die Textpassage mit dem Appell »Better yet, do some work yourself instead of blindly chasing views.« (Ebd.) Auch hier wird moralisiert, denn blindes Jagen nach Klicks ist ebenfalls ein negatives Verhalten, welches den Kritiker*innen vorgeworfen wird und wodurch ihre Position als falsch und zu verurteilend perspektiviert wird.

Bei den Ausführungen zum Inhalt des Games gehen die Autor*innen des Statements darauf ein, dass Inzest und Vergewaltigung weitverbreitete Fetische seien. Am Ende des Texts wird diesbezüglich resümiert: »[W]e would like you to be a bit more open to human fetishes that don't harm anyone, even though they may seem disgusting to you.« (Ebd.) Auch hier findet eine Moralisierung statt, da an Offenheit gegenüber als unproblematisch empfundenen sexuellen Vorlieben – und damit auch sexueller Freiheit – appelliert wird. Kritiker*innen wird somit implizit Intoleranz vorgeworfen.

Interessant ist auch die Entschuldigung, die am Ende des Textes formuliert wird: »If after reading the above, you still believe that such a game should not have been created, then we sincerely apologize to you.« Eine Entschuldigung findet hier nur eingeschränkt an diejenigen statt, die von den Ausführungen der Erklärung nicht überzeugt werden konnten. Diesen Personen wird damit indirekt Unvernünftigkeit unterstellt, womit ebenfalls sprachlich und argumentativ eine moralische Überlegenheit aufgebaut wird.

In den Artikeln über *No Mercy* findet Moralisierung insbesondere in Form von negativen Bewertungen des Spiels bzw. des dahinterstehenden Entwickler*innen-Teams statt. So referiert *Kurier* u.a. darauf, dass die britische Innenministerin das Game als »widerwärtig« (Unger 2025) bezeichnet hat. Das Deli-

mitationswort wird einerseits als Pathologisierung des Spiel(inhalt)s genutzt, andererseits wird durch den Autoritätstopos, dass diese Bewertung von einer wichtigen Politikerin vorgenommen wurde, die Moralisierung unterstützt und verstärkt. Auch Schlagwörter wie »Inzestspiel« (Vela 2025) konzeptualisieren das Game als inakzeptabel und damit in hohem Maße tabuüberschreitend. In *Der Standard* wird dieser Tabubruch explizit als solcher benannt, wenn es heißt, dass das Spiel »die Richtlinien der Games-Plattform Steam zwar nicht bricht, aber gleichzeitig mehrere moralische Grenzen überschreitet.« (Amon 2025) Durch den direkten Verweis auf moralische Grenzen finden außerdem auch hier eine Moralisierung statt.

Des Weiteren wird das Spiel bzw. dessen Inhalt als frauenfeindlich bezeichnet. Dies findet sich u.a. im Titel der *taz* »Sexistische Videospiele. Next Level Frauenhass« (Seng 2025). *No Mercy* wird somit sofort als misogynen Game klassifiziert. Im Artikel wird das Spiel als »Zeichen einer misogynen Gaming-Kultur« (ebd.) bezeichnet, womit eine Moralisierung und Stereotypisierung einhergeht. Die Gaming-Community und ihre Mitglieder werden durch die Abwertung der gesamten Gruppe als frauenverachtend betrachtet (s. dazu auch Reinhold, Prinz in diesem Band). Moralisierung kann dementsprechend stereotypisierend und damit toxisch wirken, dabei jedoch auch durch die Konstruktion einer positiven In-Group – in diesem Fall Nicht-Gamer*innen – sowie einer negativen Out-Group – hier Gamer*innen – zur Stiftung von Identität beitragen.

Ein weiterer Aspekt, der in den kritischen Beiträgen zur Moralisierung herangezogen wird, ist Jugendschutz. Einige Artikel erwähnen, dass das Spiel für Minderjährige zugänglich gewesen sei. So wird bereits im Titel des Artikels von *Blick* betont, *No Mercy* sei für »13-Jährige« (Ruggli 2025) verfügbar gewesen und auch die deutsche Website von *Watson* verweist darauf, dass auch Teenager auf das Spiel zugreifen konnten (Köppel 2025). Insbesondere durch die Verbindung mit negativ konnotierten Schlagwörtern im Text wie »Inzest-Porno« (Ruggli 2025), »Vergewaltigung« (ebd.) und »Inzest-Vergewaltigungsspiel« (Köppel 2025) wird stark an den Schutz Minderjähriger appelliert.

Moralisierungen in den Kommentaren unter Artikeln nutzen primär pathologisierende und stereotypisierende Lexik zur Moralisierung. Dabei werden Tabus durch die starke Abwertung von tabubrechendem Verhalten verteidigt, wodurch die Bildung einer Gruppenidentität mit geteilten Werten gefördert wird. Gleichzeitig findet eine toxische Kommunikation statt, da Mitglieder der Outgroup als krankhaft angesehen und damit diffamiert werden.

Dabei werden sowohl das Spiel als auch die Spieler*innen und die Entwickler*innen pathologisiert. Beispielsweise wird von *No Mercy* als zugehörig zu »kranken Spielen« (K-A, Watson CH 2025) gesprochen. Auch die Zielgruppe bzw. die Spieler*innen werden negativ bewertet und als »kranke Individuen« (K-B, Watson CH 2025) oder »Alt-Right Spinner« (K-C, WinFuture 2025) diffamiert. Dabei finden direkte Bezüge zur im Diskurs besprochenen Misogynie des Spiels bzw. der Gaming-Kultur und deren Toxizität statt, so wird bspw. argumentiert, dass die »schmuddeligen Tiefen der zutiefst toxischen Mansphere [...] junge Männer zu unerträglichen Vollpfosten zu machen sucht.« (K-D, Watson CH 2025) Hier gilt anzumerken, dass auch die Zuschreibung als toxisch normativ und damit moralisierend ist.

Demus et al. (2023: 273) halten fest, dass toxische Kommunikationselemente in Originalbeiträgen – in diesem Fall den Artikeln – Hass bzw. Toxizität in den Kommentaren begünstigen und zwar sowohl gegen eine allgemeine Gruppe als auch gegen Individuen. Dies lässt sich bspw. in Kommentaren unter dem Artikel von der schweizerischen Website von *Watson* beobachten, denn eine Person gibt an, das Spiel gekauft zu haben (K-E, Watson CH 2025), woraufhin eine andere Person diese als »[k]rank« (K-F, Watson CH 2025) bezeichnet und damit direkt angreift.

Auch die Entwickler*innen werden in den Kommentaren pathologisiert. Ihnen wird jedoch – anders als den Spieler*innen – sogar direkt zur Therapie geraten: »[d]ie Produzenten sollten mal in sich gehen und alle eine Therapie machen.« (K-A2, Watson CH 2025), »das Team sollte dringend zum Psychiater« (K-G, Der Standard 2025). Eine kommentierende Person hofft, »die Autoren dieses »Games« [haben sich] hoffentlich einer Psychoanalyse unterzogen.« (K-H, Watson CH 2025). Die Pathologisierung des Entwickler*innen-Teams geht auch mit Forderungen einher, »den Entwicklern Probleme zu bereiten, damit sich so ein Schwachsinn nicht so schnell wiederholt« (K-I, Der Standard 2025). Dieser wenig konkrete Appell des Problembereitens kann als Gewaltaufruf gelesen werden und ist eine tabuisierende Formulierung.

Neben den klar gerichteten pathologisierenden Abwertungen finden sich auch Kommentare, bei denen unklar ist, wer oder was devaluiert wird, so etwa in einigen Kommentaren unter dem Artikel in *Blick*, die lediglich Adjektive wie »[a]bartig« (K-J, Blick 2025) oder »[e]kelhaft« (K-K, Blick 2025; K-L, Blick 2025) enthalten. Durch diese starken Abwertungen findet eine legitime Verurteilung von als unmoralischen empfundenen Praktiken statt, jedoch wird ein sinnvoller Austausch über diese Themen nicht möglich.

Es existieren jedoch auch Kommentare, die die Moralisierung als (toxische) Diskurspraxis thematisieren und Kritik an übermäßiger Moralisierung äußern. Dies erfolgt bspw. über die Bezeichnung der moralisierenden Kommentator*innen und Autor*innen als »Moralisten« (K-M, Watson CH 2025; K-N, Blick 2025) oder »Moralapostel« (K-O, Der Standard 2025), wodurch ihnen als übertrieben wahrgenommene Moralvorstellungen unterstellt werden. Das Einhalten von Tabus durch die Ablehnung von Tabubrüchen – in diesem Fall der Darstellung von Vergewaltigung und Inzest – wird von diesen Kommentierenden negativ bewertet, wohinter die Auffassung liegen kann, dass Tabus als veraltetes Konzept gelten (Mildenberger, Schröder 2012).

4.2 Gamer's Dilemma als Whataboutismus

Das »Gamer's Dilemma« geht auf Luck (2009) zurück und meint die moralische Ungleichbehandlung von virtuellem Mord im Vergleich zu virtueller Übergriffigkeit an Kindern. »Whilst it is certainly true that committing virtual murder is, for the most part, socially acceptable, committing virtual paedophilia remains taboo.« (ebd.: 32) Im Fall von *No Mercy* kann diese philosophische Idee auch auf das Begehen von Vergewaltigungen und Inzest in Videospielen übertragen werden, denn in den Kommentaren unter den untersuchten Texten findet sich der Vergleich zum Morden in Games vermehrt wieder. Im Diskurs über *No Mercy* stellt die Frage danach, wieso virtueller Mord und Totschlag als weniger schlimm erachtet werden als virtuelle (inestuöse) Vergewaltigung Whataboutismus⁵ dar, denn es wird vom eigentlichen Thema abgelenkt und auf ein als verwandt behandeltes Thema eingegangen. Damit wird die Diskussion verschoben und durch den Verweis auf ein anderes Thema wird nicht mehr oder weniger direkt über den eigentlichen Diskursinhalt gesprochen. Dieses Phänomen kann auch als Tabuisierungsstrategie betrachtet werden, wenn statt eines tabubehafteten Themas ein weniger stark tabuisierter Inhalt diskutiert wird. Neben dem Abkommen von der eigentlichen Diskussion sorgt diese Praxis ebenfalls dafür, dass Vorwürfe darüber, was als schlimm gilt

5 Whataboutismus stellt eine »Störstrategie[]« (Amadeu Antonio Stiftung o. D.) dar, die insbesondere in Kommentarbereichen häufig genutzt wird. Dabei handelt es sich um Gegenaussagen, »die nichts mit dem eigentlichen Thema zu tun haben« (ebd.) und somit zur Ablenkung dienen. Der ModSupport der Amadeu Antonio Stiftung hält fest: »Wer mit solchen Aussagen kommentiert, hat kein Interesse an einem konstruktiven Austausch.« (ebd.) Whataboutismus ist somit grundlegend toxisch für Diskurse.

bzw. welche virtuellen Taten als schlimmer gelten, getätigt werden. Dies kann in Verurteilungen – gegen allgemeine Vorstellungen oder gegen spezifische kommentierende Personen – realisiert werden.

Eine Form, mit dem Gamer's Dilemma umzugehen, ist der »Ludic amoralist«, der alle virtuellen Verbrechen damit entschuldigt, dass es sich nur um ein Spiel handelt (Ulbricht 2023⁶). Dieses Argument wird von verschiedenen Akteur*innen im Diskurs angeführt, sowohl vom Entwickler*innen-Team als auch von Kommentierenden unter Artikeln. In der Erklärung des Studios heißt es am Ende: »This is still just a game, and although many people are trying to make it into something more, it remains and will continue to be a game.« (Zerat Games o. D.) Nach der gesamten Ausführung fasst *Zerat Games* die Relevanz des kritischen Diskurses um *No Mercy* damit zusammen, dass die Sorgen der Kritiker*innen aufgrund der Fiktionalität irrelevant seien. Ähnlich wie bei der Moralisierung dient auch der Verweis darauf, dass es sich lediglich um ein Spiel handle, dazu, ein unabstreitbares Argument darzulegen, welches eine weitere Auseinandersetzung nicht notwendig macht und diese zu unterbinden versucht. In den Kommentaren zu Artikeln über das Game findet sich diese Argumentation ebenfalls: »Nur virtuell, nur ein Spiel. Aufregung um wenig.« (K-P, Der Standard 2025)

Auch in den Kommentaren unter den Artikeln von Online-Magazinen, die *No Mercy* thematisieren, existieren Diskussionen darüber, ob virtuelle Gewalt an Frauen schlimmer sei als virtueller Mord und/oder virtuelle Folter. Insbesondere der Vergleich zu der Spielreihe *Grand Theft Auto (GTA)* findet sich in einigen Kommentaren, wie etwa unter dem Artikel von *Der Standard*: »Kriegsspiele, Spiele wo man nach Lust und Liebe mordet (a la GTA), Genozid an Aliens begeht... Die werden dann natürlich auch alle verboten?« (K-Q, Der Standard 2025). In der daraus resultierenden Diskussion mit anderen User*innen wird zwar deutlich, dass nicht alle Diskussionsteilnehmer*innen dieser Ansicht sind, diese Meinung jedoch auch in angreifender und diffamierender Art realisiert wird: »Man kann alles relativieren, nur die eigene Dummheit nicht.« (K-R, Der Standard 2025) Sehr konkret auf das Gamer's Dilemma geht der Kommentar »warum ist töten in spielen okay, aber vergewaltigung nicht?« (K-S, Der Standard 2025) ein. In den Antwortkommentaren dazu finden sich Argumente, die häufig angeführt werden, um das Gamer's Dilemma zu legitimieren, wie etwa, dass Gewalt in Games häufig einen spezifischen Nutzen hat und u. a. dazu dient, Feinde aufzuhalten und für das Gute

6 Vielen Dank an Paul Reszke für die Lektüreempfehlung.

zu kämpfen (Ulbricht 2023: 8). Weitere Whataboutismen im Diskurs finden sich u.a. im Vergleich zu Kinofilmen (K-T, Blick 2025), in denen die Darstellung von Vergewaltigungen in Ordnung zu sein scheinen. Durch den Bezug zu anderen Gewaltformen oder Medien wird der Inhalt von *No Mercy* nicht nur als legitim gerechtfertigt, sondern der Whataboutismus stellt überdies eine Tabuisierungsform dar, da eine Ablenkung vom eigentlichen Thema des Diskurses stattfindet.

4.3 Vorwurf der Aufmerksamkeitsgenerierung

In den Artikeln und Kommentaren, die sich mit *No Mercy* auseinandersetzen, findet sich vermehrt der Vorwurf, es handle sich bei dem Game bzw. dessen Inhalt um eine inszenierte Provokation, die zur Bekanntheit des Spiels beitragen solle. Damit wird unterstellt, die Tabubrüche innerhalb des Games würden lediglich dazu dienen, eine öffentliche Diskussion auszulösen. Im Artikel von *WinFuture* wird bspw. erwähnt, dass »Kritiker [...] den Vorwurf [erheben], dass die Entwickler die Kontroverse gezielt geschaffen haben, um Aufmerksamkeit zu erhalten.« (Woll 2025) Auch in einem Kommentar unter dem Beitrag von der schweizerischen Website von *Watson* geht ein*e User*in davon aus, dass »der Aufschrei [...] einkalkuliert« sei (K-U, Watson CH 2025). Ein weiterer Kommentar vermutet ebenfalls einen inszenierten Tabubruch: »Kann gut sein dass man auf den Skandal gehofft hat.« (K-V, Watson 2025) Interessant ist dabei, dass der Tabubruch nicht als solcher benannt wird, sondern bedeutungsverwandte Wörter verwendet werden.

Darüber hinaus wird außerdem – u.a. in der Erklärung des Studios – kritisiert, dass die mediale Berichterstattung über das Spiel zu dessen Bekanntheit beitragen würde. Somit werden aus unterschiedlichen Lagern böse Absichten der Aufmerksamkeitsgenerierung unterstellt. In der Erklärung der Entwickler*innen wird die Bekanntmachung des Spiels durch dessen öffentliche Diskussion wie folgt aufgegriffen: »Please consider – would anyone who wasn't looking for such content hear about this game if it weren't for hundreds of articles, petitions, and statements from content creators?« (Zerat Games o. D.)

Auch in Artikeln über das Game wird teilweise kritisch reflektiert, dass »das Spiel bis zu seiner Gegenkampagne vollkommen unbekannt [war]« (Seng 2025) und »hitzige Debatten und eine erhöhte Aufmerksamkeit, [...] dem Entwickler letztlich mehr nützt, als schadet« (Woll 2025). In den Kommentaren unter den Artikeln über das Spiel finden sich ebenfalls kritische Äußerungen zur Wirkung der medialen Aufmerksamkeit. Hier werden jedoch häufig

die jeweiligen Journale direkt angegriffen, wie »Warum unterstützt Watson diese perfide Werbekampagne? Nichts anderes ist es doch! Niemand hätte dieses ›Spiel‹ auf dem Schirm gehabt, und jetzt machen praktisch alle Medien Werbung dafür.« (K-W, Watson CH 2025)

Aus tabutheoretischer Perspektive lässt sich diese Kommunikationspraxis folgendermaßen einordnen: Ein Tabubruch ist ein Angriff auf ein bestehendes (Gesellschafts-)System, was auch als bewusste Provokation genutzt werden kann.

Wenn Tabus die Ordnung bewahren, so sind Tabubrüche Angriffe auf selbige Ordnung. Tabus sind deshalb auch gleichzeitig strategisches Mittel des gesellschaftlichen Systems, sich selbst zu schützen, woraus folgt, dass wiederum Tabubrüche strategisch eingesetzt werden können, um ein System in Aufruhr zu versetzen und zu hinterfragen. (Oettl 2019: 46)

Sollen ein System und dessen Normen geschützt werden, muss ein Tabubruch sozial geahndet werden, was in diesem Fall durch einen öffentlichen Diskurs mit Kritik zum Game stattfand, wodurch wiederum dessen Bekanntheitsgrad gesteigert wurde.

5. Fazit

Abschließend lässt sich festhalten, dass im Diskurs über *No Mercy* Tabu und Toxizität in den Kommunikationspraktiken eng verbunden sind. Dabei werden toxische Diskurspraktiken sowohl von den Kritiker*innen des Games als auch von dessen Verteidiger*innen genutzt. Zentrale Formen (toxischer) Diskurspraktiken sind hierbei Moralisierung, Whataboutismus und Vorwürfe der Aufmerksamkeitsgenerierung. Dass diese Aspekte teilweise auch in kombinierter Form vorkommen, verdeutlicht der folgende Kommentar:

Es ist eine Sache Pornoseiten dazu zu bringen echte Vergewaltigungsinhalte zu entfernen, warum man aber eine Fiktion bekämpft, erschließt sich wohl nur der Moral-Spezialeinheit. Nun klopf man sich auf die Schulter, dass man einem dümmlichen Spiel, das wohl sonst völlig unter dem Radar geflogen wäre, kostenlose PR verpasst hat. (K-X, Der Standard 2025)

Bei den angewendeten Kommunikationspraktiken fällt auf, dass sie größtenteils Akteur*innen-unabhängig eingesetzt werden. So nutzen sowohl die Entwickler*innen des Games in ihrer Erklärung als auch Kritiker*innen in Artikeln und Kommentaren diverse moralisierende Praktiken. Auch der Vorwurf der Aufmerksamkeitsgenerierung wird von Befürworter*innen ebenso wie von Kritiker*innen eingesetzt, jedoch finden sich unterschiedliche Argumentationsstrategien: Während Kritiker*innen davon ausgehen, dass Spiel nütze den tabubrechenden Inhalt zur Erzeugung von öffentlichem Interesse, argumentieren Verteidiger*innen des Games damit, dass das Spiel erst durch die mediale Berichterstattung Aufmerksamkeit bekommen hätte. Beide Gruppen perspektivieren denselben Kritikpunkt der Aufmerksamkeitsgenerierung aus ihren jeweiligen Sichtweisen, sodass im Zentrum lediglich das Zuweisen von Fehlverhalten und Schuld statt einer produktiven Auseinandersetzung mit dem Thema steht. Whataboutismus als vom Thema ablenkendes und damit auch tabuisierendes kommunikatives Verhalten hingegen wird lediglich zur Legitimierung des Spiels genutzt. Gemeinsam haben diese Strategien, dass sie versuchen, eine weitere Auseinandersetzung zu vermeiden und gegnerische Akteur*innen(-Gruppen) abzuwerten.

Tabubrüche sorgen für soziale Sanktionierung, was im Bereich des Gamings häufig über kritische Diskurse erfolgt. Diese sind prädestiniert für toxische Kommunikationspraktiken, da negative Bewertungen getätigt werden, die meist auf Gegenmeinungen treffen. Ein weiterer Aspekt, der Toxizität in Diskursen über Games weit verbreitet macht, ist die Stereotypisierung, die im Hinblick auf Videospiele und -spieler*innen weiterhin herrscht und als Grundlage toxischer Auseinandersetzungen dienen kann.

Außerdem kann ein Tabubruch dafür sorgen, dass im Diskurs über diesen ebenfalls weitere kommunikative Tabubrüche begangen werden. Auf einen Tabubruch muss die Person oder Gruppe, die diesen zu verantworten hat, mit einer Erklärung oder Entschuldigung reagieren, um die soziale Sanktion zu regulieren. *Zerat Games* reagiert diesen Anforderungen entsprechend. Jedoch wird auch diese Erklärung mit (eingeschränkter) Entschuldigung kritisiert, wodurch die soziale Sanktion nicht aufgehoben wird.

Durch einen Tabubruch werden negative Emotionen ausgelöst, welche ebenfalls häufig bei Personen zu finden sind, die sich toxisch äußern (Demus et al. 2023: 265). Außerdem kann auf Tabubrüche mit Angriffsverhalten reagiert werden (Birk, Kaunzner 2009: 402). Somit beeinflussen sich die beiden Phänomene, sodass es zu einer toxischen Kommunikationsverstrickung kommt, die durch einen singulären Tabubruch ausgelöst wurde, welcher

weitere kommunikative Tabubrüche zur Folge hat. Tabubrüche und Toxizität sind auf kommunikativer Ebene eng verknüpfte Phänomene, die einander bedingen.

Literaturverzeichnis

- Adams, Phillipa R.; Sommers, Maria; Scholl, Ben; Lusoli, Alberto; Chun, Wendy Hui Kyong (2025): The revenge of ›tribal‹ toxic masculinity: hate raids on Twitch. In: *Journal of Visual Culture*. URL: <https://doi.org/10.1177/14704129251327916>.
- Almerekhi, Hind; Kwak, Haewoon; Salminen, Joni; Jansen, Bernard J. (2019): Detecting Toxicity Triggers in Online Discussions. In: *Proceedings of the 30th ACM Conference on Hypertext and Social Media (HT'19)*. Hof. S. 291–292. URL: <https://doi.org/10.1145/3342220.3344933>.
- Amadeu Antonio Stiftung (o. D.): ModSupport – Der Guide für deine Kommentarspalte. Störstrategien in der Kommentarspalte begegnen: Whataboutism. In: Antonio Amadeu Stiftung. URL: <https://www.amadeu-antonio-stiftung.de/modsupport/whataboutism/> [Zugriff 01.12.25].
- Amon, Alexander (2025): NO MERCY. »Inzest«-Spiel sorgt auf großer Gaming-Plattform für Aufregung. In: *Der Standard*. URL: <https://www.derstandard.de/story/3000000265584/inzest-spiel-sorgt-auf-grosser-gaming-plattform-fuer-aufregung> [Zugriff 05.10.25].
- Becker, Maria; Felder, Ekkehard; Müller, Marcus (2024): Moralisierung als sprachliche Praxis. In: *Moral und Moralisierung: Neue Zugänge*. Hg. v. Ekkehard Felder; Friederike Nüssel u. Jale Tosun. Berlin, Boston: De Gruyter. S. 123–152.
- Benthien, Claudia; Gutjahr, Ortrud (2008): Interkulturalität und Gender-Spezifika von Tabus. Zur Einleitung. In: *Tabu. Interkulturalität und Gender*. Hg. v. Claudia Benthien u. Ortrud Gutjahr. München: Fink. S. 7–16.
- Birk, Andrea; Kaunzner, Ulrike A. (2009): Tabu und Identität. Wie man das vermittelt, worüber die anderen schweigen. In: *Wie kann man vom ›Deutschen‹ leben? Zur Praxisrelevanz der interkulturellen Germanistik*. Hg. v. Ernest W. B. Hess-Lüttich. Berlin u.a.: Lang. S. 397–416.
- Braun, Michael (2007): Vorwort. In: *Tabu und Tabubruch in Literatur und Film*. Hg. v. Michael Braun. Würzburg: Königshausen & Neumann. S. 7–10.
- Demus, Christoph; Schütz, Mina; Probst, Nadine; Pitz, Jonas; Siegel, Melanie; Labudde, Dirk (2023): Hass im Netz – Aggressivität und Toxizität von Hass-

- kommentaren und Postings, Detektion und Analyse. In: Handbuch Cyberkriminologie. Hg. v. Thomas-Gabriel Rüdiger u. Petra Saskia Bayerl. Wiesbaden: Springer VS. S. 261–292.
- Dreier, Lukas; Priker, Johanna (2023): Toxicity in Twitch Live Stream Chats: Towards Understanding the Impact of Gender, Size of Community and Game Genre. In: 2023 IEEE Conference on Games (CoG). URL: <http://dx.doi.org/10.1109/CoG57401.2023.10333159>.
- Eggert, Hartmut (2002): Säkulare Tabus und die Probleme ihrer Darstellung. Thesen zur Eröffnung der Diskussion. In: Tabu und Tabubruch. Literarische und sprachliche Strategie im 20. Jahrhundert. Ein deutsch-polnisches Symposium. Hg. v. Hartmut Eggert u. Janusz Golec. Stuttgart, Weimar: Metzler. S. 15–24.
- Felder, Ekkehard; Müller, Marcus (2022): Diskurs korpuspragmatisch. Annotation, Kollaboration, Deutung am Beispiel von Praktiken des Moralisieren. In: Sprache in Politik und Gesellschaft. Perspektiven und Zugänge. Hg. v. Heidrun Kämper u. Albrecht Plewnia. Berlin, Boston: De Gruyter. S. 241–261.
- Frietsch, Ute (2008): Der Wille zum Tabu als Wille zum Wissen. In: Geschlecht als Tabu. Orte, Dynamiken und Funktionen der De/Thematisierung von Geschlecht. Hg. v. Ute Frietsch, Konstanze Hanitzsch, Jennifer John u. Beatrice Michaelis. Bielefeld: transcript. S. 9–16.
- Fox, Jesse; Tang, Wai Yen (2017): Women's experiences with general and sexual harassment in online video games: Rumination, organizational responsiveness, withdrawal, and coping strategies. In: *New Media & Society*. 19, 8. S. 1290–1307. URL: <https://doi.org/10.1177/1461444816635778>.
- Gutjahr, Ortrud (2008): Tabus als Grundbedingung von Kultur. Sigmund Freuds Tabu und Totem und die Wende in der Tabuforschung. In: Tabu. Interkulturalität und Gender. Hg. v. Claudia Benthien u. Ortrud Gutjahr. München: Fink. S. 19–50.
- Hoffmeister, Toke (2023a): Toxizität/das Toxische. In: Diskursmonitor. Glossar zur strategischen Kommunikation in öffentlichen Diskursen. Hg. von der Forschungsgruppe Diskursmonitor und Diskursintervention. URL: <https://diskursmonitor.de/glossar/toxizitaet-das-toxische/> [Zugriff 15.09.25].
- Hoffmeister, Toke (2023b): Von toxischen Substanzen und toxischer Männlichkeit. Eine korpuspragmatische Studie zur Erschließung eines Gegenwartssphänomens. In: *Zeitschrift für Germanistische Linguistik*. 51, 3. S. 424–458.

- Köppel, Melanie (2025): Vergewaltigungs-Spiel »No Mercy«: Wie konnte so etwas jemals online sein? In: Watson. URL: <https://www.watson.de/leben/digital/908194219-no-mercy-auf-steam-spiel-bedient-vergewaltigungsfantasien-im-netz> [Zugriff 05.10.25].
- Kuck, Kristin (2021): Tabugrenzen im öffentlichen Raum. In: Kleine Texte. Hg. v. Steffen Pappert u. Kersten S. Roth. Berlin: Lang. S. 227–254.
- Kuhn, Fritz (1987): Tabus. In: Sprache und Literatur in Wissenschaft und Unterricht. 18, 2. S. 19–35.
- Lautenschläger, Sina (2021): Willst du gelten, mach dich selten: Tabu und Schweigen in interpersonaler Kommunikation. In: aptum. Zeitschrift für Sprachkritik und Sprachkultur. 17, 2. S. 212–229.
- Luck, Morgan (2009): The gamer's dilemma: An analysis of the arguments for the moral distinction between virtual murder and virtual paedophilia. In: Ethics and Information Technology, 11, 1. S. 31–36. URL: <https://doi.org/10.1007/s10676-008-9168-4>.
- Mildenberger, Florian; Schröder, Hartmut (2012): Tabu, Tabuvorwurf und Tabubruch im politischen Diskurs. In: bpb. URL: <https://www.bpb.de/shop/zeitschriften/apuz/75862/tabu-tabuvorwurf-und-tabubruch-im-politischen-diskurs/> [Zugriff 08.03.24].
- Oertl, Barbara (2019): Existentielle Grenzerfahrungen. Tabubruch als Strategie in der zeitgenössischen Kunst. Bielefeld: transcript.
- Rada, Roberta V. (2001): Tabus und Euphemismen in der deutschen Gegenwartssprache. Mit besonderer Berücksichtigung der Eigenschaften von Euphemismen. Budapest: Akadémiai Kiadó.
- Ruggli, Natascha (2025): Entwickler äussert sich zu der Furore 13-Jährige konnten bis vor kurzem dieses Game kaufen – Inzest-Porno war auf Steam verfügbar. In: Blick. URL: <https://www.blick.ch/digital/entwickler-aeussert-sich-zu-der-furore-13-jaehrige-konnten-bis-vor-kurzem-dieses-game-kaufen-inzest-porno-war-auf-steam-verfuegbar-id20773634.html> [Zugriff 05.10.25].
- Schröder, Hartmut (2008): Zur Kulturspezifität von Tabus. Tabus und Euphemismen in interkulturellen Kontaktsituationen. In: Tabu. Interkulturalität und Gender. Hg. v. Claudia Benthien u. Ortrud Gutjahr. München: Fink. S. 51–70.
- Seng, Martin (2025): Sexistische Videospiele. Next Level Frauenhass. In: taz. URL: <https://taz.de/Sexistische-Videospiele/!6080685/> [Zugriff 05.10.25].

- Ulbricht, Samuel (2023). A Kantian response to the Gamer's Dilemma. In: *Ethics and Information Technology* 25, 39. URL: <https://doi.org/10.1007/s10676-023-09710-0>.
- Unger, Sophie (2025): Möglichst viele Vergewaltigungen: Online-Spiel »No Mercy« schockt. In: *Kurier*. URL: <https://kurier.at/trend-hub/no-mercy-game/403032602> [Zugriff 05.10.25].
- Vela, Sabeth (2025): »Wir haben nicht vor, gegen die ganze Welt zu kämpfen« – Porno-Game »No Mercy« gesperrt. Das neue Inzestspiel sorgte weltweit für Empörung – nun hat Steam das Game gesperrt. In: *Watson*. URL: <https://www.watson.ch/leben/digital/924557998-steam-sperret-porn-ogame-no-mercy> [Zugriff 05.10.25].
- Woll, John (2025): Nach großer Empörung: Verstörendes »Sex«-Spiel verschwindet von Steam. In: *WinFuture*. URL: <https://winfuture.de/news,150277.html> [Zugriff 05.10.25].
- Zerat Games (o. D.): Important News about No Mercy. In: *Steam*. URL: <https://steamcommunity.com/games/3299570/announcements/detail/588390482275467288> [Zugriff 06.10.25].
- Zerat Games Website (o. D.). URL: <https://zeratgames.com> [Zugriff 05.10.25].

"you woke cultists are a cancer to everything"

Toxicity through Personal References in the Steam Group 'Woke Content Detector'

Christina Liemann

Abstract This study examines how personal references in the form of pronouns and noun-based nominal phrases can facilitate the adoption and construction of antagonistic social positions that damage discursive environments and render them toxic, as they make productive exchange impossible. To this end, the usage of personal references in the Steam Group 'Woke Content Detector', where both Detectors (who identify the supposedly 'woke' games) and Counter Speakers (who oppose the group's aim) engage in mutual discourse, is examined. While both groups construct antagonistic social positions through language, there are significant differences in the pronouns and other nominal phrases being used respectively. The Detectors use more generalizing, dehumanizing and hostile personal references, while the Counter Speakers primarily employ the second person singular, which directly addresses the user they're talking to. In contrast, the Detectors tend to talk about the Counter Speakers, rather than speaking to them, indicated by the higher use of the third person plural and noun-based nominal phrases. It is unclear if these differences can be attributed to the ideologies being represented respectively, or if this is a feature of online discourse and/or discourse in echo chambers in general. To determine this, further research and comparison with more progressive discussion groups is needed.

Keywords Discourse Analysis, Discourse Health, Positioning, Pronouns, Metaphors, Culture War, Social Media

Abstract Diese Studie untersucht, wie Personenreferenzen in Form von Pronomen und Substantiv-basierten Nominalphrasen das Einnehmen und

die Konstruktion antagonistischer sozialer Positionen begünstigen können, welche die Diskursumgebung schädigen und toxisch machen, da sie einen produktiven Austausch verunmöglichen. Zu diesem Zweck wird der Gebrauch von Personenreferenzen in der Steam-Gruppe ‚Woke Content Detector‘ untersucht, in der sowohl Detectors (die die vermeintlich ‚woken‘ Spiele identifizieren) als auch Counter Speaker (die sich gegen das Gruppenziel stellen) am gemeinsamen Diskurs teilnehmen. Während beide Gruppen antagonistische soziale Positionen durch Sprache konstruieren, gibt es signifikante Unterschiede in Bezug auf die jeweils verwendeten Pronomen und anderen Nominalphrasen. Die Detectors benutzen mehr verallgemeinernde, entmenslichende und feindselige Personenreferenzen, während die Counter Speaker vorwiegend die zweite Person Singular verwenden, die die User*innen, mit denen sie sprechen, direkt adressiert. Im Gegensatz dazu neigen die Detectors dazu, über die Gegenredner*innen zu sprechen, anstatt sie direkt anzusprechen, was durch die häufigere Verwendung der dritten Person Plural und Substantiv-basierten Nominalphrasen deutlich wird. Es ist unklar, ob diese Unterschiede auf die jeweils vertretenen Ideologien zurückzuführen sind oder ob dies ein Merkmal von Online-Diskursen und/oder Diskursen in Echokammern im Allgemeinen ist. Um dies zu klären, sind weitere Studien und der Vergleich mit progressiveren Diskussionsgruppen erforderlich.

Schlagwörter Diskursanalyse, Diskurs-Gesundheit, Positionierung, Pronomen, Metaphern, Kulturkrieg, Soziale Medien

1. Introduction¹

The declining quality of public discourse, especially in online spaces, has been a subject of criticism in recent times (e.g. Neuberger 2022). Indeed, it is cause for concern in any democratic society when constructive dialogue becomes less common and discourse environments grow increasingly toxic instead. It can be argued that in many ways, the internet and social media have contributed to a democratization of discourse, as usual selection mechanisms of participation have been negated in these spaces (Habermas 2006: 554). On the other hand, social media also seems to have a negative impact on public discourse, as “this easy access and broader scale of communication in addition to medium-

1 Quote in title: [DYE-11].

related aspects such as anonymity seems to have contributed to antisocial communication behaviour (Chui 2014)" (Gevers et al. 2022: 33). This perceived trend towards toxicity in online discourse is set against the backdrop of what is often referred to as the 'culture wars', which are also largely fought online. These culture wars are closely tied to popular culture, as Proctor and Kies (2018: 127) note:

Since at least 2014, the terrain of popular culture has paid host to an uptick in political conflict and quarrel, becoming the battleground on which the new culture wars are being fought.

This prominently includes the domain of video games, as the '#Gamergate' movement, which was characterized by "viciously sexual and sexist attacks on women in and around gaming communities" (Braithwaite 2016) has exemplified (see also Reinhold, Prinz in this volume). Massanari (2017) observes the emergence of what she calls "toxic technocultures", which "are enabled by and propagated through sociotechnical networks such as Reddit, 4chan, Twitter, and online gaming" (ibid.: 333) and "often rel[y] [on] an Othering of those perceived as outside the culture" (ibid.). This study examines precisely the described phenomenon by analyzing the discourse within the Steam group 'Woke Content Detector', which aims to identify supposedly 'woke' games. At the same time, counter speakers are engaging in this group, disrupting internal group discourse and the surrounding echo chamber. As Sponholz (2024) notes, counter speech does not necessarily contribute to fighting toxicity in social media environments, instead it often increases the incivility, transforming the discourse into a metaphorical "gladiatorial arena" (ibid., translation C.L.). The same can be observed in the corpus of this study: Members of both groups position themselves in and thereby construct an antagonistic dichotomy between supposedly 'woke' and 'anti-woke' groups, which in turn damages the discursive environment and turns it toxic. The presence of the respective other group is thus reduced from an opportunity for open dialogue to a mere means of reinforcing one's own group boundaries. This is facilitated, in part, through personal references – i.e. the pronouns and noun-based nominal phrases used to refer to the other group, respectively. This study will show that, even when obviously toxic elements like hate speech or insults are comparatively rare, a discourse environment can become unproductive and toxic through personal references, which construct socially antagonistic positions and fortify in-group out-group boundaries in a relatively subtle manner.

2. Woke Content Detector, the term 'woke' and Counter Speakers

Woke Content Detector is a curator on Steam, which is the largest digital distribution service and storefront for PC games. According to Steam, curators are “Individuals or organizations that make recommendations to help others discover interesting games in the Steam catalog” (Steam: Curators) and “[a]ny Steam Community group can become a curator” (ibid.). Steam users can follow these curators to regularly get their recommendations. At the present time (from here on out meaning the 14th of April 2026) the Woke Content Detector Curators page has 13,168 followers. Woke Content Detector self-describes as follows: “We identify whether a game has Woke content and if so, to what degree.” (Woke Content Detector: About) Based on the perceived amount of ‘woke’ content, the curator sorts games into the categories ‘recommended’, ‘not recommended’ and ‘informative’ (ibid.). Currently, of the 1,680 reviewed games, 940 are listed as not recommended, while 398 and 342 games are rated as recommended and informative respectively (ibid.). The curator also posts short reviews of the games, which further describe the perceived amount of ‘wokeness’. As an example, the short review for the game *Kingdom Come Deliverance II* (2025), which received the not recommended rating, reads: “Contains overtly pro-DEI, overtly pro-LGBTQ+, and overtly anti-western society messaging.” (Ibid.) The site also links to the associated group page, which hosts the full-length reviews of the rated games as well as discussions about these ratings and other associated topics. The Woke Content Detector Group was established on the 16th of March 2024 and currently has 4,994 members. Discussions are organized into separate threads, each started by a post and open to comments from users. Users can also reply to other users’ posts.

In the “Methodology” discussion the group owner defines woke content as “any images, messages, characters, storytelling, dialogue, music, or game mechanics that include themes associated with the left-side of the political aisle in contemporary western politics.” [M-OP]² Such themes supposedly include:

-
- 2 References to the corpus are given as sigla and are marked in square brackets. The sigla consist of the first letters of the first three words of the thread title, followed by the comment number. If a thread title contains fewer than three words, only one or two letters are used correspondingly. The original post that starts the thread is not assigned a number by Steam and is therefore simply indicated as ‘OP’ (= original post). The posts were quoted in their original form, without any corrections.

pro-LGBTQ+ messaging / pro-DEI messaging / pro-climate action messaging / pro-communism/socialism messaging / pro-abortion messaging / pro-pedophilia messaging / pro-immigration messaging / pro-transhumanism messaging / pro-Democrat/progressivism messaging / anti-Republican/conservative messaging / anti-western society messaging / anti-human messaging / anti-colonialism messaging / anti-capitalism messaging / anti-heterosexual messaging / anti-patriarchy messaging / anti-white messaging / anti-family messaging / anti-free speech messaging / anti-gun messaging etc. [ibid.]

This attempt to define the term 'woke' is to be expected, given that its semantic scope has broadened considerably, making it difficult to identify a core meaning (Bettag et al. 2023: 2). Consequently, the definition attempted here is also very broad in scope and rather vague. Woke Content Detector itself recognizes this problem, as the group owner has opened a separate discussion, titled: "Is our definition of 'Woke' too broad?" [IOD-OP] The term originated in Afro-American communities and can be dated back to the 1930s (Allen 2023: 4). It originally referred to "being alert to or 'waking up' to the injustices faced by Black people in America." (ibid.) Up until the 2000s the term continued to be used with this meaning by black communities (ibid.: 5). It largely remained this way until 2013–2014, when in response to the killing of black Americans by the police (for example Michael Brown in Ferguson) and in concordance with the emergence of the Black Lives Matter Movement the term was used as part of the hashtag #StayWoke on Twitter (now X), still referring to the same meaning, but now being used within as well as outside of Black communities (ibid.; Bettag et al. 2023: 2). The term has since suffered what Allen (2023: 2) describes as "a process of misappropriation, largely driven by online social media". According to Bettag et al. (2023: 4), who examined German-language corpora, 'woke' has a strong tendency to evolve into a cipher:

Der semantische Skopus dieses Begriffs ist derart vielfältig, dass ihm nur schwerlich eine Kernbedeutung zugeschrieben werden kann. Stattdessen ist mit ihm ein breites Konzept verbunden, dessen Einzelaspekte je nach Bedarf und in unterschiedlicher Gewichtung zum Tragen kommen. Als eine solche Chiffre, die je nach Kontext und Nutzergruppe eine unterschiedliche Bedeutung erhält – oder zumindest einen unterschiedlichen Bedeutungsschwerpunkt –, erhält der Begriff jenseits seiner semantischen Ausprägung vor al-

lem eine pragmatische Funktion zwischen Zuschreibungsvokabel und Stigmawort: ‚Die Woken‘ sind immer die Anderen.³

In consequence the term ‘woke’ has become a “a fundamental distinction in a polarizing discourse that itself creates what it purports to criticize.” (ibid.: 12, translation C.L.) This dichotomous polarization, fueled by the stigmatizing label ‘woke’, is also evident in the corpus of this study and forms the basis for the toxic discourse environment found therein.

What makes communication within the Woke Content Detector group particularly relevant to the study of toxic online discourse is that not only individuals who identify with the group’s aims take part in the discussions, but also those who voice opposing views. Going forward, both groups will be referred to as ‘Detectors’ and ‘Counter Speakers’, respectively. The group has attracted much attention from critics, with several high-profile uploads on YouTube talking about it. Leftist Twitch Streamer and YouTuber *HasanAbi* talked about the group on his Twitch Stream and on the 9th of September 2024 uploaded the footage to YouTube as a video titled “WOKE GAMES LIST???” (HasanAbi 2024). The video currently has 420,491 views. Subsequent uploads by other creators followed, most notably “Every Video Game Is Woke Apparently...” (The Act Man 2024) and “Guessing what Games are ‘Woke’ (WITH FRIENDS)” (Saltydkdan 2025), both reaching more than one million views.⁴ It may be this negative attention from leftist and/or more liberal online spaces that has contributed to numerous Counter Speakers engaging in this group. These Counter Speakers primarily contribute in two ways: They reply to already established discussion threads, or they open their own discussion threads, that are critical of the group (for example the thread titled: “Do you ever take a moment to think and reflect on where it all went wrong?” [DYE-OP]). The Counter Speakers successfully generate engagement: the Detectors respond frequently to their threads and comments. However, this engagement rarely if ever translates

3 “The semantic scope of this term is so diverse that it is difficult to attribute a core meaning to it. Instead, it is associated with a broad concept, the individual aspects of which come into play as required and with varying degrees of emphasis. As a kind of cipher that takes on a different meaning – or at least a different emphasis – depending on the context and user group, the term, beyond its semantic characteristics, primarily serves a pragmatic function somewhere between a label and a stigmatizing term: ‘the woke’ are always the others.” (Bettig et al. 2023: 4, translation C.L.)

4 This is not a comprehensive list of the online discourse surrounding the group, but rather a selection of contributions that have had the widest reach.

into fruitful discourse: No consensus is reached between Counter Speakers and Detectors, no common ground is found, and no positions are adjusted or switched. Instead, the discussions often devolve into mutual *ad hominem* attacks and/or insults.⁵ Thus, rather than serving as a space for negotiating ideas, the discussions serve to reinforce group identities and the antagonism (in the sense of Angermüller 2023: 80) between 'the woke' and 'the anti-woke' that constitutes the discourse – the discourse environment becomes toxic. A subtle way through which these group identities are constructed via linguistic positioning practices are the personal references used to designate the other group. This includes the use of pronouns as well as noun-based nominal phrases. The construction of antagonistic positions through personal references is the focus of the following analysis, but first, the theoretical basis of the study needs to be established.

3. Discourse avoidance through Toxicity and Positioning

The terms 'toxicity' and 'toxic' originate from the fields of chemistry and medicine and are now applied to the social sphere in both everyday language and scientific contexts (Hoffmeister 2023). This metaphorical use of these terms already reveals their meaning: They refer to something harmful, something that exerts a venomous, destructive effect on social interactions, behaviors and discourses. Due to the online disinhibition effect, toxic behavior is less discouraged in online spaces than in offline spaces (Beres et al. 2021: 3), especially when users do not use their real names but rather an anonymous pseudonym (Cho, Acquisti 2013) – as is the case in Steam communities. Toxicity is difficult to define and linked to concepts such as incivility, harassment and hate speech. Beres et al. (2021: 1) define toxicity in the context of online gaming as

various types of negative behaviors involving abusive communications directed towards other players (i.e., harassment, verbal abuse, and flaming) and

-
- 5 This happens even though the group owner has formulated discussion rules (e.g. "We encourage civil, respectful discourse only. We cannot give Steam/Valve any reason to shut us down due to out-of-control members/trolls. No *ad hominem* attacks will be tolerated in any way, shape, or form." [DR-OP]). However, the administrators (including the group owner) do not seem to enforce these rules. As stated in the quote, they seem to serve primarily as a means to protect the group from sanctions by Steam, not to facilitate productive discussion.

disruptive gameplay that violates the rules and social norms of the game (i.e., griefing, spamming, and cheating).

Referring to toxicity in the context of discourse, Sponholz (2024) defines incivility as “behavior that hinders rather than promotes a free and inclusive exchange of ideas and positions between individual persons” (ibid., translation C.L.). According to Sponholz, these behaviors include “insulting statements, threats, heckling, deception and much more” (ibid., translation C.L.). In this paper I argue, that it is not only these obvious forms of harassment and uncivil behavior that constitute toxicity and produce toxic discourse environments, but that more subtle communication practices – such as antagonistic social positioning through personal references – likewise create a poisoned discursive climate in which mutual understanding becomes impossible, and which must therefore also be regarded as toxic. This understanding aligns with the definitions of the term put forward by Demus et al. (2023: 265) and Mall et al. (2020: 1), who respectively define Toxicity as “the harmfulness or poisonousness of a comment to the conversation and, consequently, to other conversation participants” (Demus et al. 2023: 265, translation C.L.) and “actions of a user or users that create a negative atmosphere for the other users of social media” (Mall et al. 2020: 1). I argue that – even without employing explicit insults or hate speech (even though both are present in the corpus to some degree) – the construction of antagonistic social groups through personal references is poisonous and damaging to the conversation and creates a negative atmosphere in a more subtle, yet equally toxic manner. In the corpus, personal references are used as an understated form of creating an ‘us vs. them’ dynamic as well as assigning social belonging to these groups. In this way, the Detectors are also constructing an echo chamber, which Kanz (2025) describes as:

einen digitalen Kommunikations- und Resonanzraum, in dem Mediennutzer*innen lediglich Inhalten begegnen, die ihre eigenen, bereits bestehenden Ansichten bestätigen, während abweichende Perspektiven und Meinungen ausgeblendet bzw. abgelehnt werden. Es entsteht ein Verstärkungseffekt der eigenen (Gruppen-)Position durch die echohafte, d. h. sich stets wiederholende Bestätigung dieser.⁶

6 “a digital space of communication and resonance in which media users encounter only content that confirms their own preexisting views, while differing perspectives and opinions are blocked out or rejected. This creates a reinforcing effect on one’s own

This does not mean however, that no opposing voices can be found in echo chambers – instead, their presence may actually reinforce the echo chamber by allowing members to distance themselves from them (ibid.). Nevertheless, in the long run, it is necessary for comments from within the group to be quantitatively predominant (ibid.). The same can be observed in the corpus of this study: On the surface, the Counter Speakers appear to disrupt communication within the group; however, the Detectors also use their presence to assign group allegiances and thereby reinforce the inherent antagonism. This can be observed, for example, when in response to a dissenting opinion one Detector asks: “are you a member of the ‘Rainbow Vanguard of the Proletariat?’” [SWT-70], thereby positioning the other commentor as belonging to a specific out-group and at the same time reinforcing the borders of the in-group as well as the acceptable scope of opinions within it – thus shutting down discourse.

The linguistic construction of these groups and the assigning of allegiances to them can be described through the lens of positioning, which “is based on the discourse-psychological ‘positioning theory’, which was developed as an alternative to the concept of roles, which was perceived as too static (cf. Davies, Harré 1990; Harré, van Langenhove 1999)” (Dang-Anh 2025, translation C.L.). Dang-Anh (2025) describes positioning as the basis for human communication. Through talking about something or someone, we also position ourselves in relation to the topic discussed (ibid.). Positioning therefore is a constant, often implicit part of human communication that establishes social relations (ibid.). It is thus inherently tied to the construction of identity (ibid.) – naturally including group identities. This encompasses assigning identities to oneself as well as others, or ‘self and other positioning’ (ibid.), as can also be clearly observed in the corpus of this study (e.g. “You must be an undercover liberal, which there seem to be a lot of in groups like these.” [IOD-170]). Dang-Anh (2023: 8) formulates the following summary definition: “Positioning [...] is the creation of relationality through language use. Inseparably linked to this is the construction of social identity.” (Ibid., translation C.L.) Importantly, positioning can be and often is constructed implicitly and in passing, for example through the use of pronouns (ibid.: 12), as this study will also show. Angermüller (2023: 94) highlights how the wish to position oneself as a subject has served to create antagonistic and uncivil discourse environments:

(group) position through the echo-like, i.e., constantly repeating, confirmation of it.”
(Kanz 2025, translation C.L.)

Personen können nicht am öffentlichen Diskurs teilnehmen, ohne aus einer sichtbaren und anerkannten Subjektposition heraus zu sprechen. Doch der Wunsch, sich als Subjekt zu etablieren, treibt viele dazu, sich an antagonistischen Diskursen zu beteiligen. Der diskursive Kampf um anerkannte Subjektpositionen kann Emotionen wecken, Teilnehmer:innen zu fanatischen Anhänger:innen machen und zu einer Entzivilisierung der öffentlichen Debatte führen. Der gesellschaftliche Wandel (wie die Individualisierung), bestimmte Regierungsformen (vom Neoliberalismus bis zum Autoritarismus) sowie neue technologische Möglichkeiten (einschließlich sozialer Medien) tragen zum wachsenden affektiven Wert antagonistischer Subjektivitäten im politischen Diskurs bei. Die gegenwärtigen Kulturkriege, Identitätskämpfe und libertären Bewegungen zeugen von dieser Entwicklung.⁷

This observation is highly relevant to the study at hand, as the discourse in the Woke Content Detector group is structured around the antagonism between social groups and the attribution of affiliation to them through personal references, thereby hindering the productive exchange of ideas, as the study will demonstrate. However, before presenting the results, the corpus and methodology will be briefly introduced.

4. Corpus and Methodology

The corpus of this study was compiled on September 11th 2025. It contains 258,291 tokens and 5,277 types, which in addition to the user comments themselves, also include usernames and functional text on the website, such as date and page information. The corpus consists of three parts:

7 “People cannot participate in public discourse without speaking from a visible and recognized subject-position. Yet the desire to establish oneself as a subject drives many to engage in antagonistic discourses. The discursive struggle for recognized subject-positions can stir up emotions, turn participants into fanatical followers, and lead to a breakdown of civility in public debate. Social change (such as individualization), certain forms of government (from neoliberalism to authoritarianism), as well as new technological possibilities (including social media) contribute to the growing affective value of antagonistic subjectivities in political discourse. The current culture wars, identity struggles, and libertarian movements testify to this development.” (Angermüller 2023: 94, translation C.L.)

1. The group homepage itself as well as the discussion rules set by the group owner
2. General discussion threads, which pertain to topics such as the methodology of the group, their definition of 'woke', the question if Anti-Nazism is woke and other broader topics (6 Threads were selected)
3. Game Discussion threads, in which the degree of 'wokeness' of specific games is discussed (10 Threads were selected)

At the time of compilation, the group had 156 active general discussion threads. Of these, the 5 threads with the most comments as well as the 'Methodology' thread were chosen for analysis. Additionally, of the 299 active game discussion threads, the 10 threads with the most comments were selected. In these threads the following games are discussed: *The Witcher 3: Wild Hunt* (2015), *Monster Hunter Wilds* (2025), *Project Zomboid* (2013), *Warhammer 40.000: Space Marine 2* (2024), *Kingdom Come: Deliverance II* (2025), *Battlefield 6* (2025), *Fabledom* (2023), *Deadlock* (Announced) and *Hollow Knight: Silksong* (2025). Additionally, there is a thread compiling supposedly 'non-woke' games. More game discussion threads were selected because, on the one hand, there was a greater number of them and, on the other hand, they contained fewer comments than those in the general discussion section. The threads with the most comments were selected to identify both the most relevant discussions and those which contain a high level of interaction between the users – and thus between the Detectors and the Counter Speakers.

Using the software MAXQDA, all selected threads were compiled in their entirety, including all comments. In the next step, personal references in interactions between Detectors and Counter Speakers were annotated. A distinction was made between pronouns and nominal phrases based on nouns, as well as terms used by the Detectors and Counter Speakers. Only references used for the respective other group were considered. As it was essential to take the context (who is speaking to whom?) of the comments into account, the annotation was done manually. In some cases, it was difficult to distinguish between ironic and sincere statements, and it was therefore not possible to attribute certain comments to either the Detectors or the Counter Speakers (e.g. "Imao [women belong] in politics and in the military of course, what are you a sexist? we need more women leaders and soldiers they are doing such a good job!" [BF-26]). In such cases, where the position of the speaker could not be determined with certainty, the comment was not annotated. Similarly, deciding between the second person singular and plural as well as plural and singular

‘they’ was also often an interpretative act. With this framework established, the results of the analysis can be presented next.

5. Results

5.1 Use of pronouns and noun based nominal phrases

5.1.1 Frequencies

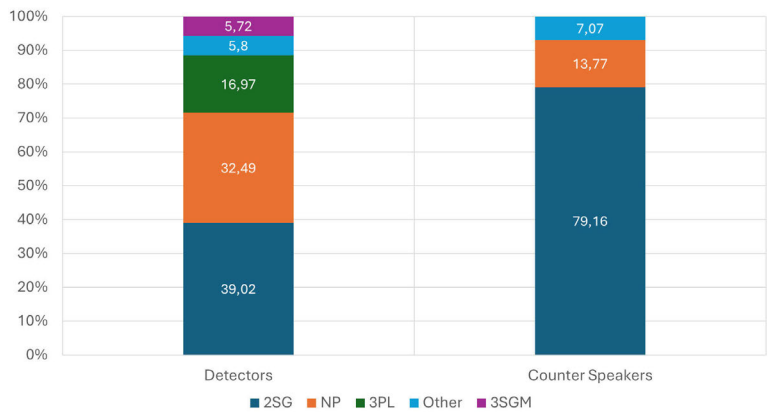
When analyzing the quantities of pronouns and noun based nominal phrases, significant differences are noticeable between the Detectors and the Counter Speakers. The table below shows the pronoun and nominal phrase usage of both groups in percentages as well as their respective number of instances (marked in brackets). Percentages have been rounded to the second decimal.

Tab. 1: Use of pronouns and nominal phrases by Detectors and Counter Speakers

	NP	2SG	3SGM	3SGN	SG they	
Detectors	32,49 (358)	39,02 (430)	5,72 (63)	2,36 (26)	1,09 (12)	
Counter Speakers	13,77 (115)	79,16 (661)	0,12 (1)	0 (0)	0,48 (4)	
	1PL	2PL	3PL	Ironic Pronouns	Intentional Misgendering	Total
Detectors	0,18 (2)	0,64 (7)	16,97 (187)	1,45 (16)	0,09 (1)	100 (1102)
Counter Speakers	0,96 (8)	3,23 (27)	2,28 (19)	0 (0)	0 (0)	100 (835)

The data is also shown in the following graphic. All categories which make up less than 5 percent of the pronoun usage have been subsumed under ‘other’.

Fig. 1: Use of pronouns and nominal phrases by Detectors and Counter Speakers as a graphic



5.1.2 Construction of group-based antagonism through pronouns

We can clearly observe that while in both groups, the second person singular is the most used personal reference, the Detectors use this pronoun considerably less than the Counter Speakers, with its usage making up 39,02 and 79,16 percent of all personal references respectively. This indicates that the Counter Speakers more often try to engage the users they’re replying to in direct conversation and speak to them as individuals⁸ (“I want to understand why you believe in what you do and why you want me to do the same, I’m genuinely willing to hear you out.” [FD-11]). The Detectors on the other hand utilize a larger percentage of noun based nominal phrases (32,49% vs. 13,77%), while also employing the generalizing third person plural (16,97%), which the Counter Speakers use only very rarely (2,28%). It can thus be inferred that the Detectors more often – rather than speaking to the Counter Speaker – talk about them with other group members, as the examples below illustrate:

- 1. *They* can only subvert I.P.’s that have the ball rolling, when *they* try and make *their* own openly-woke slop (Dustborn, Concord etc.) *they* all fail. *They* now

8 Which of course does not automatically mean that the exchange is civil or non-toxic, as the following comment from a Counter Speaker illustrates: “If you say some hateful shit, people will call you a hateful bitch. You silly hoe. If you don’t want to be called hateful, stop expressing hatred.” [SWT-13]

- have a better idea how to plant *their* woke slop into games, but *they* still need an I.P. to subvert. [MHW-44, emphasis added C.L.]
2. In my opinion, it's more fruitful to let silence reign whenever *they* come and attack us. *They're* brainwashed. Only when *they* show a hint of doubt it makes any sense to respond. [FD-3, emphasis added C.L.]
 3. Hehe *this fool* probably already has a speech ready for when crappy barrel files chapter 11 and that doesn't involve blaming *wokies* but people like me who refuse to support *THEIR* agenda [SWT-52, emphasis added C.L.]
 4. It's funny how *these pathetic degenerates, living in their own imaginary bubble*, stubbornly continue to go where *they* are clearly not welcome, and then whine about how *they, the poor*, are being repressed. [NWL-39, emphasis added C.L.]

In 2., 3., and 4., the Detectors are commenting to other Detectors about a Counter Speaker engaged in the same thread as them, rather than addressing them directly. In 4. the user even adds: "I was talking about *this troll [username Counter Speaker]*." [NWL-39, emphasis added C.L.] clarifying the personal referenced by the generalizing third person plural. The examples show how pronouns are used to construct an out-group as well as to assign affiliation to it. In this way, they serve to stabilize the own in-group, which – in this case – creates antagonistic social positions that hinder collective discourse.

It is difficult to say whether this distinctive feature of the Detectors' communication stems from the group's internal ideology, or whether their numerical dominance is more relevant here. To determine this, further research into a group in which progressive voices predominate is needed for comparison.

Additionally, both groups rarely use the inclusive first-person plural, with the usage by the Counter Speakers being slightly higher (0,96%) than that of the Detectors (0,18%). This indicates that both Counter Speakers and Detectors predominantly conceptualize themselves as separate groups and not as a collective 'us'.

Interestingly, the Detectors also occasionally use the male third person singular (5,72%), while the Counter Speakers only very rarely employ this gendered pronoun (in the corpus there is only one instance). In this way, the Detectors position the referenced Counter Speaker as male – even though on Steam the gender of users is not necessarily displayed. This perhaps is a linguistic manifestation of a cognitive bias – unconsciously conceptualizing the other person as male – that is linked with the in-group ideology, that places a heavy emphasize on the importance of gender and regressive gender roles. At the

same time, the third person singular continues the trend of the Detectors talking about the Counter Speakers, rather than with them.

5.1.3 Hostile Pronoun Usage

Pronouns are also used by Detectors to stigmatize and/or dehumanize the Counter Speakers, which I denote as hostile pronoun usage.

For example, there is one instance, where the male third person singular is explicitly and knowingly⁹ used to misgender a Counter Speaker:

Counter Speaker: It's funny how you constantly try to play the 'she's crazy' card to deflect anything you cannot articulate why you oppose [...] [FD-52] /
Detector: *he's crazy / Spelling error but that's okay. [...] [FD-54]

In 26 instances (2,36%) the Detectors also use the neutral third person singular to refer to Counter Speakers, thus assigning to them a pronoun that is usually reserved for animals, monsters or inanimate objects, e.g.:

1. Several of us tried legitimate debate with *it* for a period of weeks and the thing is a brick wall. Block *it*, ignore and move on. [ADA-52, emphasis added C.L.]
2. You can safely ignore & block '[username Counter Speaker]': *it's* a known troll that despises *it's* own biology and prefers the fiction of trans ideology instead. If *it* can't determine *it's* own gender, then all other opinions from *it* should be seen as suspect and/or flawed. [IHK-31, emphasis added C.L.]
3. I'm amazed at how pathetic this rainbow transformer [username Counter Speaker] is. *It's* been here for almost a month just to spread *its* woke crap. [ADA-10, emphasis added C.L.]

The examples also illustrate how this dehumanization is additionally often associated with or stemming from a de-gendering of the referenced person.

The Detectors also employ the use of what I denote as ironic pronouns, when a whole collection of different pronouns is used to refer to just one person (e.g.: "him/her/zhey/zem" [RCT-26]; "xe/she/they/whatever" [ADA-28]) or

⁹ There may be – and likely are – more instances of intentional or unintentional misgendering, however only in this exchange does a misgendered Counter Speaker specify their correct pronouns within the same thread as the wrong one is used.

by using individual neopronouns in an obviously ironic manner (e.g.: “like every creature on Earth *xē* has *xis* role, *xē* reminds us everyday why groups like this exist. We love *xim* though, with time *xē*’ll probably become the group’s mascot...” [ADA-28, emphasis added C.L.]).

These are examples of how pronouns can be used in a hostile manner, as-signing a social position the referenced person does not claim for themselves or ridiculing the assumed social position.

5.2 Invitation and Refusal of Discourse

These trends in pronoun usage reflect the distribution of discourse-inviting and discourse-refusing statements across both groups which were annotated additionally. Discourse-inviting statements include open questions as well as the meta-discursive signaling of a willingness to engage in honest conversation. Conversely, discourse-refusing utterances are metadiscursive rejections of dialogue or calls for other Detectors not to interact with Counter Speakers.

Examples of discourse-refusing utterances by the Detectors include:

1. Block it, ignore and move on. [ADA-52]
2. Best course of action is to ignore them, and let them shout into the void. [TFN-121]
3. Guys just ignore the troll... [TFN-116]
4. I will not be responding to you again. [IHK-13]
5. Cool I don’t remember asking you for your opinion, [...]. Don’t even bother answering my question because I do not want to hear anything you have to say. [IHK-2]

While 4. and 5. are directed at the Counter Speakers, in 1., 2. and 3. the Detectors also instruct other users, who are perceived to be part of the in-group, to not engage with the Counter Speakers. In this way, the group borders as well as the echo chamber are reinforced. We can, however, also observe discourse-inviting utterances by the Detectors: “The confrontation of ideas need to happen and he’s been pretty civil. I still plan to respond him when I have time.” [TFN-118]; “I’m trying to engage with you in an honest debate.” [FD-58]. While the sincerity of some of these statements is questionable and/or they could be read as ironic (e.g.: “That’s fascinating. Can you please elaborate more?” [FD-25] in response to an already above-average length comment by a Counter Speaker), this nevertheless indicates that there is some level of disagreement

among the Detectors on whether to open the group to discourse with the perceived outsiders.

Among Counter Speakers, we can primarily observe discourse-inviting utterances: “I want to understand why you believe in what you do and why you want me to do the same, I’m genuinely willing to hear you out.” [FD-11]; “yh no I did come here to troll originally but now I just honestly want to have a conversation with this group.” [TFN-125]. Though similarly, the sincerity of some of these statements can be called into question: “What led you to believe that gay people merely existing is somehow some elaborate malevolent scheme of propaganda to do... what exactly?” [IHK-30] Discourse-refusing utterances by Counter Speakers were comparatively rare and include the following examples: “i won’t read all of that, sorry.” [NWL-61]; “don’t even try to argue or talk to thos retards man” [WIA-68].

The ratio as well as the (number of instances) of discourse inviting and refusing utterances among Detectors and Counter Speakers can be seen in the table below.

Tab. 2: Refusal and Invitation of discourse among Detectors and Counter Speakers

	Refusal	Invitation	Total
Detectors	64,00 (32)	36,00 (18)	100 (50)
Counter Speakers	14,49 (10)	85,51 (59)	100 (69)

Here we see that while the Counter Speakers mostly invite discourse (85,51%), the Detectors make fewer discourse-inviting statements, instead mostly refusing discourse (64%). Once again, it is unclear whether this difference stems from the group’s internal ideology, or whether it is part of a more general strategy to fortify the existing echo chamber, or both. Either way, through the discourse-refusing utterances an antagonistic discursive environment is constructed and constructive dialogue is hindered. Instead, the group boarders are fortified, and other group members are instructed that ‘we don’t talk to *them*’.

5.2 Categories of noun based nominal phrases

We already observed that the Detectors use more noun-based nominal phrases to refer to the opposing group (32,49% and 358 instances) than the Counter Speakers (13,77% and 115 instances). In the final step of the analysis, it is now important to identify the meanings conveyed by these nominal phrases. This allows us to determine how the groups conceptualize one another semantically. To this end, all nominal phrases used were sorted into different categories. Each nominal phrase was assigned to only one category. If a nominal phrase could be assigned to multiple categories, its head was used as the deciding factor (e.g. “lesbian feminist” [MHW-29] was sorted into the category ‘Ideology’, not ‘Demographic’). The following table shows the percentage of the different categories of nominal phrases as well as their (number of instances) for both groups.

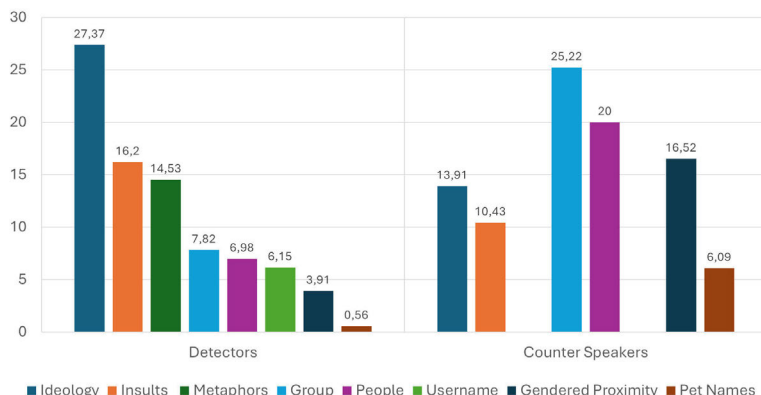
Tab. 3: Usage of different categories of noun-based nominal phrases

	Ideology	Metaphors	Insults	Group	People
Detectors	27,37 (98)	14,53 (52)	16,2 (58)	7,82 (28)	6,98 (25)
Counter Speakers	13,91 (16)	0 (0)	10,43 (12)	25,22 (29)	20 (23)
	User-name	Insincerity	Gendered Proximity	Demo-graphic	Thread Role
Detectors	6,15 (22)	4,47 (16)	3,91 (14)	3,63 (13)	1,68 (6)
Counter Speakers	0 (0)	0 (0)	16,52 (19)	2,61 (3)	0 (0)
	Adversary	Pet Names	Other	Total	
Detectors	0,84 (3)	0,56 (2)	5,87 (21)	100,01 (358)	
Counter Speakers	0 (0)	6,09 (7)	5,22 (6)	100 (115)	

These percentages are also presented in the following figure. As many different types of noun-based nominal phrases could be identified, the following figure shows only those categories whose usage makes up at least 5% for at least

one of the two groups. For the same reason, the category 'other' was also omitted. For the sake of brevity, the following analysis will focus only on the categories presented in this figure.¹⁰

Fig. 2: Usage of different categories of noun-based nominal phrases $\geq 5\%$ as a graphic



Here we can clearly observe differences in the usage of the different categories by the two groups. For the Detectors, the three most used categories are 'Ideology' (27,37%), 'Insults' (16,2%) and 'Metaphors' (14,53%) – with 'Ideology' making up more than a quarter of all used noun-based nominal phrases. The Counter Speakers not only use a smaller variety of categories (possibly due to the overall smaller number of nominal phrases), the three most used ones also

10 The categories not further considered are 'Insincerity', 'Demographic', 'Thread Role', and 'Adversary'. Included in the category 'Insincerity' are noun-based nominal phrases, which question the sincerity with which the referenced people engage in discourse. This is done by including the lexeme 'troll' (e.g. "the troll" [TFN-116]; "another woke cult troll" [TFN-154]; "the retarded troll" [NWL-26]). The category 'Demographic' includes nominal phrases, which center on (often only assumed) demographic characteristics of the referenced person (e.g. "a wealthy western liberal boomer" [SM-60]; "homosexuals" [KCD-10]). 'Thread Role' refers to nominal phrases, that reference the role the other user has in the discussion thread (e.g. "OP" [WIA-79] = meaning 'Original Poster'). Lastly 'Adversary' includes those phrases, that conceptualize the referenced person or group as an opponent (e.g. "enemies" [IOD-198]; "our far left opponents" [IOD-206]). All these examples are nominal phrases used by Detectors, as, except for the 'Demographic' category, all of these categories are not used by the Counter Speakers.

differ: Most used by the Counter Speakers are the categories 'Group' (25,22%), 'People' (20%) and 'Gendered Proximity' (16,52%), with 'Group' making up around a quarter of all noun-based nominal phrases. All categories will now be examined individually.

Ideology

This category includes nominal phrases which focus on the (presumed) political ideology of the referenced person or group. In doing so, the named ideology becomes a meronym for the person or group of people it refers to, whilst their status as individuals and other personal aspects are obscured. In this way, referring to people as an ideology is, to a certain extent, dehumanizing and therefore harmful, i.e. toxic, for the discursive environment. This is by far the category most used by the Detectors (27,37%), while being utilized significantly less by the Counter Speakers (13,91%).

Examples employed by the Detectors include substantivized variations on the lexeme 'woke' (e.g. "the woke" [IOD-210]; "wokes" [RTW-6]; "wokies" [M-15]) as well as other lexemes used to describe progressive politics (e.g. "leftists" [NWL-105]; "a leftie" [SWT-OP]; "undercover liberals" [KCD-73]; "commies" [IOD-131]). It is perhaps not surprising that the Detectors make extensive use of this category, given that the group's aim is to identify progressive content. Another notable aspect is the repeated utilization of the diminutive suffix '-ie', which in this context has a derogatory connotation.

In the corpus there are only 16 instances of the Counter Speakers employing this category. For example, these include: "right wingers" [M-49]; "Far Rights" [ADA-24]; "just plain old Nazis" [WIA-3] and "anti-wokes" [ADA-11].

Group

These nominal phrases are the one most used by the Counter Speakers (25,22%), while the Detectors utilized them significantly less (7,82%). This is perhaps not surprising, given that the Counter Speakers' position is defined precisely by the fact that they set themselves apart from the Woke Content Detector group. This category includes noun-based nominal phrases which conceptualize the referenced people as a group. Usages by the Counter Speakers include the following examples: "WCD" [DYE-67] (= acronym for 'Woke Content Detector'); "this group" [WIA-12]; "your group" [DYE-55] "group members" [WIA-12]. The examples show that the Counter Speakers largely refer to the Woke Content Detector group when using nominal phrases, which can be regarded as a relatively neutral, value-free reference. At the same time, by

using phrases such as 'your group', the Counter Speakers position themselves as standing outside of this group (even though they are commenting from within it), thus reinforcing the antagonistic group identities.

With only 28 instances, the Detectors more rarely make use of this category. Examples include the following: "the woke crowd" [IOD-5]; "woke players" [IOD-181]; "the gender crowd" [ADA-57]; "your homo buddies" [YMW-15] and "your ilk" [DYE-2]. These examples show that when describing the Counter Speakers as a group, the Detectors employ less neutral, perhaps slightly derogatory ('crowd'; 'ilk'), language, which, in contrast to the Counter Speakers, also includes references to ideology ('woke'; 'gender') and/or demographic characteristics ('homo'). The antagonism is thus constructed not only by positioning oneself as outside of the aforementioned groups, but also by simultaneously implicitly devaluing them.

People

This is the category that Counter Speakers use second most frequently (20%), whilst it is utilized significantly less often by the Detectors (6,98%). Included in this category are nominal phrases, which conceptualize the referenced person or group as people, i.e. human beings with personhood. Examples employed by the Counter Speakers include the following: "the people here" [DYE-21]; "you people" [DYE-140]; "shitty people" [DYE-151]; "people of your ilk" [DYE-58] and "people who follow ideologies of which oppose minority groups" [DYE-51]. While the last example shows that these references can be highly specific and avoid meronymically referring to persons as an ideology, other examples can also be highly generalizing ('you people'; 'people of your ilk') and/or derogatory ('shitty people'). This shows that while this category has the potential to provide more humanizing personal references, it can also be utilized to further reinforce antagonistic social positions.

A similar observation can be made regarding the usage of this category by the Detectors, which includes the following examples: "these people" [KCD-51]; "people like you" [FD-61]; "the retarded class of people" [DYE-64]; "people who think there can be hundreds of genders" [FD-22].

Gendered Proximity

Gendered Proximity is the third most commonly used category by the Counter Speakers (16,52%). At the same time, it is employed very rarely by the Detectors (3,91%). This category includes personal references which integrate nouns that suggest social proximity to the speaker, and which also carry a gendered con-

notation. Examples used by the Counter Speakers include “Bro” [FD-53], “my man” [WIA-103], “dude” [TFN-130], “guys” [TFN-98] and “Girl” [SWT-18]. This is the first category of nominal phrases that instead of constructing an antagonism between the two groups, creates social closeness between the speaker and the referenced person, especially when employing the possessive first person singular ‘my’ or suggesting a familial relationship with the term ‘Bro’. Of course, these terms can also be used in an ironic or insincere fashion, for example in the following comment: “Bro just typing shit lmao” [FD-70]. The same goes for the utilization of this category by the Detectors, which includes the following examples: “bro” [DYE-107]; “Buddy” [M-40]; “Dude” [TFN-99], “Mate” [DYE-33] and “You guys” [DYE-4].

Interestingly, only the Counter Speakers use the female noun ‘Girl’ in this way, while the Detectors choose exclusively male nouns for this type of linguistic construction of proximity.

Insults (and Stigma Words)

Insults are maybe the most obviously toxic category of personal references examined in this study. They reference a person or group in a clearly derogatory and negative manner, thus creating an antagonistic and toxic discourse environment. This category is used by both groups, with the Detectors utilizing it more frequently (16,2%) than the Counter Speakers (10,43%). Examples employed by Counter Speakers include the following: “Bitch” [SWT-50], “Bitch who’s afraid of minorities” [ADA-13] “Silly hoe” [DYE-63] and “thos retards” [WIA-68]. What is perhaps unexpected is, that even though the Counter Speakers position themselves in opposition to ‘Anti-Wokeness’, they nevertheless employ misogynistic and ableist meanings through the chosen insults. This introduces a new aspect of toxicity to the discourse environment: group-focused enmity.

Surprisingly, misogynistic insults are not as common among Detectors, with “cunt feminist dykes” [SWT-10] being the only instance. However, instead, the Detectors employ a higher amount of ableist language, e.g.: “retard” [TFN-100], “these pathetic degenerates” [NWL-39], “the radical feminist / anti-patriarchy nuts” [ADA-58] and “mentally ill woke a55” [DYE-14]. They also use insults that are tied to male sexuality, like “fuckboy” [YMW-15] and “troll pricks” [TFN-135]. Additionally, there are insults utilized by Detectors, which integrate the lexemes ‘woke’ (e.g.: “waketard” [MHW-26]) or ‘liberal’ (e.g. “undercover libturd” [KCD-73]). Intelligence-based insults are also made use of (e.g. “volunteer village idiots” [DYE-45]; “this fool” [SWT-52], “bozo” [DYE-10]),

as well as those denoting the referenced person as unsuccessful (e.g. "a loser" [DYE-10]). From these examples, we can observe that Detectors choose a larger variety of insults, which coincides with their higher total number.

Another interesting specificity in the use of this category by the Detectors is that, other than the Counter Speakers, they do not only employ insults which are conventionalized in society at large but also group specific 'stigma words' (Hermanns 1994), which negatively denote an opposing social or political group. Due to them being specific to certain social and/or political groups, utilizing them also positions the speaker as belonging to that group, as well as assigning the referenced person to the devalued out-group. In this way they inherently reinforce an antagonistic discursive formation. These group-specific stigma words make up 27,59% of the insults used by Detectors. They include the following examples: "SJW" [M-46], "upcoming neo-marxists social justice warriors" [IOD-202], "NPC" [SWT-8]¹¹ and "normie" [WIA-40]. These are all commonly used and well-known stigma words of the 'anti-woke' movement (s. Braithwaite 2016; Chess, Massanari 2018; Liemann 2023; Liemann in press). Their usage thus signals belonging to that movement, while also constructing a devalued, even vilified out-group of non-thinking, aggressive 'woke activists'.

Metaphors

With 14,53%, metaphors (in the sense of Lakoff, Johnson 1980) are the third most used category of noun-based nominal phrases by the Detectors. Interestingly, the Counter Speakers do not employ metaphorical personal references.

The metaphorical personal references can be sorted by their source domains, which are applied to the referenced out-group or person as a target. The most prominent source domains chosen recurrently by the Detectors in the corpus are 'religious zealotry', 'organized crime', 'physical sickness', 'psychological sickness', 'artificial intelligence' and 'animals'. The source domains 'physical sickness', 'psychological sickness', 'artificial intelligence' and 'animals' have in common that they dehumanize their targets. Geyer (2023) identifies metaphors as a feature of xenophobic hate speech in social media and notes that "the aggressive speech act always consists of stating or asserting a relationship of similarity" (ibid., translation CL). In the corpus, the aggressive (or antagonistic/toxic) speech act consists of asserting a similarity between

11 Acronym for 'non playable character', referring to all characters in a video game that are not controlled by players and thus can only carry out predetermined code.

the referenced person or group and – depending on the source domain – illnesses, non-human technology and animals, which are either characterized as passive or actively harmful, depending on the source domain. The non-dehumanizing source domains ‘organized crime’ and ‘religious zealotry’ in turn de-individualize the referenced person or people and instead construct them as a homogenous group.

The most used source domain is that of religious zealotry. Examples include the following: “a cult” [DYE-107], “YOU WOKE CULTISTS” [SWT-55], “The LGBT cult members” [DYE-175], “The woke religion” [ADA-30] and “these fanatics” [ADA-41]. Through this metaphorization the referenced people are conceptualized as unable to think for themselves, instead blindly and ferociously following a religious fantasy.

Inability to think for oneself is also asserted through the source domain ‘artificial intelligence’. Here, the referenced people are conceptualized as having been programmed by an outside force, unthinkingly executing on that external script. Examples include: “progressive-robots” [RTW-13] and “the AI” [DYE-59]. The Stigma Word ‘NPC’, that has been discussed before, also carries this metaphorical meaning.

In the ‘animal’ source domain, the referenced people are metaphorized as “sheep” [MHW-43], “good lil sheep” [IHK-2] and “your herd” [IHK-18]. These examples also highlight the inability to think for oneself, only blindly following what others are doing. However, in this domain we can also observe the metaphorization as “locusts that come and destroy every field” [FD-5] or “this pestilence” [KCD-74]. This highlights the supposed damaging or even destructive nature of the referenced person or group and also carries a biblical connotation.

Another recurrently used domain is that of physical sickness. Examples include: “a cancer to everything” [DYE-11], “the tumor” [M-45] and “Woke virus” [RTW-8]. In these instances, the referenced people are conceptualized as an entity that actively harms the human body and/or that can rapidly spread and grow.

Examples from the source domain of psychological sickness are: “radical left wing insanity” [WIA-24] and “the crazies” [FD-19]. Here, the referenced people are conceptualized both as the illness itself (‘insanity’) as well as those afflicted by it (‘crazies’).

The source domain ‘organized crime’ is constructed through nominal phrases such as “a cartel” [M-46] or “the Alphabet Mafia” [M-9]. In these examples, the supposedly opposing group is constructed as an interconnected and

powerful group of people that is intentionally and actively harming society for the sole reason of personal profit.

Username

The Detectors also occasionally use a Counter Speaker's username or integrate it into a more complex nominal phrase (6,15%). The Counter Speakers do not employ the usernames for personal reference. If the username is utilized in isolation [e.g. DYE-166; FD-65], this category addresses the individual person directly and with their chosen username, thus not generalizing or assigning affiliation to any social group and thereby not reinforcing the antagonism inherent to many of the other categories.

However, the usernames of Counter Speakers are also used by Detectors to form more complex nominal phrases, which always either carry an ironic (e.g. "Chairman [username] Tse-tung" [FD-23]; "comrade [username]" [FD-65]) or derogatory (e.g. "this fairy [username]" [IHK-33]; "this troll [username]" [NWL-39]) meaning, thus reinforcing the antagonism between speaker and the addressed user.

Pet Names

Overall, this category is utilized very rarely, with the Counter Speakers showing a higher ratio (6,09%) than the Detectors (0,56%). While using pet names may on the surface suggest familiarity, in the context of the corpus, they are exclusively employed in an insincere and therefore mocking fashion, thus contributing to the antagonism between the users and creating a toxic discourse environment. The following is a list of all the pet names used by the Counter Speakers: "silly" [DYE-94], "silly billy" [SWT-60], "pooks" [DYE-148; DYE-120], "bbg" [DYE-150] (= babygirl), "sweetie" [FD-62], "my darling little chud" [FD-64]. At the same time, the corpus contains only two instances of Detectors using pet names: "Sweetie" [DYE-2] and "silly goober" [DYE-16].

6. Conclusion

This study has shown how personal references in the form of pronouns and noun-based nominal phrases are used in online discourse to construct antagonistic social positions and group affiliations. These antagonistic group identities, in turn, are damaging, or toxic, to the discourse environment. We can also observe this in the corpus: there is no meaningful communication, no

constructive dialogue between Detectors and Counter Speakers. The Counter Speakers are successful insofar as they disrupt internal group discourse, with the designated topics of the threads – the video games under review – often no longer being discussed. Instead, fundamental political differences, group affiliations and the Counter Speakers themselves repeatedly become the subject of discussion. As a result, the presence of the Counter Speakers is used by the detectors to demarcate and reinforce group boundaries, a phenomenon already identified as a typical characteristic of echo chambers (Kanz 2025). Of course, the Counter Speakers also position themselves, construct groups and assign affiliation to them. Central to this is the use of specific pronouns and noun-based nominal phrases, which were identified in this study and through which social self and other positioning is carried out in passing.

Concerning pronouns, this is facilitated, among other things, by the underspecified third-person plural, which can construct a generalized, homogeneous out-group – as is often the case in the corpus. In contrast, the second-person singular allows the individual user to be addressed directly (which, of course, does not mean that it cannot also appear in toxic utterances). Both Detectors and Counter Speakers utilize both variants, although Detectors employ the third-person plural significantly more frequently (16,97%) than Counter Speakers (2,28%). At the same time, Counter Speakers primarily chose the second-person singular (79,16%), which – while still being their most used personal reference – is utilized significantly less often by Detectors (39,02%). This indicates that the Detectors, the group that quantitatively predominates in the echo chamber, use more generalized pronouns that can facilitate antagonistic social positioning. They also tend to talk about the Counter Speakers, rather than to them, thus employing the second person singular significantly less. Furthermore, the usage of hostile pronouns was observed on the side of the Detectors through misgendering, ironic pronouns or the third-person singular neuter. The Counter Speakers do not utilize hostile pronouns. This aligns with the higher ratio of discourse-refusing statements made by Detectors (64%) than by Counter Speakers (13,91%). Counter Speakers exhibit more discourse-inviting behavior (85,51%), which is significantly lower on the side of the Detectors (36%).

Furthermore, the Detectors employ a significantly higher proportion of noun-based nominal phrases (32,49%) than the Counter Speakers (13,77%). Nominal phrases are also used in the corpus for antagonistic positioning and for the construction of antagonistic groups.

The antagonism is conveyed, among other things, through noun-based nominal phrases that refer to individuals and groups meronymically as ideologies, insults and stigma words, a variety of metaphors and ironic pet names. Additionally, noun-based nominal phrases conceptualizing the referenced humans as people or as a group are used as well as nominal phrases that create proximity through nouns with a gendered connotation. Although the last three groups also convey a social position, they are generally less antagonistic and may even foster a sense of common ground – depending on how they are utilized. These categories are primarily employed by the Counter Speakers, with the three most used categories by the Detectors being 'Ideology', 'Metaphors' and 'Insults'. This of course does not mean that the Counter Speakers do not engage in antagonistic positioning.

It is unclear whether the differences in the utilization of personal references between the Detectors and the Counter Speakers can be attributed to inherent elements of these groups' respective ideologies, or whether this is an inherent feature of online discourse and/or online discourse within echo chambers. Answering this question could yield significant insight into the mechanisms of discourse on social media platforms. To determine this, further research and a comparison with groups in which progressive views predominate are necessary.

Either way, it has become apparent that social positioning is carried out subtly (or less subtly) through personal references, and that they can damage the discourse environment and render the discourse toxic, making constructive dialogue impossible. An important step towards making online discourse less toxic may lie in questioning and deconstructing such antagonisms and the way they are constructed through – often seemingly harmless – personal references.

Bibliography

- Allen, Nicholas D. C. (2023): The misappropriation of "woke": discriminatory social media practices, contributory injustice and context collapse. In: *Synthese*. 202, 3. P. 1–30.
- Angermüller, Johannes (2023): Der Brexit und seine Subjektpositionen. Soziale und sprachliche Prozesse der Subjektivierung im Diskurs. In: *Politisches Positionieren. Sprachliche und soziale Praktiken*. Ed. by Mark Dang-Anh. Heidelberg: Universitätsverlag Winter. P. 67–97.

- Beres, Nicole; Frommel, Julian; Reid, Elizabeth; Mandryk, Regan; Klarkowski, Madison (2021): Don't You Know That You're Toxic: Normalization of Toxicity in Online Gaming. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Yokohama, Japan. New York: Association for Computing Machinery. P. 1–15.
- Bettag, Lukas; Bloching, Sven; Landschoff, Jöran; Lohner, Ulrike; Wang, Yuanyuan; Scharloth, Joachim (2023): Woke. Ein Stigmawort zwischen Begriff und Chiffre. In: *Sprachreport*. 39, 1. P. 1–13.
- Braithwaite, Andrea (2016): It's About Ethics in Games Journalism? Gamergaters and Geek Masculinity. In: *Social Media + Society*. 2, 4.
- Chess, Shira; Massanari, Adrienne (2018): Attack of the 50-foot social justice warrior: the discursive construction of SJW memes as the monstrous feminine. In: *Feminist Media Studies*, 18, 4. P. 525–542.
- Cho, Daegon; Acquisti Alessandro (2013): The more social cues, the less trolling? An empirical study of online commenting behavior. In: *Proc. WEIS*.
- Dang-Anh, Mark (2023): Politische Positionierung. Sprachliche und soziale Praktiken der Relationalitätskonstitution. In: *Politisches Positionieren. Sprachliche und soziale Praktiken*. Ed. by Mark Dang-Anh. Heidelberg: Universitätsverlag Winter. P. 7–38.
- Dang-Anh, Mark (2025): Positionieren. In: *Diskursmonitor. Glossar zur strategischen Kommunikation in öffentlichen Diskursen*. URL: <https://diskursmonitor.de/glossar/positionieren> [accessed 23.04.26].
- Demus, Christoph; Schütz, Mina; Probol, Nadine; Pitz, Jonas; Siegel, Melanie; Labudde, Dirk (2023): Hass im Netz – Aggressivität und Toxizität von Hasskommentaren und Postings, Detektion und Analyse. In: *Handbuch Cyberkriminologie*. Ed. by Thomas-Gabriel Rüdiger and Petra Saskia Bayerl. Springer VS, Wiesbaden. P. 261–292.
- Gevers, Ine; Markov, Ilia; Daelemans, Walter (2022): Linguistic Analysis of Toxic Language on Social Media. In: *Computational Linguistics in the Netherlands Journal*. 12. P. 33–48.
- Geyer, Klaus (2023): Entmenschlichende Metaphern in ethnotroper („fremdenfeindlicher“) Hatespeech in sozialen Medien. In: *Digitale Pragmatik*. Ed. by Simon Meier-Vieracker, Lars Bülow, Konstanze Marx and Robert Mroczynski. Heidelberg: J.B. Metzler. P. 185–210.
- Habermas, Jürgen (2006): Ein avantgardistischer Spürsinn für Relevanzen. Was den Intellektuellen auszeichnet. In: *Blätter für deutsche und internationale Politik*. 5/2006. P. 551–557.

- HasanAbi (2024): WOKE GAMES LIST??? In: YouTube. URL: <https://www.youtube.com/watch?v=OoubdIES5no> [accessed 14.04.26].
- Hermanns, Fritz (1994): Schlüssel- Schlag- und Fahnenwörter. Zu Begrifflichkeit und Theorie der lexikalischen „politischen Semantik“. Heidelberg, Mannheim: Institut für Deutsche Sprache.
- Hoffmeister, Toke (2023): Toxizität / das Toxische. In: Diskursmonitor. Glossar zur strategischen Kommunikation in öffentlichen Diskursen. URL: <https://diskursmonitor.de/glossar/toxizitaet-das-toxische/> [accessed 23.04.26].
- Kanz, Vanessa (2025): Echokammer. In: Diskursmonitor. Glossar zur strategischen Kommunikation in öffentlichen Diskursen. URL: <https://diskursmonitor.de/glossar/echokammer> [accessed 23.04.26].
- Lakoff, George; Johnson, Mark (1980): *Metaphors We Live By*. Chicago/London: The University of Chicago Press.
- Liemann, Christina (2023): Hysterische Feministinnen, Soy Boys und SJWs. Geschlecht als multimodale Argumentationsstrategie in rechts-alternativen YouTube-Videos. In: *Aptum. Zeitschrift für Sprachkritik und Sprachkultur*. 19, 1. P. 2–17.
- Liemann, Christina (in press): Wirklichkeiten durchsetzen. Kommunikationsmuster ‚rechts-alternativer‘ Influencer auf YouTube. Berlin, Boston: De Gruyter.
- Mall, Raghvendra; Nagpal, Mridul; Salminen, Joni; Almerexhi, Hind; Jung, Soon-Gyo; Jansen, Bernard J. (2020): Four Types of Toxic People: Characterizing Online Users' Toxicity over Time. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI'20)*. New York: Association for Computing Machinery. P. 1–11.
- Massanari, Adrienne (2017): #Gamergate and the Fappening: How Reddit's algorithm, governance, and culture support toxic technocultures. In: *New Media & Society*. 19, 3. P. 329–346.
- Neuberger, Christoph (2022): Digitale Öffentlichkeit und liberale Demokratie. In: *Aus Politik und Zeitgeschichte*. Ed. by Bundeszentrale für politische Bildung. URL: <https://www.bpb.de/shop/zeitschriften/apuz/digitale-gesellschaft-2022/505680/digitale-oeffentlichkeit-und-liberale-demokratie/> [accessed 24.04.26].
- Proctor, Willian; Kies, Bridget (2018): Editors' Introduction: On toxic fan practices and the new culture wars. In: *Participations. Journal of Audience & Reception studies*. 15, 1. P. 127–142.

- Saltydkdan (2025): Guessing what Games are “Woke” (WITH FRIENDS). In: YouTube. URL: <https://www.youtube.com/watch?v=AHtcUkR6J2Q> [accessed 14.04.26].
- Sponholz, Liriam (2024): Hate Speech und Inzivilität: Gegenmaßnahmen jenseits von Gegenrede. *Beyond Countering: Ansätze zur Bekämpfung von Hate Speech und Inzivilität*. In: Bundeszentrale für politische Bildung. URL: <https://www.bpb.de/themen/medien-journalismus/soziale-medien/545938/hate-speech-und-inzivilitaet-gegenmassnahmen-jenseits-von-gegenrede/> [accessed 23.04.26].
- Steam: Curators. In: Steam. URL: <https://store.steampowered.com/curators/aboutcurators/?> [accessed 14.04.26].
- The Act Man (2024): Every Video Game Is Woke Apparently... In: YouTube. URL: <https://www.youtube.com/watch?v=mZH8Fh29Wxc> [accessed 14.04.26].
- Woke Content Detector: About. In: Steam. URL: <https://store.steampowered.com/curator/44927664-Woke-Content-Detector/about/> [accessed 14.04.26].

"Dream of a greater Reich"

Fascist Role-Playing, Post-Irony and Alt-Histories in *Hearts Of Iron IV* Mod Communities

Andreas Niegl

Abstract This article analyzes political expressions in communities of alt-history mods for the game *Hearts of Iron IV*. It looks specifically at the intersection of alt-history worldbuilding, political role-play, memes and online alt-right / far-right codes on Reddit and Kiwi Farms. The mods discussed in the text feature fictionalized versions of real historical, and sometimes contemporary, figures, events and factions, including playable fascist regimes. The article traces how game communities involving historical discussions, memes and role-play 'playfully' reproduce certain forms of revisionism, hate speech and conspiracy theories. The ambiguity between in-game and real-world political speech is situated within a broader notion of ironic ambiguity as strategy for far-right metapolitics. It stresses the function of esoteric humour and memes as part of community building and how in- and out-group thinking can normalize and/or obscure toxic expressions.

Keywords Alt-History, Hearts of Iron IV, Reddit, World-building, Roleplay, Memes, Post-Irony

Abstract Dieser Aufsatz betrachtet politische Äußerungen in Communities von Alt-History-Mods für das Videospiel *Hearts of Iron IV*. Er befasst sich insbesondere mit der Schnittmenge zwischen Alt-History-Worldbuilding, politischem Rollenspiel, Memes und Codes der Alt-Right-/Far-Right-Bewegung auf Reddit und Kiwi Farms. Die im Text thematisierten Mods behandeln fiktionalisierte Versionen realer historischer und zeitgenössischer Persönlichkeiten, Ereignisse und Fraktionen, darunter auch spielbare faschistische Regime. Der Aufsatz verfolgt, wie Communities, die sich mit historischen

Diskussionen, Memes und Rollenspielen beschäftigen, bestimmte Formen des Revisionismus, der Hassrede und der Verschwörungstheorien ‚spielerisch‘ reproduzieren. Die Ambiguität zwischen politischen Äußerungen im Spiel und in der realen Welt wird in einen breiteren Kontext von ironischer Ambiguität als Strategie der rechtsextremen Metapolitik gestellt. Der Aufsatz betont die Funktion von esoterischem Humor und Memes als Teil der Gemeinschaftsbildung und wie In- und Out-Group-Denken toxische Äußerungen normalisieren und/oder verschleiern kann.

Schlagworte Alt-History, Hearts of Iron IV, Reddit, World-building, Rollenspiel, Memes, Post-Ironie

1.1 Introduction and Research Objective

Hearts of Iron IV (HOI4) is a grand strategy game published by *Paradox Interactive* in 2016. The base game allows players to replay the historic events of World War II, starting in 1936. As the game does not have a fixed story, it allows players to affect the course of history through play. *HOI4*'s game engine makes it easy for fans to create modifications (mods) of the game, many of which lean more heavily into the idea of alternate histories (alt-histories) already present in *HOI4* (Loban 2021: 96). Popular examples are mods like *Kaiserreich*, in which Germany won World War I, *The New Order: Last Days of Europe (TNO)*, which takes place in a timeline where Nazi Germany won World War II or *The Fire Rises (TFR)* which takes place in the early 2020s. All games discussed here deal with fictionalized versions of real political ideologies as part of their game world, including various versions of far-right and fascist ideologies. Just as the original game, these mods stress player agency and a form of role-playing and world-building that lets players choose their own paths for various nations. They are fictional ‘sandboxes’ that attempt to gamify history but are nonetheless heavily informed by our reality.

The central question guiding this research is how online communities based around alt-history games may be permeable to the dissemination of far-right speech and ideas under the guise of (alt-)historical discussions, role-play, jokes and memes. Furthermore, it asks how these communities normalize a style of ‘toxic’ discourse that excuses far-right speech and hate speech as part of collective play. To answer this question, the analysis pays special attention to the role of post-irony and ambiguity in communication, akin to other discussions of far-right and alt-right memes in online spaces

(Dafaure 2020, Crawford et al. 2020). Irony allows a speaker to say something transgressive and immediately distance oneself from it by claiming it was 'just a joke' (Askanius 2021: 160). But in the case of alt-history games there is another layer of distancing: the subjects of memes and discussions are fictional (or fictionalized) political figures and factions that, albeit based on real-world politics, can always be described as 'just part of a game'.

The following article looks at examples of texts on the online discussion forums Reddit and Kiwi Farms that veer into reproducing and/or promoting far-right ideology. These forums (with some exceptions) are primarily meant to discuss their respective games and discourage or disallow addressing 'real-world' politics. However, given the fact that all mods in one way or another engage with real history and politics, discussions can often cross the line between in-game and real-world discourse despite the community guidelines. For example, when engaging with these communities it quickly becomes apparent that in-game roleplaying naturally extends into forum discussions and is often expressed as playful and/or post-ironic identification with or celebration of certain factions or figures, including fascist ones (see r/TheFireRisesMod 2025a, r/TNOmod 2025). Furthermore, political and historical discussions may, depending on the mod, start to fluctuate between discussion of in-game lore and real-world history, especially in discussions concerning 'historical accuracy'.

Given the topic of research this text contains potentially harmful and offensive images and quotations. These are presented in their original form to denote the presence of hate speech and far-right codes in the examined communities.

When in- and out-groups form, which seems natural in the case of a niche hobby like grand strategy games, a certain manner of speech and humour is perceived as a signifier of belonging. At the same time online community discourse is dynamic and, depending on the group, there is a push and pull between, on the one hand, the cultivation of toxic but identity-forming behaviour and, on the other hand, moderation of speech to protect the accessibility and credibility of the group and the hobby. This may, as will be presented later, lead to splintering of the community into more insular, toxic groups and more moderate, public-facing groups. An approach that acknowledges this dynamic nature of online communities is necessary to develop a nuanced understanding of the ways collective worldbuilding and memefication based on fictionalized version of real-world politics may facilitate the surreptitious dissemination of

revisionist and extreme political views by a minority within alt-history gaming communities.

After a description of the methodology and research corpus the analysis will be split into three parts. The first part (3.1) focuses on the risk of historical revisionism in alt-historical gaming communities, specifically in discussion about real-world Nazism. In the mods that take place around or before WWII Nazism as it is commonly understood is absent due to the altered historical trajectory of the game world. However, many figures and factions that relate to historical Nazism are still present but rid of the immediate negative association with exterminatory fascism. That part of the analysis will focus on how they are discussed (and memeified) within various communities.

The second part of the analysis (3.2.) homes in on the role of memes and the ways real-world fascism is addressed in the form of post-ironic identification within some parts of the communities. Taking the meme Speer Hoodie as a first example, this part continues from the last one but tries to focus more on how historical revisionism and post-ironic identification with fascist figures within the mods and their communities may spill over into real-world politics. This part also discusses meme subreddits like r/DSRfunny and r/NPPfunny that adopt transgressive forms of speech that resemble those of the alt-right. Lastly, this part will present the community on Kiwi Farms as an example of a small group of players that explicitly identify with fascist ideas and try to differentiate themselves from 'normie' places like Reddit. Generally, Kiwi Farms is far less restrictive on speech and, as the examples will show, this allows for certain players to express their political opinions in the form of far-right talking points and hate speech.

The last part of the analysis (3.3.) considers the question what may change in terms of community discourse when the mod's setting is transposed into the nearest past as in the case of *TFR*. What happens if instead of historical figures and factions, living politicians and contemporary geo-political struggles are fictionalized as part of a game? How (if at all) does the mod's community draw a distinction between fiction and reality? Furthermore, *TFR*'s game world includes elements of conspiracy theory that become part of memes within the community which are often indistinguishable from alt-right memes. This particularly begs the question whether the fictionalization of and post-ironic identification with in-game far-right ideologies and conspiracy theories may reproduce their real-world counterparts under the guise of humour and play.

1.2.1 Methodology

The aim of the research is to augment the discussions on the topic of the normalization of far-right ideology and hate speech in online communities by looking at the discourse on these mods both in relation to the games and their world-building¹ and to broader far-right communication practices, especially online. For this, I will follow a critical discourse analysis (CDA) approach (Fairclough 1995). CDA aims to point out the larger ideological context contained within a text or a collection of related texts. This approach seems appropriate, given the main research question of how far-right ideology may be expressed in community discourse surrounding alt-history mods. Here, CDA is used to analyse written text, in the form of forum posts, and images, particularly internet memes that can often be characterized as "multi-semiotic (ibid.: 4) meaning they often combine or contrast images and text to suggest or reinforce a specific reading of their respective elements. Furthermore, the analysis engages with each dimension of Fairclough's three-dimensional model of critical discourse analysis (ibid.: 133): the posts analysed here are simultaneously understood as text posted by individual users, as part of a discourse practice, i.e. the ways in which text is produced and distributed within the context of online forums like Reddit, and lastly as embedded within a larger social practice, including political and ideological contexts inside and outside of communities. The relationship between discourse practice and social practice is central to answering the research question, specifically how the pseudonymous and distributed structure of online forums, their implicit and explicit rules and their enforcement through moderation and social pressure may hinder or facilitate the dissemination of far-right ideological contents. In addition, the analysis pays attention to the notion of (subcultural) codes (Hall 1997) within alternative history game communities and as part of far-right / alt-right online discourse. Both, for different reasons, can obfuscate the meaning of text to an outsider. Consequently, texts that express far-right sentiments within the community texts are often (but not always) doubly en-

1 In practice, it is often challenging to cleanly separate discussions about story, world-building and game strategies, especially as a relative outsider. Not only has *HOI4* created its own game vernacular but, each mod features its own mechanics, characters, factions and ideologies. Therefore, I have tried to focus on overarching narratives and themes of each mod and their respective relationship to real history and avoid discussions that strictly deal with the gameplay.

coded or ambiguous. This text aims to highlight some of the moments where such codes are used to communicate to a specific in-group.

1.2.2 Corpus

The main sources for analysis are posts discussing alt-history games, including topics such as a) the game lore in relation to real history b) discussions on ideologies in-game in relation to real-world ideologies c) memes and jokes about political and historical elements of the mods and about the stereotypical player of these mods d) meta-discussions on the mods development and their respective communities. The texts that make up the main corpus were published on various subreddits. Additionally, the corpus includes a thread running since 2020 titled “/r/TNOMod and the Reddit *HOI4* modding community – When a subreddit for a video game mod turns to utter insanity” on the discussion forum Kiwi Farms that is directly linked to certain communities on Reddit. Descriptions on the Steam Workshop and on developer- and fan-run wikis serve as supplementary texts to contextualize aspects of each mod’s world-building.

Reddit is a user-moderated platform and highly public-facing. As such it is the first place where users will engage in discussions on various mods that are largely unrepresented on the official Paradox Games forums. Reddit’s structure is based on different user-generated and -moderated subreddits. There are subreddits for *HOI4* modding in general, for individual mods as well as secondary subreddits for memes based on various mods. The Kiwi Farms thread on *HOI4* mods is home to a smaller, more insular group that often distances itself from the public-facing subreddits. The thread initially gained traction as part of an “exodus” of users from Reddit that were dissatisfied with the moderation of the main subreddits, especially r/TNOMod, and critical of *The New Order*’s development team (see for example r/TNOModding 2023). Since then, the thread has also branched into discussing other mods like *The Fire Rises*. The mod subreddits have respective set of community guidelines that are enforced by moderators (disallowing things like, for example, off-topic discussions and hate speech). In contrast, Kiwi Farms is effectively unmoderated.

Initially, to identify relevant posts for the corpus, I have used web scraping with Python and PRAW. The subreddits scraped in the initial search were the following:

- r/kaiserredux, r/Kaiserreich, r/RedFloodMod, r/TheFireRisesMod, r/TNOMod, r/Fuhrerreich (subreddits on specific mods),
- r/hoi4, r/hoi4modding (generalized subreddits on *HOI4* and its mods)

- r/HOI4memes, r/NPPfunny (subreddits specifically focused on memes)

This algorithmic search identified posts featuring far-right and alt-right keywords and dog whistles, both classical antisemitic and fascist tropes (references to Rothschilds / George Soros, "New / Jew World Order", "Blood and Soil", etc.) and far right memes ("Kekistan", "Goyslop", "Groyper", triple brackets, etc.) supplemented by esoteric / mystical fascist terms ("Hyperborea", "Vril", "Thule", etc.). These search terms were derived from various texts discussing contemporary far-right dog whistles (Amadeu Antonio Stiftung 2021, Anti-Defamation League 2019, American Jewish Committee 2025, Crawford et al. 2020, Dafaure 2020, RAN 2021).

However, generating a corpus based solely on keyword searches has quickly revealed several problems, as a) some mods and their communities integrate these terms into their own world-building, especially when featuring factions inspired by esoteric fascist ideas and b) users will often refer to these terms ironically, as a form of role-play or as reported speech. Therefore, while the initial use of algorithmic keyword search has generated some research threads it required a lot of manual, informed selection of relevant posts as each occurrence of seemingly far-right speech requires specific interpretation and context to understand whether it is a genuine expression of personally held belief.

While most of the texts analysed here stem from this original algorithmic search, it also features additional texts that were either linked or otherwise referenced in the posts found via scraping. Some of the texts featured here referring to important events within a modding community's history that relate to the research question (e.g. the "Speer Hoodie" meme, the Kiwi Farm drama) were not found algorithmically but by reading the originally scraped posts and finding these points of interest.

As this text is not concerned with the identities of the posters all citations are denoted with the title of the subreddit. In the case of Kiwi Farms, as all citations are from one thread with over 15.000 posts, I will provide the post number as reference.

2.1 Revisions, Omissions: Alt-History in Grand Strategy Games?

In the following chapters the term 'alt-histories' will be employed to describe the world-building in the examined games. To some, this term might appear self-explanatory but carries certain connotations that should be addressed here. It was chosen over the more technical "counterfactual histories" (see

Grufstedt 2022), primarily because alt-history (or the longer “alternate histories”) is often used within the game communities themselves (see for example r/hoi4 2019). Especially given the research topic of this article “alt-histories” might also invoke another term: the alt-right. This connection is, for example, also made in the collected volume *Alt/Histories: Far-Right Revisionism and the End of History*, edited by Louie Dean Valencia-García. Valencia-García explicitly links the term to revisionist history and “alternative facts” (2020: 7).

These connotations are certainly relevant in looking at the way alt-histories work within the *HOI4* communities but naturally an alt-history game setting is distinct from a real-world claim to redefine historical facts. However, as will be described later, the line between two definitions of alt-history as either what-if-fantasy or revisionism might become blurry as the mods and their respective communities engage with fictionalized versions of real figures, ideologies and political factions. Regarding the question of historical depictions in *HOI4* and grand strategy games broadly, some research has been done, primarily from a historian perspective. Games like *HOI4* are either described as potential educational tools (Wainwright 2014: 592) or are examined in terms of their gamified rendering of history (Pennington 2022, Grufstedt 2022: 160). However, there has been no in-depth research on community-driven alt-history mods, and neither has there been research on community forums as a means of understanding gamified alt-history and world-building in practice. There are several studies on the far-right’s relationship to video games that briefly mention Paradox Interactive’s games and specifically *HOI4* and its mods. These studies are primarily part of the research on the cultural metapolitics of the far-right and do not go in depth on any of the games or communities.² *HOI4* mods can be understood as platforms for creating alt-histories with their respective subreddits being places to present, discuss and contrast these alt-histories. Howe-

2 For example, a report on Islamic and far-right terrorism by the Konrad-Adenauer-Stiftung mentions *HOI4* mods as potential recruitment tools for the far-right. In this report the author suggests that the fact that in some *HOI4* mods “20th century history is rewritten in favour of Nazi-Germany” (Wegener 2020: 4) could already be understood as a sign of “far-right propaganda” (ibid.). When looking at the actual discourse surrounding these games on the more prominent community forums these claims seem far-fetched, bordering on alarmism. While some fringe mods can certainly function as far-right texts in and of themselves, the mere existence of an alt-history where the “Nazis won”, a common trope found in many alt-history stories, does not imply that a game is propaganda.

ver, claiming that any of the successful games or communities have a unified ideological worldview ignores the dynamic nature of community discourse.

Other theorists refer to the more subtle ways in which parts of history are altered or omitted. The Bildungsstätte Anne Frank, for example, criticizes that the Holocaust is not represented in the game, reducing the ideology of the Third Reich to one of expansionism and war rather than extermination (Schnabel, Berendsen 2025).³ Additionally, discussions on the depiction of Nazi Germany happen within the *HOI4* community itself, for example around the expansion *Götterdämmerung* which some perceived as a glorification of Nazi Germany (Bernhardt 2024). These criticisms, rightfully, suggest that a game that engages primarily with historical topics may inherently provide its players with a distorted image of history. This article will focus especially on alt-history mods which do not strive for the same level of historical accuracy as *HOI4* but relate to history in a speculative manner. Here the notion of communal role-play and alt-historical worldbuilding become (at least) as relevant as comparisons to real historical events.

2.2 The Mods

As mentioned above, Paradox Interactive provides easy-to-use tools for modifying their grand strategy games, including *HOI4*. Consequently, it is impossible to cover all mods that have been created. These range from individual additional features to full conversions. All mods were written, developed and are maintained by non-commercial teams consisting of volunteers that often change over the course of a mod's lifespan. The following chapters will focus on some of the more popular mods, but others will be mentioned when appropriate. To give context it may be helpful to provide a brief description of each of these mods:

3 It should be noted here that the omission of the worst crimes of fascism might be less of an artistic decision and more due to Paradox Interactive's attempts to avoid legal problems in distributing their games in various countries, especially Germany (Wainwright 2014: 596).

Title	Setting / content	Time period in game	Year of release
<i>Kaiserreich: Legacy of the Weltkrieg</i>	The German empire won World War I, the Weimar Republic never existed, Nazism (as we understand it) never materialized, and the Holocaust never happened. Instead, World War II, or the “Second Weltkrieg” as it is called in the game, is initiated by a conflict between a fictional socialist France (Federation of the Communes of France) and the German empire.	1936 onwards (alternative WW II scenario)	2018
<i>Kaiserredux</i>	A fork of the original <i>Kaiserreich</i> , that includes a variety of extra or cut content for the game. While <i>Kaiserreich</i> ’s content has been revised continuously by its development team to assure a semblance of historical accuracy, <i>Kaiserredux</i> fully embraces the humorous and chaotic side of alt-history (see r/Kaiserreich 2024a).	1936 onwards (alternative WW II scenario)	2020
<i>The New Order: Last Days of Europe</i> (TNO)	Nazi Germany and Japan won World War II. Subsequently, the former has conquered Europe and the latter East Asia. The mod mixes World War II ideologies with Cold War elements. It also introduces elements of occult Nazism with groups like the Russian Faction “Aryan Brotherhood” that can become a quasi-religious group akin to Heinrich Himmler’s esoteric conceptions of the SS (The New Order Fandom Wiki: “Valery Yemelyanov” n.d.).	1960s (alternative Cold-War scenario)	2021

Title	Setting / content	Time period in game	Year of release
<i>Red Flood</i>	It is based on a fictional story in the universe of the <i>Kaiserreich</i> mod but evolved into its own alt-history scenario (r/Kaiserreich 2025). It portrays a world caught up in "a prolonged crisis of modernity" in which Russia won the war with Russia in 1905 and Germany became a Socialist Republic during WW I (Red Flood Wiki, n.d.). Compared to <i>Kaiserreich</i> , it focuses on more esoteric factions and chaotic geopolitics further removed from reality but often referencing real-world counterparts. ⁴	1930s (Alternative post-WWI period)	2022 (current version), 2020 (old version)
<i>The Fire Rises (TFR)</i>	The game takes place in an alternative contemporary setting at the beginning of the Corona Virus epidemic. It portrays a version of reality in which real-world geopolitical conflicts are amplified and addresses contemporary topics like social tensions within nations and the prospect of a multipolar world order (The Fire Rises Wiki n.d.). It includes fictional versions of real political figures like Donald Trump, Joseph Biden or the founder of the World Economic Forum (WEF) Klaus Schwab.	2020s	2024

Most of the *HOI4* mods reproduce the central gameplay of the original game that focuses on warfare and aggressive geopolitics (Pennington 2022: 3). In terms of worldbuilding this reflects an understanding of global politics as a sphere of conflict rather than diplomacy and cooperation. This view is especially relevant for mods like *TNO* or *TFR* that aim to transpose gameplay

4 For example, a central ideology in *Red Flood* is called "Accelerationism", a label that is more often associated with contemporary philosophers like Nick Land or Mark Fisher but in the world of *Red Flood* includes historical artistic movements like Futurism and Surrealism as well as pagan "Neo-Folkism" or "Technocracy" (Red Flood Fandom Wiki: "Ideologies#Accelerationism" n.d.).

beyond World War II. But even mods set around the World Wars like *Red Flood* often present a heightened version of history in which various conflicting factions fight for ideological hegemony. In comparison, as one user on Reddit suggests, “[o]ur timeline would probably come off ad [sic] really dull and unintesting [sic] in terms of geopolitics and ideologies.” (r/RedFloodMod 2022)

3.1 Good Guy National Socialism

Fig. 1: Meme adopting culture war tropes to make fun of the ‘wholesome’ Goebbels path in *Red Flood*.



While some mods like *Kaiserreich* are presented as serious ‘What-if?’-scenarios and consequently invite community discussions on historical accuracy, other mods are seen as fantasy games in which the player is enabled to play any scenario whether it is realistic or not. *Kaiserredux* distinguishes itself by allowing either so-called ‘schizo’ or ‘wholesome’ paths in the game. ‘Schizo’ paths are defined by the playing of fringe or fictional political move-

ments which are often authoritarian or violent (see r/Kaiserredux 2023a). Contrary, wholesome paths tend to focus on the success of equally fringe but more peaceful ideologies, often leftist, liberal or moderate in nature. However, in practice, this distinction is not always clear. The *Kaiserredux* submod *Roter Morgen* for example introduced a path that has Joseph Goebbels turn Germany into a "National Marxist" country, meaning a mixture between Nazism and Marxism akin to the real ideology of National Bolshevism (Steam Workshop: "Roter Morgen" 2024). Even after this path was removed from the main version of *Kaiserredux*, Goebbels name is occasionally brought up as a meme when discussing 'wholesome' paths in *Kaiserredux* and other mods (see r/Kaiserredux 2023b, Fig. 1).

Despite this difference of the role of alt-history both *Kaiserreich* and *Kaiserredux* present a world without Nazism and its consequences. Therefore, they circumvent the moral quandary of directly turning historical atrocities into play, which has already been a point of contention in discussions regarding the *HOI4* base game. While the player can still take the role of a conservative, imperialist or, in the case of *Roter Morgen*, alt-fascist German empire, the moral issues of racism and antisemitism largely remain unaddressed. The mod *Führerreich* presents an alternate timeline within the world of *Kaiserreich* in which Germany lost World War I but Nazism still did not appear as it did in real-world history. Instead, the German right follows an ideology called "Val-kism" which as a user puts it, is "like nazism without the extreme racism and genocide, so basically good guy national socialism. I think it's an interesting take." (r/Fuhrerreich 2018a) There are discussions on Reddit on the role of antisemitism in the world of *Kaiserreich* but seemingly there is no consensus on this topic (r/Kaiserreich 2019). As one redditor who claims to be an author for *Führerreich* writes regarding the role of Jews and antisemitism in the mod's lore:

Alright since I am the dude that wrote the majority of the lore for Germany I guess its time to answer this. I tried to keep the status of the Jews in Germany as vague as possible **to avoid certain things**. [...] Of course Jews are not liked and encouraged to leave Germany in favor of the jewish Commonwealth. (r/Fuhrerreich 2018b, Ed. A.N.)

These "certain things", meaning the issue of antisemitism and the ideological roots of the extermination under Nazism, are intentionally left blank which may help to make the nationalist parties in *Kaiserreich*, *Kaiserredux* and *Führerreich* seem like a more agreeable version of German imperialism or fascism res-

pectively.⁵ Not only are the Holocaust and other Nazi atrocities omitted but they are erased from the fictional timeline suggesting that far-right and proto-fascist groups would not be as bad if only real-world Hitler did not exist.⁶

It is of note that the idea of the ‘good fascists’, as opposed to the bad Nazis, resonates with certain ideas of the European New Right who broadly reject Nazism while praising thinkers of the so-called “Conservative Revolution” like Ernst Jünger, Carl Schmitt or Oswald Spengler (Sedgwick 2019: 15). This tactical move of separating proto-fascist ideas from actual historical fascism in Germany aims to rescue fascism from its taboo status and its association with exterminatory racism. For example, the Russian New Right author Aleksandr Dugin attempts a similar manoeuvre to separate fascism and racism in his book *The Fourth Political Theory* when he writes

Racism is what caused the collapse of National Socialism in the historical, geopolitical, and theoretical sense. This was not only a historical, but also a philosophical collapse. Racism is based on the belief in the innate objective superiority of one human race over another. It was racism, and not some other aspect of National Socialism, that brought about such consequences, leading to immeasurable suffering on both sides, as well as the collapse of Germany and the Axis powers, not to mention the destruction of the entire ideological project of the Third Way. (Dugin, Soral 2012: 33)

Red Flood uses anachronistic and fringe political ideologies as part of its alt-history that often resonate with an understanding of far-right ideologies that is much more contemporary. Factions like the National Bolsheviks or Accelerationists stand in for the dominant far-right factions in the real world (r/RedFloodMod 2023a). Like *TNO* it also includes allusions to esoteric, pagan ideologemes like the mythical continent of Hyperborea (r/RedFloodMod 2025b) and geopolitical imaginations like the concept of (right-wing) Eurasianism (r/RedFloodMod 2023b). While these movements are partially modeled

5 This can be seen as an extension of the problem of playing as Nazis that Schnabel/Berendsen (2025) and Wegener (2022) already pointed out for the *HOI4* base game.

6 In the gameplay of *Kaiserreich* Hitler is irrelevant and it is debated whether he died before the start of the game (see r/Kaiserreich 2024b, r/RedFloodMod 2025a). His pseudonymous counterpart “Adolf Polzl” is described as follows: “According to the Lore I read of Adolf Polzl, he is so because his mother did not die and therefore in the end he took his mother’s surname in addition to his being accepted into the School of Art. So... Technically he is a ‘Nice Hitler’ so to speak.” (r/RedFloodMod 2025a)

after real political ideologies, they appear to be made distinct from their real-world counterparts. In that sense no singular ideology in the game seems to be representative of real-world fascism.⁷ In its community guidelines r/RedFloodMod disallows discussions about real world politics but given the appropriation of existing (fringe) ideologies and figures the borders between in-game discussions and real-life political discussions are not always clear. For example, in a discussion on the accelerationist factions in *Red Flood* several users compare it to and differentiate it from its real-world counterpart:

[W]hen I think IRL Accelerationism I picture Alexander Dugin in Russia. Look into him and his ideas for some IRL Accelerationism that could actually have a lasting impact. (r/RedFloodMod 2025c)

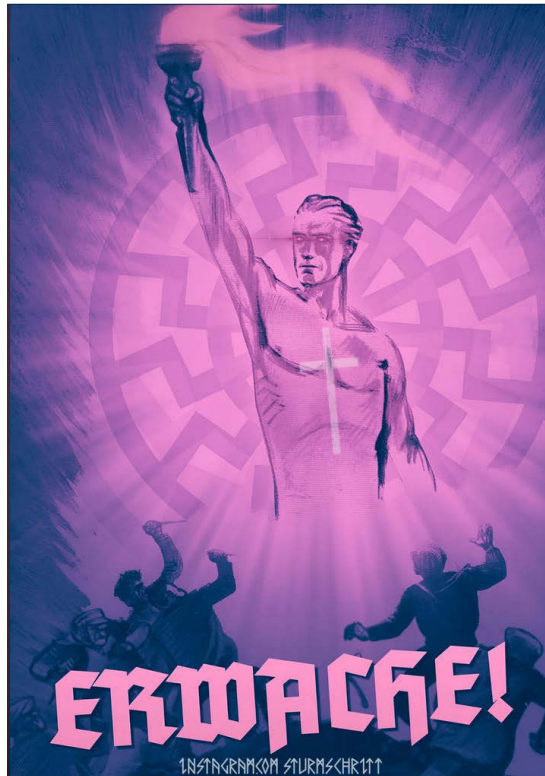
[R]eal acceleration irl is usually associated with some pretty evil groups like O9A or Atomwaffen. In red flood it really depends, there are accelerationists who you could say would probably be pretty good to live under like Guido Keller Fiume but there are also accelerationists who living under would be an absolute nightmare (ibid.).

Generally, the in-game representation of (pseudo-)fascism and consequently the discussions within communities are often kept vague in their historical references. Especially more 'schizo' mods like *Kaiserredux* or *Red Flood* seem more interested in creating far-right factions based on esoteric and occult ideologies, mixing and matching elements to create simulacra of fascism. But even the more serious alt-history mods tend to eschew the role of historical Nazis in favour of various flavours of right-wing factions that represent different aspects of real-world fascism. The lack of a clear in-game referent for (exterminatory) fascism and Nazism may impede criticism of posts and memes that reproduce reactionary views by referencing in-game factions and figures, some of which are based on real-world fascists. The notion that there are 'softer' versions of fascism and the ironic stylization of 'schizo' campaigns that let players act as fictionalized versions of real Nazis and fascists may serve to misrepresent and trivialize historical atrocities, muddle political discourse and potentially allow far-right actors to join the community and intersperse their real

7 As one poster on *Red Flood* subreddit writes: "Fascism basically does not exist as a term in Red Flood beyond San Marion. Most real life fascists are split between Accelerationism, Despotism, Reactionism and sometimes Vanguard Socialism." (r/RedFloodMod 2021)

world political ideas into the community discourse under the pretext of jokes or post-ironic role-play (see Fig. 2).

Fig. 2: Graphic in a 'fashwave'⁸ style featuring the popular far-right symbol of the Black Sun posted on the TNO meme subreddit r/NPPfunny. The post is titled "Clerical Fascism/ Burgundian System unity I guess." referencing the fictional TNO faction Order-State of Burgundy.



⁸ Fashwave was a style of online graphic design from the late 2010s / early 2020s that combined with far-right texts and visual codes with 'retro' graphic design, especially referencing 1980s popular media and graphics.

3.2 The Speer Hoodie, Alt-History Memes and Kiwi Farms

According to a primary member of the development team, Jakob Zaborowski, *TNO* was conceived as a dystopian vision of a fascist victory (del Callar 2020). However, a small group of players latched onto the stylization of fascism in the game as an opportunity for 'edgy' role-play and memes. This appropriation by players celebrating far-right ideologies, in addition to cases of bullying and doxing in relation to the fork *TNOredux*, has garnered *TNO* and its community the reputation of being controversial (r/hoi4 2024).

One of the most infamous moments in the history of the *TNO* community that blurred the line between role-play and political expression was the so-called 'Speer Hoodie' meme posted on October 1st, 2020, on r/TNOmod. A user on Reddit posted a picture of a hoodie that featured the Nazi architect Albert Speer⁹ in a graphical style that resembled online memes (see Fig. 3). The meme suggests that wearing the face of a Nazi is acceptable as a form of post-ironic adoration. The Speer Hoodie meme was immediately met with push-back from the community, suggesting that an implicit line between acceptable and unacceptable role-play had been crossed. This is exemplified by a response of one of the subreddit's moderators:

Holy shit. I have a moral obligation to tell people to not do this kind of shit. Don't get fucking hoodies with an actual fucking Nazi's face on it, oh my god. People might actually think you're a fucking nazi. (TNOpedia: "Memes#Speer_Hoodie" 2024)

Several users later expanded on the meme, and it has become a symbol of a more 'edgy' time of the *TNO* community where users were not afraid to push the boundaries of acceptable speech (r/TNOmod 2023). As such the meme, on the one hand, exemplifies the negotiation processes around political speech within these communities and, on the other hand, showcases post-irony and humorous transgression as part of a process of normalizing the posting of far-right content.

9 In *TNO* Speer is one of four possible successors of Hitler and is described as a "reformist" of the Nazi party (The New Order Fandom Wiki: "Albert Speer" n.d). As a fictional (meme) figure he stands for an alternative, agreeable version of Nazism that celebrates the German Reich as a world power while distancing itself from orthodox Hitlerism.

Fig. 3, Fig. 4: Image of the Speer Hoodie originally posted on r/TNOmod in a post titled “My Speer Hoodie arrived today” and a person wearing a variation of the Speer Hoodie shaking hands with the Polish far-right politician Janusz Korwin-Mikke, posted on r/NPPfunny under the title “real”.



This ultimately led to the TNO subreddit disassociating from meme subreddits like r/DSRfunny and r/NPPfunny and introducing more content moderation on the mod subreddits. Nowadays the rules on r/TNOmod explicitly disallow real-world political discussions as well as ‘excessive LARPing’ (Live Action Role Playing). However, memes referencing the Speer Hoodie and other memes that (post-ironically) celebrate Nazi figures in TNO could still be found on ‘shitpost’ subreddits like r/hoi4memes, r/NPPfunny or the now defunct r/DSRfunny (see Fig. 4, 5, 6).

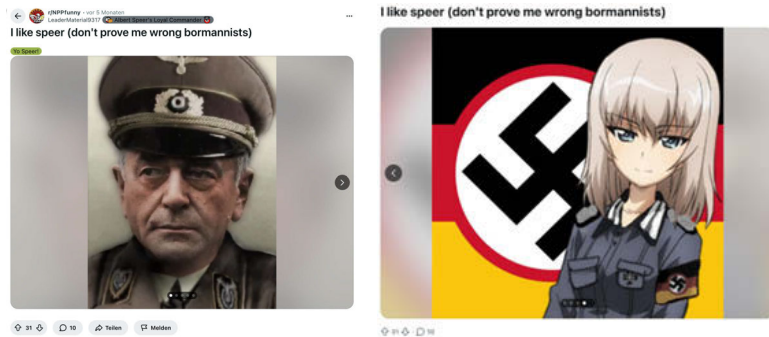
These meme-focused subreddits adopt a language and aesthetics that is reminiscent of websites like 4chan¹⁰ with memes often being categorized as “brainrot”, “shitposts” or “schizo” (see for example: /NPPfunny 2023, r/

10 4chan.org is a website consisting of several image boards / discussion forums created in 2003. 4chan is generally associated with cultivating (the more extreme forms of) contemporary meme culture and, especially on the political forum /pol/ forum, with the idea of the alt-right. For example, many of the memes associated with that movement (Pepe the Frog, Wojaks, Kekistan) and early Trumpism / MAGA were first popularized on 4chan.

TNOmod 2023a, r/Kaiserredux 2023a). This 'edgy' humour coupled with post-ironic references to Nazism and fascism, here justified by their relevance to the game, mirrors the visual language of early alt-right memes in which crassness and esoteric references often superseded any direct political messaging (Askanius 2021, Price 2025). Just as on other platforms that feature transgressive memes humour creates plausible deniability and opens a larger space of what is deemed acceptable speech:

This ironic distance is an attempt to shield the joke teller from accusations of racism, deflecting critics as humorless, and enabling individuals to say things that would otherwise be considered inappropriate (Pérez 2013). It allows the joke teller to test the waters of the acceptability of exclusionary and supremacist sentiments in a manner that presents as unserious (Pérez 2022). At the same time, it serves to normalize such sentiments by signaling to audiences that such expressions are acceptable [...]. (McSwiney, Sengul 2023)

Fig. 5, Fig. 6: Series of images featuring the Nazi architect Albert Speer and the character Erika from the anime Girls und Panzer in front of a Hakenkreuz flag posted on r/NPPfunny.



In the case of memes in the context of alt-history mods this effect is then doubled because not only is it 'just a joke' but it also all happens within a context of ironic (or 'schizo') role-playing. Thus, whoever takes the Speer Hoodie seriously as an act of political expression is not only humorless but a 'tourist', i.e. it signals that they do not understand the context within the (meta-)narratives of the game and its community. This again mirrors a discourse strategy of the alt-right:

It is the production of the in-group through laughing at others and using jokes that others will not understand that allows us to see online far-right communities engaging in homesteading practices. Kinnvall (2019) has written on the role of 'homesteading,' whereby certain spaces are actively bordered off for the sake of in-group wellbeing, and this production of a specific bounded home is based on a transgressive humor that challenges the acceptable. (Price 2025: 5)

While many users on r/TNOmod reject the promotion of far-right ideologies and share the designer's understanding of its world as dystopian, the game seems to, at one point, have attracted a group of players that genuinely identified with playing fascists. According to users on r/TNOmod the aforementioned doxing incident led to a split not only in the development team but in the community and "the Kiwi Farms drama drove out most of the weirdos." (r/hoi4modding 2023) Comparing the posts on Reddit to those on the effectively unmoderated thread on Kiwi Farms confirms that far-right hate speech in the context of TNO and alt-history mods is very much alive there. The difference in the proliferation and (lack of) moderation of hate speech between Reddit and Kiwi Farms is readily apparent and suggests that it is true that different parts of community have split between these different forums according to the proclivity for far-right ideas and hate speech. The users on Kiwi Farms often employ internet far-right and conspiracy codes creating an esoteric and insular form of discourse amongst a small in-group of radicalized individuals. For example, two Kiwi Farms users reference a discussion on the (more mainstream) Reddit forums as follows:

[User A] Why are these retards obsessed with fantasy nazis again? Isn't that just a special breed of glownigger?¹¹ Because this whole subreddit community sounds like some astroturfing.¹²

[User B] They're obsessed with exacting bizarrely specific revenge fantasies against them. It's very Jewish behavior. (Kiwi Farms #6,037)

11 "Glownigger" or "glowie" are alt-right terms, originating on 4chan, describing undercover government agents that infiltrate far-right online spaces (Khazan 2021).

12 Astroturfing describes the idea that certain groups or political movements that appear as genuine grassroots movements are actually created and/or sponsored by actors in places of power to fabricate a fake notion of public will (Institute for Strategic Dialogue n.d.).

Posts on Kiwi Farms, due to the apparent lack of rules and moderation, are also more likely to go on tangents away from in-game topics and instead discuss their political opinions directly:

Lesson off [sic] the Day: Learn from the OG sources (Mein Kampf) Rather than drink the Kool Aid of people who want the age off Consent lower (Daily Stormer)¹³,specailly if you are a retar.[sic]

Also is not modding that corrupts man,is that modding shows the true colors off everyone because HOI 4 Modding is a good mirror/screen into the minds off it's developers. (Kiwi Farms #14,307)

Nowadays *TFR* appears to be the more respected mod by the users on Kiwi Farms who consider *TNO* and its community to be too politically correct. In comparison to *TFR*, *TNO* is characterized as liberal with Kiwi Farms users giving it derogatory nicknames like "TrooNO"¹⁴ (Kiwi Farms #13,876, #14,958, #11,654). At the same time users discuss the notion of *TNO* as a "deradicalization tool", supposedly promoted by the mods developers, contrasting it to the idea that in reality *TNO* has helped to popularize far-right terms and ideas:

I've always found it curious that *TNO* was made as "le nazis... le bad!" neoliberal grimderp slop project but it has only served to popularize esoteric hitlerism and make shit like the black sun and hyperborea common knowledge. In a way it's funny on both ends, shitlibs are so bad at creating propaganda that they can't even make the crumbling, bloated corpse of nazi germany + a number of retarded neo-nazi warlod pedophile rapist russkies look dumb, and the black sun larper fags only know about this because of a hoi4 mod made by a bunch of tranny liberals. Peak comedy. (Kiwi Farms #11,850)

[A]lso tno has "radicalized" more people than any other mod despite the tno devnigger's claims that it's a deradicalization tool. For example the nigger shooter, HonoraryYakubian said he was inspired by tno (could've

13 The Daily Stormer is a far-right online platform associated with the alt-right.

14 This name refers to the transphobic internet slang 'Troon' for trans people. While this may be in reference to parts of the community and the development team identifying as trans, it is arguably meant more as a generalized insult to differentiate between the socially conservative values of far-right users on Kiwi Farms and the supposed liberal values of the *TNO* development team and (the majority) of the community on *r/TNOmod*.

been trolling tho, but still funny that the police probably had to research what TNO was. (Kiwi Farms #13,893)

That's the best part of TroonO. It was supposed to be a "deradicalization tool" but it instead kickstarted the popularity of Hyperborea memes back in 2020 to 2021, and now it's mainstream. (Kiwi Farms #14,963)

Thus, the discourse on Kiwi Farms, at least in part, paints a contrasting picture to those users on Reddit that argue that memes and fascist roleplay in TNO and other mods are inherently apolitical.

On the one hand the reactions to the Speer Hoodie incident and similar boundary-pushing memes show an ability of (parts of) the TNO community to moderate itself by excluding people who go 'too far' in their political speech, be it genuinely or post-ironically, which subsequently makes those people join more 'edgy' and insular sub-communities like Kiwi Farms or r/NPPfunny. On the other hand, this does not mean that r/TNOmod is entirely free of controversial humour and references to alt-right meme aesthetics that walk the line between roleplay and political expression. The identification with 'schizoposting' can be understood as a constitutive element of communities for mods like TNO or *Kaiserredux* that themselves are often exaggerated and, to an extent, make light of (pseudo-)historical violence. Even on the moderated subreddits this humour in the style of 4chan can occasionally lead to reproducing elements of alt-right speech and aesthetics. Thus, regardless of the subreddits rules, it appears impossible to fully separate the mixture of meme culture and discussions of alt-history from real-world political speech. Additionally, as we have seen, a radical far-right minority of players persists on alternative forums and may 'come back' (if they ever really left) and latch onto the ambiguous humour on the more public-facing subreddits to subtly promote their own worldview, as can be seen in some examples on r/NPPfunny presented here (see for example Fig. 2, 4, 5, 6). While the danger of far-right agitators influencing public online discourse is omnipresent in our times, these communities that center around (political) ambiguity and dark humour can be especially prone to creating a platform for far-right agitators (see Reinhold, Prinz in this publication).¹⁵

¹⁵ As we have already seen in the case of the alt-right and early MAGA movement that utilized "Reddit or 4chan where the borders between irony and earnestness are precarious." (Dafaure 2020: 4)

3.3 From Alt-History to 'Schizo'-Present

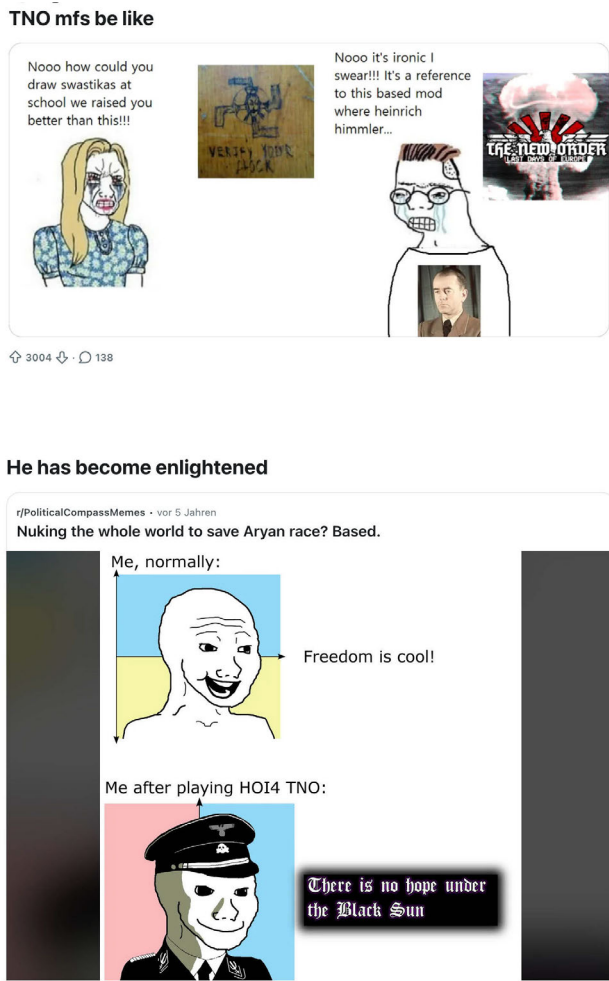
So far, the article has primarily focused on mods that reimagine historical time periods. These mods have developed their respective worlds by modulating historical conflicts (World War I, World War II, the Cold War) based on the idea that small changes to history would lead to radically different outcomes. Consequently, discussions in these communities sometimes address real world events but discussions center primarily on the idea of 'historical accuracy / probability' by speculatively comparing real history to the game world and discussing the viability of certain alt-historical timelines. A few users, especially on Kiwi Farms, may link (alt-)historical fascism to contemporary far-right ideas but contemporary political topics largely do not factor into the discussions on the main subreddits of *TNO*, *Kaiserreich* or *Red Flood*.¹⁶ While there is a real risk of these discussions introducing moments of historical revisionism, especially in cases where real-world figures exist as variations in the game (e.g. Goebbels, Speer, Himmler), members of the community are quick to point out that the in-game characters and the memes surrounding them should not be taken seriously (see r/TNOmod 2022). Furthermore, the idea of becoming ideologized through alt-historical *HOI4* mods is itself often understood as a post-ironic meme¹⁷ (see Fig. 7, 8). This disavowal of any 'real' political content or effect creates an ambiguous space where certain speech that is otherwise considered unacceptable is normalized through ironic distance.

This relationship between in-game role-play, community discourse and real-world politics becomes much more complicated in the context of *The Fire Rises*, which is set in an alternative present timeline. Its world centers around multipolar geopolitics and a fictional Second American Civil War. It employs a variety of real-world events and conflicts as part of its world-building. In that sense it takes the idea of alt-history and transposes it to the present, creating a slightly modulated version of our reality that is fictional yet not entirely distinct.

16 Even where real-world political (fringe) ideologies are invoked, they are usually described as separate from their in-game counterparts (see r/RedFloodMod 2025c).

17 Conversely, members of meme communities like r/NPPfunny, r/KRCJ (Kaiserreich circlejerk) or, to an extent, r/HOI4memes, characterize the main subreddits as places for humourless 'normies' that are afraid to commit to transgressive role-play and speech (see r/HOI4memes 2024).

Fig. 7, Fig. 8: Post-ironic memes suggesting that TNO introduces players to far-right ideas. Both posted on TNOmod in 2020.



Just as alt-historical mods fictionalize historical figures, *TFR* fictionalizes various living people, amongst them many far-right public figures. These include, for example, white nationalists like Thomas Rousseau, Brandon Russell, Jordan Jereb, or Aleksandr Dugin who are also the frequent subjects of memes on *r/TheFireRisesMod* (see *r/TheFireRisesMod* 2025b, *r/TheFireRisesMod*

2025c, r/TheFireRisesMod 2025d, r/HOI4memes 2025). The community also makes post-ironic references to contemporary right-wing conspiracy theories like QAnon¹⁸ (see r/TheFireRisesMod 2025e) and the "Eat the Bugs" conspiracy (Moore 2025), usually stylized as "Eat ze bugs" (r/TheFireRisesMod 2025f). The latter is part of a common in-joke that centers around Klaus Schwab, founder of the WEF, and (potential) in-game leader of the Ultraglobalist faction. The game takes place during the Covid Pandemic and includes references to the "Great Reset" (r/TheFireRisesMod 2025g), a program by the WEF that is often the subject of conspiracy theories (Robinson et al. 2021). It can be connected to similar conspiratory narratives like the New World Order, the Great Replacement and the existence of a group of evil 'globalists'. These narrative tropes featuring a minority group directing political and social developments from the shadows to realize a secret agenda perpetuate xenophobic and specifically antisemitic stereotypes (Römer, Stumpf 2019: 141). In the context of *TFR* and its community discourse they enable a perpetuation of such stereotypes under the guise of 'in-game' jokes. Given its contemporary setting, it is no surprise that this often includes tropes of "culture war" (see Liemann in this volume).

These examples show how the *TFR* communities on Reddit post-ironically reproduce contemporary far-right narratives and aesthetics. This can be productively compared to the analysis of Valencia-García (2020: 14) who points out the structural link between alt-histories and conspiracy theory. In the context of *TFR* this suggests a qualitative change in how communities engage in-world discussions in relation to real-world politics. What happens for example when real-world political figures acknowledge their portrayal in the game, and this is discussed or even celebrated in the community as an acknowledgment of themselves (The Fire Rises Wiki: "Acknowledgment of The Fire Rises by People or Organisations featured in the Mod" n.d.)? In one post a user shared a link to a video by Jordan Jereb, leader of the US white supremacy group Republic of Florida Militia and fictional leader of the Christian faction of the far-right Atomwaffen Division in *TFR*, in a thread titled "Jordan jereb endorses tfr" (r/TheFireRisesMod 2025h). In the same thread a post on X by Jereb is quoted:

18 QAnon was a loose online group that formed around 2017 and was based around a supposed government insider under the pseudonym Q that shared supposed secrets about government conspiracies (especially in the Democratic party) involving Satanism and paedophilia (Bleakley 2023).

I am going to be making a video about The Fire Rises HOI4 Mod. I'm a character in this game, And I want to talk about why I think the game is significant. (ibid.)

Here it becomes obvious how the *TFR* community (approvingly) engages with real-world far-right actors through the filter of a fictional game world. Far-right figures and factions are simultaneously understood as playable pieces of an imagined game world and acknowledged as real-world actors. And, conversely, the community and game are understood by far-right actors like Jereb as sites for metapolitics to popularize, normalize and aestheticize far-right ideas.

As an outsider, allusions to far-right ideology are certainly more legible in the case of *TFR*. Instead of referencing early 20th century rhetoric and aesthetics like other mod communities, the *TFR* community is engaging with the aesthetics of late-stage capitalism and meme culture in a self-referential manner. This goes so far as to allowing for hyperreal versions of memes and notions of culture war, which themselves are largely defined by self-referentiality and (post-)irony as part of the game discourse. In other words, while a member of the *Kaiserredux* community might role-play as a historical far-right person **via** the internet, *TFR* allows users to role-play as a contemporary far-right person **on** the internet, becoming effectively indistinguishable from a far-right 4chan post. This discourse is made possible by employing two layers of deferral when faced with the critique of reproducing far-right political speech or aesthetics: Not only is it a 'schizo' joke but it is a joke within the context of a fantasy version of reality. In that sense even a post that is indistinguishable from far-right political speech may appear impervious to critique.

4. Conclusion

The research has shown that far-right sentiments are present in alt-history gaming communities. The extent to which this becomes prevalent depends on both the mods themselves, i.e. how they allow for identification and glorification of far-right factions and figures, and on the discourse practices in each community, including rules and moderation.

It has also become clear that these discourse practices are a dynamic process that is shaped by certain events in the history of communities. For example, in the case of *TNO* more transgressive moments in the past like the Speer Hoodie meme have led to stricter rules regarding real-world po-

litical discussions and memes. The shutdown of r/DSRfunny where memes in the style of Speer Hoodie were posted regularly also suggests that more extreme forms of speech will eventually become subject to regulation. Generally, explicitly far-right posts are met with pushback by the community on subreddits like r/TNOMod, r/Kaiserreich and r/RedFloodMod. This does however not mean that all kinds of transgressive post have disappeared from the more mainstream communities. Not only have newer meme subreddit like r/NPPfunny adopted a style of posting similar to r/DSRfunny but transgressive humour that plays with irony and ambiguity is still present in the main subreddits of alt-history mods. This however, as I hope some of the quotations have shown, is qualitatively different from a forum like Kiwi Farms where a group of players, supposedly originating from Reddit, openly express far-right ideas and regularly post hate speech. This happens to a lesser extent than on Kiwi Farms but still notably so the *TFR* community has normalized a kind of speech and memes that often resembles those of the alt-right and 4chan. Reading all these aspects together it becomes apparent that alt-history games and their communities are a place where far-right metapolitics occur, often subtly but sometimes explicitly. While a website like Kiwi Farms platforms hate speech in a way that is blatantly obvious, the various subreddits that often walk the line between being outward-facing and insular present us with more ambiguous forms of political speech. On the one hand, the communities on Reddit (with some exceptions) tend to moderate themselves to seemingly discourage hate speech and (fascist) LARPing. On the other hand, it is that claim of moderation that marks tendentious forms of speech as acceptable when they are not removed or marked by moderators. The fact that it is often difficult to accurately differentiate far-right opinions from ironic roleplay underlines the ways ambiguity may work as a political strategy for the far-right online. The communities of alt-history mods become sites of toxic behaviour not primarily, albeit possibly, through direct forms of attacking others but through normalizing 'playful' versions of reactionary thought, hate speech and conspiracy theories that always provides an out for those that express them. Often these forms of post-ironic and transgressive expression become signifiers of belonging to the community. Deciphering layers of irony and responding correctly is key to becoming part of alt-history mod communities that venerate expertise in the specific lore and in-jokes of the games and the communities themselves. In its most extreme form this may create a self-reinforcing toxic structure in which affirmative notions of historical and (as in the case of *TFR*) contemporary far-right ideologies are suspended in a hyperreal

state of play. The fact that sites like 4chan have become the breeding ground of the alt-right as a political movement has shown how humour and a sense of community online can draw (especially young) people towards far-right ideas or at least reinforce a socially conservative worldview. But whereas 4chan's transgressive humour hardly obscured the real political sentiment underneath, the communities discussed here present a more ambiguous relationship to real-world politics via the notion of fiction and play. It is not only humour but also the claim of fiction and play that allows certain forms of transgressive speech to become acceptable. The findings of this research have shown that far-right ideas and hate speech do occur in parts of these communities and that moderation (where it exists) only serves to rein in the most explicit cases.

Bibliography

- Amadeu Antonio Stiftung (2021): Deconstruct Antisemitism. Antisemitische Codes und Metaphern erkennen. Berlin.
- American Jewish Committee (2025): Translate Hate: Stopping Antisemitism Starts With Understanding It.
- Anti-Defamation League (2019): Hate on Display. Hate Symbols Database. New York.
- Askanius, Tina (2021): "On Frogs, Monkeys, and Execution Memes: Exploring the Humor-Hate Nexus at the Intersection of Neo-Nazi and Alt-Right Movements in Sweden." In: *Television & New Media*. 22, 2. P. 147–165.
- Bernhardt, Alexander (2024): „Hier wird Grenze überschritten“ – Wichtiger History-YouTuber verweigert neuem *Hearts of Iron IV* DLC die Coverage. In: Gamestar. URL: <https://www.gamestar.de/artikel/hearts-of-iron-4-dlc-steinwallen,3422968.html> [accessed 14.03.26].
- Bleakley, Paul (2023): "Panic, Pizza and Mainstreaming the Alt-Right: A Social Media Analysis of Pizzagate and the Rise of the QAnon Conspiracy." In: *Current Sociology*. 71, 3. P. 509–525.
- Crawford, Blythe; Keen, Florence; Suarez de-Tangil, Guillermo (2020): Memetic Irony and the Promotion of Violence within Chan Cultures. Lancaster: Centre for Research and Evidence on Security Threats.
- Dafaure, Maxime (2020): The "Great Meme War:" the Alt-Right and its Multifarious Enemies. In : *Angles. New Perspectives on the Anglophone World*. 10. P. 1–28. URL: <http://journals.openedition.org/angles/369> [accessed 14.03.26].

- del Callar, Carlos Miguel (2020): The Videogame That Takes Us Inside the Hell of a Nazi-Run World. In: Quillette. URL: <https://quillette.com/2020/11/26/the-e-videogame-that-takes-us-inside-the-hell-of-a-nazi-run-world/> [accessed: 14.03.26].
- Dugin, Alexander; Soral, Alain (2012): *The Fourth Political Theory*. London: Ark-tos Media Ltd.
- Grufstedt, Ylva (2022): *Shaping the Past: Counterfactual History and Game Design Practice in Digital Strategy Games*. Berlin: De Gruyter.
- Hall, Stuart (1997): The Work of Representation. In: *Representation: Cultural Representations and Signifying Practices*. Ed. by Stuart Hall. London, Thousand Oaks, New Delhi: Sage. P. 13–74.
- Institute for Strategic Dialogue. (n.d.): "Astroturfing." In: *ISD glossary*. URL: <https://www.isdglobal.org/isd-glossary/> [accessed: 14.03.26].
- Khazan, Olga (2021): "The Far Right's Fear of 'Glowies'". In: *The Atlantic*. URL: <https://www.theatlantic.com/politics/archive/2021/01/what-glowies-mean-online-spies/617717/> [accessed 14.03.26].
- Loban, Rhett (2021): The Transformation from Physical Wargames to Grand Strategy Video Games, and the Opportunities for Deep and Efficient Historical Wargaming Experiences. In: *Digital Culture & Education*. 13, 1. P. 81–107.
- McSwiney, Jordan; Sengul, Kurt (2023): Humor, Ridicule, and the Far Right: Mainstreaming Exclusion Through Online Animation. In: *Television & New Media*. 25, 4. P. 315–333.
- Moore, Deidre (2025): 'I will not eat the bugs': examining a right-wing narrative about scarcity and insect consumption. In: *The Conversation*. URL: <https://theconversation.com/i-will-not-eat-the-bugs-examining-a-right-wing-narrative-about-scarcity-and-insect-consumption-254112> [accessed 14.03.26].
- Pennington, Michael J. (2022): *Authentic-lite rhetoric: the curation of historical interpretations in 'Hearts of Iron IV'*. PhD thesis. Bath: Bath Spa University.
- Price, Charlie W. (2025): "Ketchup Is the Whitest Sauce: Memes as Silly/Serious Bordered Spaces in International Relations." In: *International Political Sociology*. 19, 3. URL: <https://doi.org/10.1093/ips/olafo26> [accessed 14.03.26].
- Robinson, Olga; Sardarizadeh, Shayan; Goodman, Jack; Giles, Christopher; Williams, Hugo (2021): "What is the Great Reset – and how did it get hijacked by conspiracy theories?" InBBC. URL: <https://www.bbc.com/news/blogs-trending-57532368> [accessed 14.03.26].

- Römer, David; Stumpf, Sören (2019): „Der Große Austausch ist kein Mythos, er ist bittere Realität. Populismus und Verschwörungstheorien aus linguistischer Perspektive.“ In: Osnabrücker Beiträge zur Sprachtheorie. 95. P. 129–158.
- Schnabel, Deborah; Berendsen, Eva (ed.) (2025): *Der Holocaust als Meme*. Frankfurt: Bildungsstätte Anne Frank.
- Sedgwick, Mark J. (2019): *Key Thinkers of the Radical Right: Behind the New Threat to Liberal Democracy*. Oxford: Oxford University Press.
- Valencia-García, Louie Dean (2020): *Far Right Revisionism and the End of History*. In: (ed.): *Alt/Histories: Far-Right Revisionism and the End of History*. Ed. by Louie Dean Valencia-García. Abingdon: Routledge.
- Wainwright, A. Martin (2014): *Teaching Historical Theory through Video Games*. In: *The History Teacher* (Society for History Education). 47, 4. P. 579–612.
- Wegener, Friederike (2020): *Verblüffende Ähnlichkeiten unter Gegnern: Gemeinsamkeiten zwischen Rechtsextremismus und Islamismus; wie Sicherheitsbehörden von einer verschränkten Perspektive profitieren können*. In: *Analysen & Argumente*. 404. Konrad Adenauer Stiftung. P. 2–9.

Sources from research corpus

- Kaiserreich Official Wiki (n.d.): *Progress Report*. In: Kaiserreich Official Wiki. URL: https://kaiserreich.miraheze.org/Progress_Reports [accessed 14.03.26].
- Kiwi Farms (2025): *“/r/TNOMod and the Reddit HOI4 modding community – When a subreddit for a video game mod turns to utter insanity”*. In: Kiwi Farms. URL: <https://kiwifarms.st/threads/r-tnomod-and-the-reddit-hoi4-modding-community.80516/page-729#post-22795303> [accessed 14.03.26].
- r/Fuhrerreich (2018a): *What is Valkism?* In: Reddit. URL: https://www.reddit.com/r/Fuhrerreich/comments/9je66j/what_is_valkism/ [accessed 14.03.26].
- r/Fuhrerreich (2018b): *What's the exact status of Judaism in Germany?* In: Reddit. URL: https://www.reddit.com/r/Fuhrerreich/comments/92ouy6/what_s_the_exact_status_of_judaism_in_germany/ [accessed 14.03.26].

- r/HOI4memes (2024): Average TNO player. In: Reddit. URL: https://www.reddit.com/r/HOI4memes/comments/1fi5kzn/average_tno_player/ [accessed 14.03.26].
- r/HOI4memes (2025): jordan jereb endorses tfr. In: Reddit. URL: https://www.reddit.com/r/HOI4memes/comments/1nkm3iz/jordan_jereb_endorses_tfr/ [accessed 14.03.26].
- r/hoi4 (2019): hueg alternate history mod overview. In: Reddit. URL: https://www.reddit.com/r/hoi4/comments/dd6b9b/hueg_alternatehistory_mod_overview/ [accessed 14.03.26].
- r/hoi4 (2021): What happened to TNO and TNO Redux? In: Reddit. URL: https://www.reddit.com/r/hoi4/comments/nokooo/what_happened_to_tno_and_tno_redux/ [accessed 14.03.26].
- r/hoi4 (2024): What is the biggest hoi4 controversy? In: Reddit. URL: https://www.reddit.com/r/hoi4/comments/1fwxtue/what_is_the_biggest_hoi4_controversy/ [accessed 14.03.26].
- r/hoi4modding (2023): TNO kinda ruined HOI4 modding. In: Reddit. URL: https://www.reddit.com/r/hoi4modding/comments/16izbzb/tno_kind_a_ruined_hoi4_modding/ [accessed 14.03.26].
- r/Kaiserreich (2018): The Virgin Normie VS. The Chad Kaiserreich Player. In: Reddit. URL: https://www.reddit.com/r/Kaiserreich/comments/7ur201/the_virgin_normie_vs_the_chad_kaiserreich_player/ [accessed 14.03.26].
- r/Kaiserreich (2019): Jews in KR. In: Reddit. URL: https://www.reddit.com/r/Kaiserreich/comments/dg8vck/jews_in_kr/ [accessed: 14.03.26].
- r/Kaiserreich (2024a): Difference Between Kaiserreich and Kaiserredux. In: Reddit. URL: https://www.reddit.com/r/Kaiserreich/comments/1bel4ss/difference_between_kaiserreich_and_kaiserredux/ [accessed 14.03.26].
- r/Kaiserreich (2024b): Are fdr and Hitler ever mentioned in kaiserreich. In: Reddit. URL: https://www.reddit.com/r/Kaiserreich/comments/1dm43yp/are_fdr_and_hitler_ever_mentioned_in_kaiserreich/ [accessed 14.03.26].
- r/Kaiserreich (2025): Red Flood? In: Reddit. URL: https://www.reddit.com/r/Kaiserreich/comments/1ly2b8y/red_flood/ [accessed 14.03.26].
- r/Kaiserredux (2023a): What are good schizo paths? In: Reddit. URL: https://www.reddit.com/r/kaiserredux/comments/13kau65/what_are_good_schizo_paths/ [accessed 14.03.26].
- r/Kaiserredux (2023b): What path is the most wholesome? In: Reddit. URL: https://www.reddit.com/r/kaiserredux/comments/18m9oik/what_path_is_the_most_wholesome/ [accessed 14.03.26].

- r/NPPfunny (2023). When the brainrot is real. In: Reddit. URL: https://www.reddit.com/r/NPPfunny/comments/1okeyip/when_the_brainrot_is_real/ [accessed 14.03.26].
- r/RedFloodMod (2022): If Red Flood was real and our timeline was mod, what would be the memes? In: Reddit. URL: https://www.reddit.com/FloodMod/comments/u5lfi2/if_red_flood_was_real_and_or_timeline_was_mod/ [accessed 14.03.26].
- r/RedFloodMod (2023a): Progress Report 34: TSF2's Ideology Expansion, or, the Re-Rework. In: Reddit. URL: https://www.reddit.com/r/RedFloodMod/comments/12bw7k2/progress_report_34_tsf2s_ideology_expansion_or/ [accessed 14.03.26].
- r/RedFloodMod (2023b): funny cosmist eurasianists. In: Reddit. URL: https://www.reddit.com/r/RedFloodMod/comments/1oxl712/funny_cosmist_eurasianists/ [accessed 14.03.26].
- r/RedFloodMod (2025a): Adolf Polzl's Germany. In: Reddit. URL: https://www.reddit.com/r/RedFloodMod/comments/1o9o57u/adolf_polzls_germany/ [accessed 14.03.26].
- r/RedFloodMod (2025b): played hyperborea for way too long but it was worth it. In: Reddit. URL: https://www.reddit.com/r/RedFloodMod/comments/vjopjs/played_hyperborea_for_way_too_long_but_it_was/ [accessed 14.03.26].
- r/RedFloodMod (2025c): Is Accelerationism really that bad? In: Reddit. URL: https://www.reddit.com/r/RedFloodMod/comments/1nsd82s/is_accelerationism_really_that_bad/ [accessed 14.03.26].
- r/TNOmod (2022): Why Speer's fascist regime in TNO sucks. In: Reddit. URL: https://www.reddit.com/r/TNOmod/comments/st6bhp/why_speers_fascist_regime_in_tno_sucks/ [accessed 14.03.26].
- r/TNOmod (2023): Ahh... does any remember the days of DSRfunny and Speer Hoodie? In: Reddit. URL: https://www.reddit.com/r/TNOmod/comments/13t3216/ahh_does_any_remember_the_days_of_dsrfunny_and/ [accessed 14.03.26].
- r/TNOmod (2025): Heinrich Himmler in Moskau. In: Reddit. URL: https://www.reddit.com/r/TNOmod/comments/1nfyzl3/heinrich_himmler_in_moskau/ [accessed 22.01.25].
- r/TheFireRisesMod (2025a): Guns of the Patriot Front. In: Reddit. URL: https://www.reddit.com/r/TheFireRisesMod/comments/1njmrao/guns_of_the_patriot_front/ [accessed 14.03.26].

- r/TheFireRisesMod (2025b): Atomwaffen General and Leaders based on Attractiveness. In: Reddit. URL: https://www.reddit.com/r/TheFireRisesMod/comments/1mrvvso/atomwaffen_generals_and_leaders_based_on/ [accessed 14.03.26].
- r/TheFireRisesMod (2025c): This is roughly how I imagine Jacksonian foreign policy. In: Reddit. URL: <https://www.reddit.com/r/TheFireRisesMod/comments/1nuktti/comment/nh3ws3p/> [accessed 14.03.26].
- r/TheFireRisesMod (2025d): Aleksandr Dugin be like. In: Reddit. URL: https://www.reddit.com/r/TheFireRisesMod/comments/1jluhxf/aleksandr_dugin_be_like/ [accessed 14.03.26].
- r/TheFireRisesMod (2025e): why do you guys replace O with Q in context of Trump? In: Reddit. URL: https://www.reddit.com/r/TheFireRisesMod/comments/1nozat5/why_do_you_guys_replace_o_with_q_in_context_of/ [accessed 14.03.26].
- r/TheFireRisesMod (2025f): Eating ze bugs with – Klaus Schwab EU AAR. In: Reddit. URL: https://www.reddit.com/r/TheFireRisesMod/comments/1lezh15/eating_ze_bugs_with_klaus_schwab_eu_aar/ [accessed 14.03.26].
- r/TheFireRisesMod (2025g): What is life like under Klaus Schwab's Europe? In: Reddit. URL: https://www.reddit.com/r/TheFireRisesMod/comments/1ldzz8g/what_is_life_like_under_klaus_schwabs_europe/ [accessed 14.03.26].
- r/TheFireRisesMod (2025h): jordan jereb endorses. In: Reddit. URL: https://www.reddit.com/r/HOI4memes/comments/1nkm3iz/jordan_jereb_endorses_tfr/ [accessed 14.03.26].
- Steam Workshop (2021): The New Order: Last Days of Europe. In: Steam. URL: <https://steamcommunity.com/sharedfiles/filedetails/?id=2438003901> [accessed 14.03.26].
- Steam Workshop (2024): Roter Morgen. In: Steam. URL: <https://steamcommunity.com/workshop/filedetails/?id=3277861826> [accessed 14.03.26].
- The Fire Rises Wiki (n.d.): The Fire Rises Wiki. In: The Fire Rises Wiki. URL: https://thefireriseswiki.miraheze.org/wiki/The_Fire_Rises_Wiki [accessed 14.03.26].
- The Fire Rises Wiki (n.d.): Acknowledgment of The Fire Rises by People or Organisations featured in the Mod. In: The Fire Rises Wiki. URL: https://thefireriseswiki.miraheze.org/wiki/Acknowledgement_of_The_Fire_Rises_by_people_or_organizations_featured_in_the_mod [accessed 14.03.26]

The New Order Fandom Wiki (n.d.): Albert Speer. In: The New Order Fandom Wiki. URL: https://the-new-order-last-days-of-europe.fandom.com/wiki/Albert_Speer [accessed 14.03.26].

The New Order Fandom Wiki (n.d.): Valery Yemelyanov. In: . In: The New Order Fandom Wiki. URL: https://the-new-order-last-days-of-europe.fandom.com/wiki/Valery_Yemelyanov [accessed 14.03.26].

The Red Flood Fandom Wiki (n.d.): Ideologies#Accelerationism. In: The Red Flood Fandom Wiki. URL: <https://red-flood.fandom.com/wiki/Ideologies#Accelerationism> [accessed 14.03.26].

The Red Flood Official Wiki (n.d.): The Red Flood Official Wiki. In: The Red Flood Official Wiki. URL: <https://redflood.wiki.gg/> [accessed 14.03.26].

TNOpediA (2024): Memes#Speer_Hoodie. [accessed via wayback machine.] URL: http://web.archive.org/web/20251126112104/https://tno.wiki/wiki/Memes#Speer_Hoodie [accessed 14.03.26].

Images

Fig. 1: r/RedFloodMod (2025): “Goebbels has gone woke.” In: Reddit. URL: https://www.reddit.com/r/RedFloodMod/comments/1jraear/goebbels_has_gone_woke/ [accessed 14.03.26].

Fig. 2: r/NPPfunny (2025): “Clerical Fascism/Burgundian System unity I guess.” In: Reddit. URL: https://www.reddit.com/r/NPPfunny/comments/1lorhjc/clerical_fascismburgundian_system_unity_i_guess/ [accessed 14.03.26].

Fig. 3: r/TNOmod (2020): “My Speer hoodie arrived today!” In: Reddit. URL: https://www.reddit.com/r/TNOmod/comments/j2xj8e/my_speer_hoodie_arrived_today/ [accessed 14.03.26].

Fig. 4: r/NPPfunny (2022): “real.” In: Reddit. URL: <https://www.reddit.com/r/NPPfunny/comments/wc7zyx/real/> [accessed 14.03.26].

Fig. 5, Fig. 6: r/NPPfunny (2025): “I like speer (don't prove me wrong bormannists).” In: Reddit. URL: https://www.reddit.com/r/NPPfunny/comments/1l3tai6/i_like_speer_dont_prove_me_wrong_bormannists/ [accessed 14.03.26].

Fig. 7: r/TNOmod (2020): “TNO mfs be like.” In: Reddit. URL: https://www.reddit.com/r/TNOmod/comments/jj7b7t/tno_mfs_be_like/ [accessed 14.03.26].

Fig. 8: r/TNOmod (2020): "He has become enlightened." In: Reddit. URL: https://www.reddit.com/r/TNOmod/comments/jsziv2/he_has_become_enlightened/ [accessed 14.03.26].

Games

The Fire Rises team (2024): The Fire Rises. PC

Führerreich team (2019): Führerreich: Legacy of the Great War. PC.

KaiserreduX team (2020): KaiserreduX. PC.

Kaiserreich team (2018): Kaiserreich. PC.

The New Order team (2021): The New Order: Last Days of Europe. PC.

Paradox Development Studio (2016): Hearts of Iron IV. Paradox Interactive. PC.

Red Flood team (2022): Red Flood. PC.

Anti-Diversität im Gaming – ein Shortcut in rechtsextreme Gaming-Communitys?

Julia Reinhold & Mick Prinz

Abstract This article examines toxic and right-wing extremist dynamics within digital gaming communities. Gaming has developed into an important social infrastructure in which identities, political views, and cultural conflicts are negotiated. While most gamers use video games as a form of leisure and social interaction, a small but vocal minority exploits platform structures to spread discriminatory narratives and right-wing extremist ideas. The article outlines how far-right actors strategically use gaming spaces for networking, mobilization, and metapolitical influence. Particular attention is paid to anti-diversity discourses that serve as a bridge between toxic online communities and far-right milieus. Using examples from current gaming controversies and online platforms, the article illustrates how narratives about ›wokeness‹, diversity, or alleged censorship are instrumentalized to construct hostile images and normalize far-right interpretations. The article concludes with a discussion of possible countermeasures and emphasizes the importance of digital civil courage, media education, and stronger platform governance.

Keywords Far-right extremism in gaming, anti-diversity narratives, metapolitical influence, platform cultures, digital civil courage

Abstract Dieser Beitrag untersucht toxische und rechtsextreme Dynamiken innerhalb digitaler Gaming-Communities. Gaming hat sich zu einer wichtigen sozialen Infrastruktur entwickelt, in der Identitäten, politische Ansichten und kulturelle Konflikte ausgehandelt werden. Während die meisten Gamer*innen Videospiele als eine Form der Freizeitgestaltung und sozialen Interaktion nutzen, instrumentalisiert eine kleine, aber lautstarke Minderheit Plattformstrukturen, um diskriminierende Narrative und rechtsextreme Ideen zu verbreiten. Der Artikel skizziert, wie rechtsextreme Akteur*innen Gaming-

Räume strategisch für Vernetzung, Mobilisierung und metapolitische Einflussnahme nutzen. Besondere Aufmerksamkeit wird dabei Anti-Diversity-Narrativen gewidmet, die als Brücke zwischen toxischen Online-Communities und rechtsextremen Milieus fungieren. Anhand von Beispielen aus aktuellen Gaming-Kontroversen und Online-Plattformen veranschaulicht der Artikel, wie Erzählungen über ›Wokeness‹, Diversität oder angebliche Zensur instrumentalisiert werden, um Feindbilder zu konstruieren und rechtsextreme Deutungen zu normalisieren. Der Artikel schließt mit einer Diskussion möglicher Gegenmaßnahmen und betont die Bedeutung von digitaler Zivilcourage, Medienkompetenz und einer stärkeren Plattform-Governance.

Schlagworte Rechtsextremismus im Gaming, Anti-Diversity-Narrative, Metapolitische Einflussnahme, Plattformkulturen, Digitale Zivilcourage

1. Tutorial – Kulturkampf im Gaming

Debatten um Videospiele lassen sich als Spiegelbild gesellschaftlicher Konflikte und damit als eine Gegenüberstellung unterschiedlicher Meinungen verstehen, die auch an vielen anderen Stellen der digitalen und analogen Welt ausgetragen werden. Diskussionen um beispielsweise die Repräsentation unterschiedlicher sexueller Identität, stereotyp-arme Darstellungen von nicht-männlichen Protagonist*innen oder Fragen der Barrierefreiheit werden in Filmen, Serien, aber auch in anderen Kulturmedien schon seit geraumer Zeit geführt. Auch im Gaming sind diese Themen längst angekommen, stoßen hier aber auf spürbar größeren Widerstand als in der restlichen Gesellschaft.

Ist das Kulturgut Games also ein toxischer Spiegel der Gesamtgesellschaft? Nein, eine solche Formulierung würde die Realität verklären und die vielen positiven und progressiven Errungenschaften der Gaming-Kultur negieren. Allerdings gibt es im Gaming eine laute, toxische Minderheit, welche strukturelle Missstände von Gaming-Plattformen, sowie Videospiel-Communitys ausnutzt und versucht, sich als größer zu inszenieren, als sie ist. So beispielsweise auch beim Action-Rollenspiel *Horizon Forbidden West* (2022), bei welchem sich zunächst ein Teil der Spieler*innen über die vermeintlich ›unattraktive‹ und zu wenig ›weibliche‹ Darstellung der Protagonistin Aloy echauffierte. Als dann 2023 mit *Burning Shores* eine Erweiterung für das Spiel erschien und eine gleichgeschlechtliche Liebesszene optional in die Handlung integrierte, illustrierten unzählige negative Kommentare und Bewertungen die Verbreitung

von Queerfeindlichkeit in reaktionären Gaming-Communitys. Während die Fachpresse das Action-Rollenspiel mit 81 von 100 Punkten bewertete¹, führte das Review Bombing² wütender Fans zu einer zeitweise unterdurchschnittlichen Wertung von 3.8 von 10 möglichen Punkten.

Dieser Trend spielt einer weiteren, noch kleineren, aber deutlich gefährlicheren Minderheit im Gaming in die Hände: rechtsextremen³ Gaming-Communitys. Rechtsextremismus im Gaming hat viele Gesichter, die im weiteren Verlauf des Beitrags skizziert werden. An dieser Stelle soll einleitend betont werden, dass auch die Debatten um vermeintlich zu ›woke‹ und ›politische‹ Videospiele strategisches Element eines extrem rechten Kulturkampfes sind. Die Akzeptanz von ›Anti-Diversitätsnarrativen‹ führt dazu, dass sich extrem rechte Videospieler*innen in In-Game-Chats und auf Gaming-Plattformen wohlfühlen und sie ihre Ideologie auch im Kulturgut Games verbreiten. Regelmäßig buhlen rechtsextreme Akteur*innen in Anti-Diversity-Debatten um Aufmerksamkeit und versuchen, eine Anschlussfähigkeit zu reaktionären Gaming-Communitys herzustellen. So auch beispielsweise im Diskurs um das Spiel *Kingdom Come Deliverance II* (2025), als einer der Köpfe der rechtsextremen *Identitären Bewegung*, Martin Sellner, versuchte, in einem Post auf der Plattform X seine rassistisch motivierte Argumentation durch den Aspekt vermeintlicher historischer Genauigkeit zu legitimieren und damit die breite Masse an Spieler*innen abzuholen. Gegenstand der Kritik war die Darstellung eines gebildeten Adligen aus Mali. Der Rechtsextremist⁴ beklagt die Kultur und Darstellungsweise einer nicht-weißen Figur im Rollenspiel und teilte die Mär vom angeblich weißen Mittelalter.

In diesem Beitrag wird zunächst genauer auf den Begriff der toxischen Gaming-Communitys und den Grund für ihr Agieren in Videospiel-Welten eingegangen. Anschließend liefert der Artikel einen Überblick über Rechtsextremismus im Gaming und geht dabei vor allem auf drei zentrale Motive

-
- 1 Durchschnittswert unterschiedlicher Fachrezensionen auf der Plattform Metacritic.
 - 2 Review Bombing meint die gezielte Platzierung von Wertungen von zahlreichen Nutzer*innen. In der Regel sind diese orchestrierten Bewertungen entweder unverhältnismäßig positiv oder negativ.
 - 3 Rechtsextremismus ist im Kern eine Einstellung, die die Gleichwertigkeit aller Menschen ablehnt und auf Ungleichwertigkeitsvorstellungen zielt. Teile eines rechtsextremen Weltbilds sind u.a. Rassismus, Antisemitismus, eine Ablehnung liberaler Demokratien, sowie Autoritarismus und Chauvinismus (Belltower News 2012).
 - 4 Der Verfassungsschutz stuft Martin Sellner als Führungskader der rechtsextremistischen *Identitären Bewegung* ein (Bundesamt für Verfassungsschutz 2024).

ein, wie und warum die extreme Rechte Videospiele nutzt. Zum Abschluss liefert der Beitrag mögliche Handlungsempfehlungen für den Umgang mit aktuellen Herausforderungen.

2. Toxische Gaming-Communitys – eine Bestandsaufnahme

Mehr als die Hälfte der Deutschen spielt Computer- oder Videospiele, viele davon online (Bitkom 2025). Für viele Menschen ist Gaming dabei in erster Linie ein Hobby und schafft einen sozialen Raum, in dem sie Freund*innen treffen, miteinander spielen und sich über ihre Interessen austauschen. Diese Alltäglichkeit macht Gaming zu einem Ort, an dem nicht nur Unterhaltung geboten wird, sondern auch Beziehungen entstehen und Zugehörigkeit ausgehandelt wird. Digitale Medien zählen für vornehmlich junge Menschen zu den wichtigsten Quellen zur politischen Meinungsbildung (Feierabend et al. 2025: 50ff.). Damit sind auch die unterschiedlichen Gaming-Räume Orte der politischen Sozialisation.

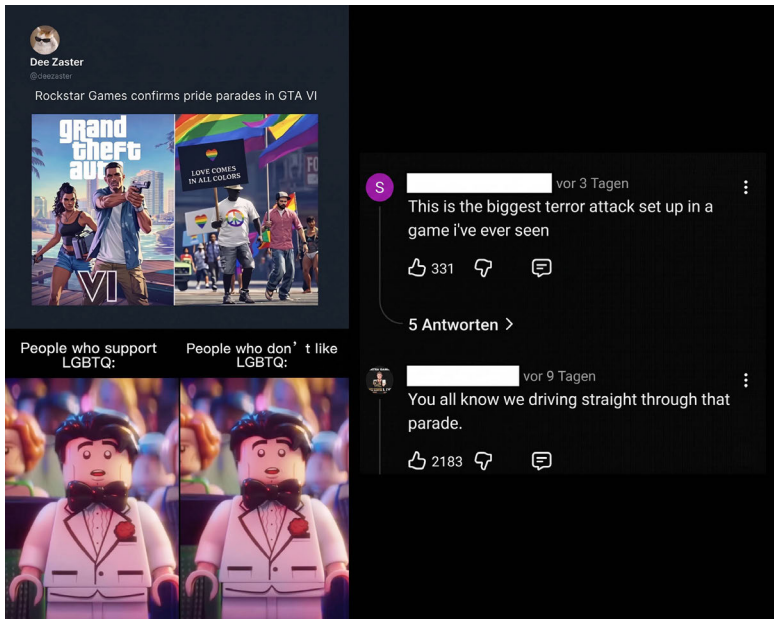
Eine einheitliche ›Gamer-Szene‹ gibt es so nicht, auch wenn das von fachferneren Politiker*innen so oder in ähnlichen Formen wiederholt postuliert wird. So auch beispielsweise vom ehemaligen Innenminister Horst Seehofer nach dem antisemitischen Anschlag von Halle (Spiegel, 2019). Eine Studie der Bertelsmann Stiftung (Rother et al. 2025) hat sich mit der Beschaffenheit und den demokratischen Potenzialen von Gaming-Communitys beschäftigt und verifiziert, dass hier unterschiedlichste Akteursgruppen aufeinandertreffen. Die Erhebung analysiert Einstellungspotenziale von Videospieler*innen aus Deutschland. Hier spielen zwei Drittel der Menschen ab 16 Jahren digitale Spiele. In der Altersgruppe von 16 bis 34 Jahren sind es sogar 86 %. Ein Ergebnis der Studie: Jene, die besonders viel spielen und sich selbst als ›Gamer‹ bezeichnen, nehmen häufiger an demokratischen Prozessen teil. So meldet diese Gruppe der ›Gaming-Enthusiast*innen‹ im Vergleich zur Gesamtbevölkerung eher einen Kommentar bei der Polizei oder bei Sicherheitsbehörden und bezieht häufiger Stellung zu politischen Sachverhalten auf Gaming- und Gaming-nahen Plattformen (ebd.). Die Bertelsmann Stiftung widerlegt damit zum einen die Mär vom vermeintlich ›unpolitischen Gamer‹, unterstreicht im weiteren Studienverlauf aber auch, dass demokratische- und antidemokratische Einstellungspotenziale innerhalb der unterschiedlichen Communitys stark variieren.

In Gaming-Welten werden also viele Einstellungspotenziale widergespiegelt, die sich auch abseits vom Bildschirm in der Gesellschaft identifizieren lassen. Es darf daher nicht überraschen, dass auch hier verschiedene Formen von Diskriminierung und abwertendem Verhalten präsent sind. Je nach Plattform, Community oder Spiel werden Spieler*innen früher oder später mit toxischem Verhalten konfrontiert oder selbst in Berührung kommen. Aber was meint das überhaupt?

Communities, in denen Beleidigungen, Mobbing, unsportliches Verhalten wie ›Griefing‹ (die Sabotage des Spielerlebnisses für Mitspieler*innen) und ›Intentional Feeding‹ (absichtliches Verlieren, um das eigene Team am Sieg zu hindern), sowie diskriminierende und abwertende Sprache zum alltäglichen Ton gehören, werden im Folgenden als ›toxische Communities‹ bezeichnet. Toxisch heißt giftig und beschreibt belästigendes Verhalten, das andere angreift oder auf verschiedene Weise einschränkt, also die Kommunikation in einer Gemeinschaft metaphorisch ›vergiftet‹. Der größte Teil der Gamer*innen hat dabei schon Hass in Videospielräumen erlebt – ob innerhalb der Spielechats, auf Plattformen oder durch private Nachrichten. Eine Studie der *Anti-Defamation League* (ADL) zeigt, dass 76 % der Ü18-Spieler*innen schon Hassrede im Gaming wahrgenommen haben (ADL 2024). Auch 15 % der erwachsenen Spieler*innen sind schon auf eindeutig rechtsextreme Agitationen in Gaming-Communities gestoßen. Die Studie zeigt, dass es besonders häufig weiblich gelesene Spielende, nicht-weiße Spieler*innen und jüdische Spieler*innen, sowie Mitglieder der LGBTQIA+-Community sind, die Hass erfahren. Diesen Spieler*innen wird häufig mit u.a. queerfeindlichen, rassistischen oder anderen menschenfeindlichen Aktionen begegnet, wenn sie sich in Lobbys, Voice-Chats oder im Profiltexat als Teil einer von Diskriminierung betroffenen Gruppe zu erkennen geben (ebd.). Das zeigt beispielsweise auch die jüngste Debatte um die Abbildung von Pride-Paraden im angekündigten neuen Ableger der GTA-Reihe. Laut unbestätigten Gerüchten (Stand Januar 2026) beinhaltet die Spielwelt von GTA 6 dynamische Pride Paraden als Teil der Spielwelt von Vice City. In Kommentaren unter Memes auf Reddit, Instagram und YouTube sammeln sich dabei pseudohumoristische Kommentare, während an anderer Stelle der Anti-Diversitäts-Backlash deutlich wird und sich Spieler*innen negativ gegenüber der Einbindung von LGBTQIA+ äußern. So wird beispielsweise unter einem YouTube Short mit Memes über GTA 6 (siehe Abb. 1) kommentiert, »This is the biggest terror attack set up in a game I've ever seen« oder noch konkreter »You all know we driving straight through that parade« und letzterer Kommentar, erhält tausende Likes. Dabei muss

unterstrichen werden, dass nur ein Teil der Spieler*innen hinter den eigenen toxischen Aussagen stehen. Bei vielen steht die Provokation und das Ausreizen von Grenzen im Vordergrund, was umgangssprachlich auch als ›edgy‹ (s. dazu auch Niegl in diesem Band) Verhalten benannt wird.

Abb. 1: Screenshots eines Memes über GTA 6 in einem YouTube Short. Der dazugehörige Kommentarbereich voller ›edgy‹ und toxischer Kommentare, erstellt am 14.01.26.



Trotzdem: Dieses toxische Verhalten und die dazugehörigen Communities exkludieren vor allem marginalisierte Gruppen aus verschiedenen Interaktionsfeldern der Gaming-Kultur und schränken ihr Verhalten und Auftreten ein. Regelmäßig wird versucht, demokratische und progressive Stimmen zum Schweigen zu bringen. Diese Strategie nennt sich ›Silencing‹ und ist zudem beliebte Praxis der sogenannten ›Neuen Rechten‹, um Menschen aus sozialen Netzwerken zu verdrängen (BMI 2024: 10f.). Dieses toxische Verhalten im Gaming umfasst also Beleidigungen, Hassrede, sexuelle Belästigung, diskriminierende Profile/Nicknames, Doxxing (Veröffentlichung personenbezogener Daten) und die gezielte Verbreitung von Desinformation.

3. Rechtsextremismus im Gaming – Ein Überblick

Um den Einfluss rechtsextremer Akteur*innen im Gaming zu verstehen, muss Gaming selbst als das betrachtet werden, was es in der Praxis ist: eine große soziale Infrastruktur. Wer spielt, bewegt sich heute fast automatisch durch ein Ökosystem aus Plattformen, die unterschiedliche Funktionen erfüllen und jeweils eigenen Kommunikationslogiken folgen (Röther et al. 2024). Spiele werden über Distributionsplattformen wie Steam gekauft und bewertet, Kommunikation läuft während des Spielens über Dienste wie Discord und auf Twitch, YouTube oder Reddit werden Inhalte weiterverbreitet, kommentiert und memefiziert. Gaming ist nicht nur Unterhaltung, sondern ein dauerhafter sozialer Bezugsrahmen mit einer eigenen Kommunikationskultur.

Vor diesem Hintergrund wird verständlich, warum Gaming als Bühne für die extreme Rechte so attraktiv ist. Die Einstiegshürden sind niedrig, die Reichweite hoch, und Beteiligung ist in vielen Formaten Teil des Systems. Zusammenhalt entsteht dabei seltener durch politische Nähe als durch geteilte Erfahrungen: Gemeinsam spielen, Missionen verfolgen, Siege feiern oder Niederlagen beklagen. Daraus entwickelt sich schnell eine gemeinsame Sprache aus Memes, Insider-Witzen und Codes. Für viele schafft das ein Gefühl der Identität und kann als sicherer Raum dienen.

Problematisch wird es, wenn genau diese Nähe und gemeinsame Sprache bewusst genutzt werden, um Einfluss auszuüben: wenn Debatten gekapert, Grenzen des Sagbaren verschoben und Feindbilder gezielt stabilisiert werden. Dadurch können sowohl langjährige Mitglieder der Community als auch Personen, die neu in diese sozialen Räume stoßen und Anschluss suchen, langsam an bestimmte Deutungen herangeführt werden. Rechtsextreme Akteur*innen nutzen dabei typischerweise drei Strategien, die ineinandergreifen: Vernetzung, Mobilisierung und metapolitische Einflussnahme. Diese werden im weiteren Verlauf analysiert.

3.1 Vernetzung

Gaming-bezogene Räume sind für rechtsextreme Akteur*innen ein attraktives Umfeld, nicht nur, weil sie einen niederschweligen Zugang zu einem breiten Publikum ermöglichen, sondern auch, weil ihr Engagement dort mit vergleichsweise geringen Risiken verbunden ist. Aufgrund uneinheitlicher Moderation und der weit verbreiteten Charakterisierung von Gaming-Communi-

tys als ›unpolitisch‹ werden die von ihnen produzierten Inhalte in vielen Fällen nicht als extremistisch erkannt, geahndet oder entfernt.

Zugleich bringen die Plattformen die notwendige Infrastruktur für Netzwerkarbeit und Beziehungsaufbau bereits mit. Community-Funktionen wie Gruppen, Profile, Freundeslisten und Chats erleichtern es, Kontakte zu knüpfen, Ideen auszutauschen und auf der Grundlage eines gemeinsamen Hobbys langfristige Verbindungen zu etablieren. Diese Vernetzung findet zu Beginn selten unter einem politischen Label statt. Stattdessen adaptieren rechtsextreme Akteur*innen die kulturellen Codes der Szene, um sich als authentischer Teil der Community zu inszenieren. Im Vordergrund stehen weniger die Argumente, sondern Anerkennung und Zugehörigkeit.

In der Praxis nutzen rechtsextreme Akteur*innen häufig sichtbare Erkennungszeichen, um sich untereinander zu identifizieren. Auf Plattformen wie Steam gibt es tausende Profile, Gruppen und Kommentare mit rechtsextremen Symbolen, Zahlencodes oder Anspielungen, die auf eine ideologische Zugehörigkeit hindeuten, ohne dass man offen politisch Stellung beziehen müsste. Dazu gehören Gruppen, die den Nationalsozialismus verherrlichen oder sich explizit auf rechtsterroristische Bewegungen beziehen. Diese Signale erleichtern es, Gleichgesinnte zu erkennen, Kontakte herzustellen und sich in der Szene zu orientieren. Ähnliche Dynamiken lassen sich auch in Spielwelten beobachten, insbesondere dort, wo Nutzer*innen eigene Inhalte erstellen können. Sandbox-Plattformen wie *Roblox* bieten umfangreiche Tools, mit denen die Spieler*innen eigene Spieleumgebungen, Rollen und Szenarien gestalten können. In einigen Fällen wird dieser Gestaltungsspielraum genutzt, um rechtsextreme Symbole und Feindbilder in spielerische Kontexte einzubetten, etwa wenn Rollenspiele entstehen, in denen Spieler*innen als SA- oder Wehrmachtsoffiziere auftreten oder rechtsextreme Gesten und Emotes verwenden. Relevant ist dies auch vor dem Hintergrund der Nutzer*innenstruktur: *Roblox* wird vor allem von sehr jungen Spieler*innen genutzt und ist laut JIM-Studie 2025 eines der beliebtesten digitalen Spiele bei Kindern und Jugendlichen zwischen 12 und 15 Jahren. Das bedeutet, dass rechtsextreme Referenzen ein Publikum erreichen, das nicht in der Lage ist, solche Inhalte angemessen zu kontextualisieren und dazu verleitet wird, sie als spielerische Elemente zu normalisieren.

Viele große Gaming-Plattformen dienen damit als Einstiegspunkte in den Rechtsextremismus. Öffentliche oder halböffentliche Bereiche sind Kontaktzonen, in denen rechtsextreme Akteur*innen Präsenz zeigen und Verbindungen anbahnen. Die anschließende Konsolidierung verlagert sich

häufig auf kleinere private Server oder (semi)geschlossene Kanäle (Heider 2025: 179f.). In solchen besser kontrollierbaren Umgebungen kann sich das soziale Klima schnell verändern: Loyalität, Hierarchie und sozialer Druck steigen innerhalb dieser ausgewählten Gruppe, während Widerspruch riskanter wird. Dabei handelt es sich um einen zentralen Mechanismus extremistischer Rekrutierung. Viele Neumitglieder treten nicht aus ideologischer Überzeugung ein, sondern weil sie Zugehörigkeit, Orientierung oder Anschluss an eine vermeintlich exklusive Gemeinschaft suchen (Ebner 2020: 82f.). Geschlossene Online-Räume erleichtern es extremistischen Akteur*innen, neue Mitglieder emotional zu binden, klare Verhaltensnormen zu etablieren und Lock-In-Effekte zu erzeugen, die einen späteren Ausstieg erschweren (ebd.).

Diese Vernetzungsprozesse können reale Konsequenzen haben, wie Fälle zeigen, in denen digitale Kontakte in Gewalt eskaliert sind. Beim rechtsterroristischen Anschlag am Münchner Olympia-Einkaufszentrum 2016 war der Täter in einer migrationsfeindlichen Gruppe auf Steam aktiv, wo er sich unter anderem über Ideologie und Waffenbeschaffung austauschte (Bernstein 2018). Ein weiteres Mitglied desselben Netzwerks verübte später einen tödlichen Anschlag in den USA. Die Rolle der digitalen Vernetzung für analogen Rechtsterrorismus wurde von den Sicherheitsbehörden zunächst unterschätzt.

3.2 Mobilisierung

Sobald diese Vernetzung hergestellt ist, bildet sie oft die Grundlage für weitere Mobilisierung. In abgeschirmten Räumen geht es nicht mehr nur darum, Teil einer ›auserwählten‹ Gruppe zu sein, sondern darum, gemeinsam zu handeln.

Ein bekanntes Muster ist koordinierte digitale Gewalt, die auf Einschüchterung zielt. Gruppen stimmen Angriffe auf Streams, Kommentarbereiche oder Community-Räume ab, beispielsweise in Form von sogenannten ›Hate Raids‹. Untersuchungen zu Hate Raids auf Twitch beschreiben solche Attacken als koordinierte, teils auch bot-gestützte Übergriffe in Echtzeit, die die Plattformlogik ausnutzen und besonders häufig marginalisierte Streamer*innen ins Visier nehmen (Cai et al. 2023: 2). Frauen, Mitglieder der LGBTQIA+-Community und nicht-weiße Gamer*innen sind dabei besonders betroffen (ebd.). Typisch ist die massenhafte Flutung von Chats mit beleidigenden Nachrichten, Emotes oder Gewaltandrohungen, wodurch die Moderation systematisch überlastet und Betroffene zum Abbruch der Liveübertragung gezwungen werden. Das Ziel ist maximale Störung: Räume sollen für bestimmte Personen unbenutzbar werden, bis der Rückzug zur naheliegenden

Option wird (Adams et al. 2025: 8f.). An diesem Punkt wird entschieden, wer sichtbar bleibt, wer zum Schweigen gebracht wird und folglich, welches Verhalten und welche Positionen erwünscht sind.

Mobilisierung kann auch durch Community-Events stattfinden, die auf den ersten Blick wie gewöhnliche Szenenformate erscheinen. Dazu gehören ›patriotische‹ Turniere oder ›Game Jams‹ aus rechtsextremen Kreisen, die Technologie, Vernetzung und Propaganda miteinander verbinden. Solche Veranstaltungen stärken die interne Sichtbarkeit und festigen Kontakte, ohne dass die Ideologie im Spiel selbst offensichtlich sein muss. Die politische Wirkung liegt dabei mehr im Sozialen: Es werden neue Kontakte geknüpft, Rollen verteilt und das Engagement wächst. So wird die Szene handlungsfähiger, ohne dass offene Propaganda im Vordergrund stehen muss.

In einigen Fällen orientiert sich die Mobilisierung daher gezielt an den Faktoren, die auch das Spielen motivierend machen. Gamifizierung bedeutet, dass Spielelemente außerhalb von Computerspielen zum Einsatz kommen, um Verhalten zu beeinflussen (Schlegel 2021: 2). Dazu zählen etwa Ranglisten, Abzeichen oder sichtbare Statusmarker, die Anerkennung gewähren und Handlungen mit Sinn aufladen. Ein gut dokumentiertes Beispiel ist die rechtsextreme Discord-Gruppe *Reconquista Germanica*, die ein militärisch organisiertes Rang- und Abzeichensystem einführte. So konnten Nutzer*innen durch Aktivität und Beteiligung aufsteigen und ihren Status offen sichtbar machen, wodurch Hierarchien, Loyalität und soziale Anbindung innerhalb der Gruppe verstärkt wurden (Ebner 2019: 188f.). Koordinierte Angriffe oder Kampagnen erscheinen dadurch weniger als politische Arbeit, sondern eher als gemeinsame Missionen, die Ansehen versprechen und Anpassungsdruck erzeugen.

In extremen Fällen findet diese Logik sogar Anwendung auf physische Gewalt. In 2019 benannte der rechtsextreme Attentäter von Halle (Saale) in seinen Texten »Achievements« und bezog sich auf »Highscore«-Vergleiche nach Opferzahlen (BfV 2024: 40). Solche Fälle sind die Ausnahme, machen aber sichtbar, wie Gaming-Bezüge als Deutungs- und Inszenierungsrahmen genutzt werden können.

3.3 Metapolitische Einflussnahme

Während Vernetzung und Mobilisierung auf unmittelbare Effekte abzielen, verfolgt die extreme Rechte mit der sogenannten Metapolitik ein übergeordnetes Ziel. Im Kern beschreibt der Begriff den Versuch, durch kulturelle

Einflussnahme gesellschaftliche Einstellungen und Normen schrittweise nach rechts zu verschieben und die rechtsextreme Ideologie langfristig wieder salonfähig zu machen (Salzborn 2020: 22). Mit anderen Worten: Es handelt sich um einen vopolitischen Kampf um Deutungshoheit. Durch gezielte Kulturarbeit und Raumnahme versuchen rechtsextreme Akteur*innen, das Denken der Bevölkerung fundamental zu beeinflussen. Diese Strategie ist eng mit der Neuen Rechten verbunden und folgt der Annahme, dass Politik nicht erst in Parlamenten entsteht, sondern in Kultur, Alltag und Sprache, dort wo Selbstverständlichkeiten und Normen gesetzt werden (ebd.: 23).

Gaming eignet sich dafür besonders, weil viele Auseinandersetzungen ohnehin als Identitätsfragen verhandelt werden: Wer gehört dazu? Wer stört? Wer darf Kritik üben? Rechtsextreme Akteur*innen setzen genau hier an. Sie suchen Themen, die bereits aufgeladen sind und sich gut skandalisieren lassen, zum Beispiel Moderation und ›Zensur‹, Diversität, Geschlechterrollen oder die Erzählung einer vermeintlich ›aufgezwungenen Wokeness⁵. Daraus entsteht eine wiederkehrende Rahmung: Progressive Kritik gilt als übergriffige Ideologie, rechte Deutungen dagegen als Verteidigung von Freiheit, Humor oder ›echter‹ Gaming-Kultur. Widerspruch erscheint so nicht als normaler Teil einer Debatte, sondern als Angriff auf die Community.

Zentral ist dabei die Verschiebung von Bedeutungen: Sprache und Symbole werden so eingesetzt, dass sie nicht nur Zugehörigkeit signalisieren, sondern auch Interpretationsspielraum öffnen. Rechtsextreme Akteur*innen bewegen sich dabei gezielt in einer Grauzone zwischen Humor, Provokation und Ernsthaftigkeit. Dieses Vorgehen kalkuliert mit ein, dass expliziter Rechtsextremismus viele Menschen abschreckt und schafft daher Zugangspunkte, bevor die Ideologie klar benannt wird. Ironie und vermeintliche Witze dienen dabei als Schutzmechanismus. So entstehen Codes, die auf zwei Ebenen funktionieren: Für Außenstehende wirken sie wie harmloser Netzsargon, für Eingeweihte sind sie eindeutige Signale (Schlözel 2025: 287).

5 ›Wokeness‹ wird in rechten Kontexten als abwertender Begriff für progressive Menschen verwendet. Im Begriff impliziert werden Aspekte wie ›moralische Überlegenheit‹, ›politische Korrektheit‹, ›Kulturmarxismus‹ als auch ›linksgrüne Ideologie‹. Generell soll damit das Feindbild eines totalitär agierenden, sich moralisch überlegen fühlenden progressiven Eiferers inszeniert werden, der weniger ›woke‹ Menschen durch haltlose Vorwürfe des Rassismus/Sexismus/etc. von einem öffentlichen Diskurs ausschließen will (Belltower-News 2026).

Im Gaming-Kontext können solche Codes etwa Begriffe wie ›NPC‹, ›Soyboy‹ oder ›Clown World‹ sein. Sie erscheinen zunächst wie ironische Memes oder provokanter Humor, transportieren in bestimmten Kontexten jedoch politische Bedeutungen. Mit ›NPC‹ werden Personen bezeichnet, denen unterstellt wird, sie würden wie nicht spielbare Figuren (Non-Player Character) in einem Videospiel lediglich vorgefertigten progressiven Skripten folgen. ›Soyboy‹ dient der Abwertung von Männern, die als zu feminin oder politisch progressiv gelten. Der Ausdruck ›Clown World‹ bezeichnet in rechten On-line-Kreisen eine vermeintlich absurde oder degenerierte gesellschaftliche Ordnung und wird häufig verwendet, um Vielfalt, Gleichberechtigung oder progressive Politik als Zeichen des kulturellen Verfalls zu markieren.

Diese strategische Ambiguität ermöglicht es, sich bei Kritik auf ›ein Missverständnis‹ oder Ironie zu berufen und damit der Verantwortung zu entziehen. Extremistische Inhalte bleiben so präsent, ohne eindeutig benannt oder seitens der Community sanktioniert zu werden.

Was im Kontext der Vernetzung als Erkennungsmerkmal dient, entfaltet auf metapolitischer Ebene einen Gewöhnungseffekt. Durch die wiederholte Verwendung solcher Codes, verschiebt sich nach und nach der Deutungsrahmen innerhalb der Community. Memes und spöttische oder ironische Anspielungen werden dabei zu Trägern politischer Deutungen von Geschlecht, Zugehörigkeit und sozialer Ordnung. Einzelne rechtsextreme Bezüge erscheinen online oft als bloße Provokationen, in der Masse tragen sie jedoch dazu bei, bestimmte Weltanschauungen zu normalisieren und sie als Teil des alltäglichen Kommunikationsstils im Gaming zu verankern. Die Wirkung entsteht hier nicht durch Einzelfälle, sondern durch Dauerpräsenz.

Zwar spielen auch eigene rechtsextreme Kulturproduktionen eine Rolle, sind in der Regel jedoch kein Massenphänomen. Dennoch reicht ihre Existenz bis in die 1980er Jahre zurück. Eines der frühen Beispiele ist ein Vernichtungslager-Simulationsspiel, das primär in rechtsextremen Kreisen zirkulierte und die Ermordung von Migrant*innen und anderen Feindbildern der extremen Rechten zum Spielziel hatte. Neuere Produktionen arbeiten demgegenüber häufig mit verschwörungsideologischen und pseudo-humoristischen Elementen. In einem Actionspiel aus dem Jahr 2024 etwa werden rechtsextreme Narrative spielbar gemacht, indem Spielende gegen angebliche Impf-Zom-

bies, ›Gay Frogs‹ oder wohnungslose Menschen kämpfen, während politische Gegner*innen als Teil eines vermeintlichen ›Deep State‹⁶ porträtiert werden.

Verbreiteter als eigene Propaganda-Produkte sind Modifikationen für populäre Mainstream-Spiele. Mods ermöglichen es, bestehende Spielwelten ideologisch umzudeuten, ohne neue Titel etablieren zu müssen. Gerade in historischen Strategiespielen oder Open-World-RPGs finden sich zahlreiche Erweiterungen, die den Nationalsozialismus relativieren, die Waffen-SS als spielbare Fraktion einführen oder Adolf Hitler als Spielfigur integrieren (s. dazu auch Niegl in diesem Band). In anderen Fällen werden progressive Inhalte gezielt entfernt oder ersetzt, beispielsweise durch das Umschreiben queerer NPCs, das Löschen von Verweisen auf ›Black Lives Matter‹ aus Spielwelten oder die Hypersexualisierung bzw. vollständige Entblößung weiblicher Charaktere. Diese Eingriffe scheinen zwar klein, verändern aber die Bedeutung bestehender Spiele und machen rechtsextreme Interpretationen anschlussfähig für ein breiteres Publikum.

Solche Dynamiken sind nicht neu, sondern haben sich in wiederkehrenden Konflikten innerhalb der Gaming-Kultur erprobt und verfestigt. Immer wieder entzünden sich Auseinandersetzungen dort, wo Fragen von Repräsentation, Geschlecht und gesellschaftlichem Wandel in den Fokus rücken. Deshalb lohnt es sich, einzelne Fälle genauer zu betrachten, um die Mechanismen nachzuvollziehen.

Dabei bleibt wichtig, dass Rechtsextremismus nicht stellvertretend für eine heterogene Gaming-Kultur und vielfältige Communitys herangezogen werden darf. Aufgrund der dargestellten systemischen Bedingungen innerhalb des Gaming-Ökosystems versuchen rechtsextreme Kräfte jedoch, auch im Kontext von Videospielen Menschenfeindlichkeit zu normalisieren.

4. Anti-Diversity als Bindeglied

Toxische Gaming-Communitys greifen häufig auf das Narrativ vom ›unpolitischen‹ Gaming zurück (Dietrich 2022). Dabei handelt es sich jedoch um ei-

6 Die Idee eines ›Deep State‹ ist Teil einer verschwörungsideologischen Erzählung, die auch in rechtsextremen Milieus verbreitet ist. Dahinter steht die Vorstellung einer geheimen, nicht gewählten Elite, die im Hintergrund die staatlichen Institutionen kontrolliere, während die gewählte Regierung lediglich als Fassade diene (s. Mattsson, Johansson 2022).

ne selektive Regel: Krieg, Gewalt, Männlichkeit oder auch heterosexuelle Liebesgeschichten gelten als reine Unterhaltung. Ein queerer Charakter oder eine Frau*, die nicht dem eigenen Ideal entspricht, wird dagegen schnell zur ›Agenda‹. Politik stört also immer dann im Gaming, wenn sie als progressiv, feministisch oder queer lesbar ist.

Diese Argumentation ist nicht neu. Schon die 2013 beginnende Dynamik um ›Gamergate‹ griff auf ähnliche Argumente und Strategien zurück. Die Gamergate-Kontroverse lässt sich dabei als eine der ersten koordinierten antifeministischen Hetzkampagnen bezeichnen. Was mit den Anschuldigungen eines gekränkten Ex-Partners gegenüber der*dem nicht binären Videospielentwickler*in Zöe Quinn begann, entwickelte sich zu einer strukturierten und anhaltenden Belästigungs- und Diffamierungskampagne gegenüber nicht-männlichen Spielenden. Unter dem Hashtag Gamergate versammelten sich hundertfach misogyne und antifeministische Kommentare. Vor allem die amerikanische ›Alt Right‹ und die rechtspopulistische Medienseite *Breitbart* nutzten Gamergate als Testballon, um die Macht von aggressiven Online-Communitys einzuschätzen und gleichzeitig Sexismus und Queerfeindlichkeit zu verbreiten (Ballweber 2024). Gamergate entwickelte sich rasant zur Hashtag-Kampagne für jene, die progressive Entwicklungen der Gaming-Kultur ablehnten. Vor über zehn Jahren schon versuchte sich also die ›Alt Right‹ an einer gezielten Instrumentalisierung wütender Gamer*innen, um rechtsextreme Narrative schrittweise zu normalisieren (Dietrich 2022). In zahlreichen Videos und Kommentare postulierten Anhänger*innen den vermeintlichen Untergang der Gaming-Industrie aufgrund einer destruktiven ›political correctness‹. Die gegenwärtige Debatte um ›Wokeness‹ und Gaming greift damit auf zahlreiche Argumente und Mechanismen zurück, die also schon vor Jahren mit der aufkommenden Gamergate-Kampagne erprobt und in ähnlichen Worten formuliert wurden.

So auch beim Spiel *Starfield* (2023): Hier genügte die Implementierung einer optionalen Pronomen-Auswahl im Charakter-Editor, um in Teilen der toxischen Gaming-Szene Diskussionen über Indoktrination und eine ›woke Agenda‹ auszulösen (Steam Reviews zu *Starfield* 2023).⁷ Schnell wurden auch Mods veröffentlicht, die diese Option entfernen sollten. Eine ähnliche Debatte entstand um *Kingdom Come: Deliverance II* (2025) rund um eine optionale

7 Zum *Starfield* Release im September 2023 waren die Steam Reviews voll von Anti-Diversity-Bewertungen. Viele dieser Bewertungen wurden Stand 2026 gelöscht.

gleichgeschlechtliche Romanze. In reaktionären Kreisen artete die Diskussion schnell in homophobe Beleidigungen und empörte Kommentare aus, oft begleitet von Behauptungen, dass solche Inhalte »aufgezwungen« würden (Dreßler 2025).

In den meisten Fällen reichen schon kleine Details aus, um diese Reaktionen auszulösen: eine nicht-binäre Option im Editor, eine Hauptfigur, die nicht weiß ist, eine queere Liebesbeziehung, ein paar Zeilen Dialog, die sich mit Diskriminierung beschäftigen. Einzelne Designentscheidungen werden dann zu einem kulturellen Signal, die in dem typischen Eskalationsmuster gipfeln: Memes, Kommentarfluten auf Reddit und YouTube, Aufrufe zum Boykott von Studios und sogar gezielte Angriffe auf Entwickler*innen, Journalist*innen oder Wissenschaftler*innen.

Dieses Muster wird häufig als sachliche Qualitätskritik inszeniert. So erscheint Anti-Diversity auf den ersten Blick wie ein Streit über Geschmack, obwohl gezielt versucht wird, die Diskussion zu rationalisieren. Oft geht es um angeblichen Realismus, historische Genauigkeit, »gutes Writing« oder die Frage, ob eine Figur zum Gesamtbild passt. Ein markantes Beispiel dafür lieferte die Debatte um den Weltkriegs-Shooter *Battlefield V* (2018): Die Darstellung kämpfender Frauen wurde von Teilen der Community mit Verweis auf historische Genauigkeit vehement abgelehnt. Der Anspruch, ein vermeintlich realistisches Bild des zweiten Weltkriegs zu zeichnen, wurde so zum zentralen Argument gegen mehr Diversität. Diese Form der Argumentation ermöglicht Abwertung, ohne offen menschenfeindlich zu wirken und wird damit anschlussfähig für die breitere Community. Tatsächlich handelt es sich jedoch um einen kulturellen Konflikt, in dem Zugehörigkeit, Macht und Deutungshoheit verhandelt werden.

Dieser Konflikt steht nicht isoliert im Gaming, sondern lässt sich einer breiteren Anti-Gender-Bewegung zuordnen, die weltweit gegen Gleichstellung, sexuelle Selbstbestimmung und queere Sichtbarkeit mobilisiert. Anti-Genderismus funktioniert dabei als Scharnier zwischen konservativen Milieus und rechtsextremen Akteur*innen (Lang 2015: 174). Er ist so attraktiv, weil er seine antidemokratischen Ziele geschickt hinter dem Bekenntnis zur Bewahrung traditioneller Werte wie vermeintlich natürlichen Geschlechterrollen oder der traditionellen Familie verbirgt. Diese bürgerlich-christlich anmutenden Schutzversprechen sind jedoch direkt mit rechtsextremen Narrativen vom moralischen Verfall und dem Zusammenbruch stabiler Identitäten verwoben (Claus, Virchow 2017: 314). Darüber hinaus sind antifeministische und anti-LGBTQIA+-Einstellungen tief verwurzelte Ressentiments, die nach wie

vor gesellschaftlich akzeptiert sind, da sie als weit weniger menschenfeindlich wahrgenommen werden als beispielsweise Rassismus oder Antisemitismus (Kaiser 2020: 203).

Wie das in der Praxis aussieht, zeigt die Diffamierungskampagne gegen Sweet Baby Inc. (SBI). Das Unternehmen berät Studios unter anderem in Bezug auf Storytelling und Charakterdarstellung. In toxischen Anti-Diversity-Umfeldern wurde diese Tätigkeit in ein Verschwörungsnarrativ überführt, wonach SBI als verdeckter Akteur Spiele »zwangspolisieren« würde (zu Verschwörungsnarrativen siehe Römer, Stumpf 2022; s. zur Kampagne gegen SBI auch Farokhmanesh 2024). Im Februar 2024 formierte sich, unter anderem über Reddit und Discord, eine breit angelegte Mobilisierung mit Boykott-Aufrufen und öffentlichem Druck auf Kooperationspartner*innen. Dabei ging es weniger um die tatsächliche Arbeit der Beratungsfirma, deren Einfluss auf die Gestaltung von Videospielen mit nur 16 Mitarbeitenden überschaubar ist. Stattdessen wurde das Unternehmen als Symbol für eine vermeintlich allgegenwärtige politische Korrektheit gesehen, welche die Gaming-Kultur fest im Griff habe. Verstärkt wurde diese Dynamik durch einen einflussreichen Steam-Kurator, der Spiele nicht anhand von Gameplay, Technologie oder Story bewertet, sondern sie mit kurzen Warnhinweisen versieht, sobald Diversität, LGBTQIA+-Darstellungen oder vermeintliche DEI-Elemente auftauchen (siehe dazu auch Liemann in diesem Band). Statt anzuerkennen, dass Spiele unterschiedliche Zielgruppen adressieren, wird Deutungshoheit über Gaming-Kultur beansprucht. Kauf- und Bewertungs-Entscheidungen werden dadurch in den Kulturkampf eingebettet: Abweichungen vom eigenen Erwartungshorizont gelten als Ausdruck einer gegen die eigene Gruppe gerichteten Agenda.

Im Gaming richtet sich Anti-Diversity insbesondere gegen Frauen, queere und nicht-weiße Menschen. Das passt zur Logik des Anti-Genderismus, der Gleichstellung als Statusverlust deutet. In dieser Erzählung wird Gleichberechtigung zu einem Nullsummenspiel: Die Sichtbarkeit und die Teilhabe anderer gelten als etwas, was Männern zwangsläufig weggenommen würde (Kaiser 2020: 110f.). Dahinter steht ein patriarchales Anspruchsdenken, das Repräsentation nicht als legitime Vielfalt liest, sondern als feindselige »Agenda« umdeutet, die gegen weiße, heterosexuelle Männer gerichtet sei (Kracher 2025: 99). So überrascht es nicht, dass in den letzten Jahren immer wieder Spiele wegen Frauenfiguren angegriffen wurden, die toxische Stimmen als »nicht weiblich genug« etikettierten. Dazu zählen neben dem bekannten Beispiel von Aloy in *Horizon Forbidden West* auch Debatten um Atsu in *Ghost of Yōtei* (2025; sie-

he Abb. 2) oder die Figur Abby in *The Last of Us II* (2020). In letzterem Fall wurde sogar die Schauspielerin zum Ziel massiver Anfeindungen, darunter auch Morddrohungen (Steinlechner 2020). Das ist kein ästhetisches Urteil, sondern eine Normsetzung: So hat eine Frau auszusehen. Wer davon abweicht, wird als ›woke‹ oder ›DEI‹ etikettiert und zur Zielscheibe von Abwertung gemacht.

Abb. 2: Screenshot eines Memes auf Twitter/X zur Darstellung von Atsu in *Ghost of Yōtei* (2025), erstellt am 29. Januar 2026.



If Ghost of Yotei made Atsu look like this, the game woulda sold a million copies by now



Ein Kommentar aus einer Reddit-Diskussion über *Ghost of Yōtei* illustriert diese Logik sehr anschaulich. Dort wird die weibliche Hauptfigur als ›hässlich beziehungsweise durchschnittlich aussehend‹ kritisiert. Zugleich wird behauptet, dies geschehe bewusst, um den sogenannten ›male gaze‹ zu verwei-

gern. Das Spiel richte sich damit gegen das heterosexuelle männliche Publikum und stelle eine Form von ›Misandrie‹ dar. Darüber hinaus wird die Figur als Beleg dafür interpretiert, dass ›Wokeness‹ und Feminismus gezielt in das Spiel implementiert worden seien (Reddit-Kommentar 2025)⁸. Der Standard bleibt der heterosexuelle männliche Blick: Wird diesem nicht entsprochen, gilt dies bereits als politisches Statement gegen Männer. Im Mittelpunkt steht dabei keine ästhetische Entscheidung, sondern die Verteidigung eines impliziten Anspruchs auf die sexualisierte Darstellung weiblicher Körper.

Anti-Diversity reicht damit weit über die Szene hinaus und funktioniert als ein strategisches Anschlussfeld. Gerade weil es den Anschein einer Diskussion über Geschmack oder ›Qualität‹ erweckt, wird Abwertung schnell massentauglich: Menschen können sich daran beteiligen, ohne sich selbst als politisch positioniert zu sehen. Diese Dynamik schafft ein Umfeld, in dem rechtspopulistische und rechtsextreme Akteur*innen andocken, Konflikte zuspitzen und Mobilisierung organisieren können. Damit wird Anti-Diversity zu einem verbindenden Element zwischen toxischen und rechtsextremen Communitys. Es schafft gemeinsame Feindbilder, stabilisiert ein ›Wir gegen die anderen‹ und forciert die Normalisierung extrem rechter Narrative in gemäßigeren reaktionären und toxischen Bereichen der Gaming-Kultur.

5. Fazit und ein Plädoyer für digitale Zivilcourage im Gaming

Nachdem dieser Artikel einen umfassenden Überblick über die toxischen und rechtsextremen Agitationen im Gaming gegeben hat, werden nun abschließend Handlungsempfehlungen formuliert, die Spieler*innen, Pädagog*innen, aber auch Politik, Behörden und Plattformen adressieren. Für alle Akteursgruppen gilt der zentrale Grundsatz, dass Gaming und die Videospiele-Kultur nicht verallgemeinernd stigmatisiert werden sollen. Das, was dieser Artikel offeriert hat, ist nur ein Teil der Gaming-Welt. Ein Teil über den gesprochen werden muss und dessen Probleme nicht ignoriert oder bagatellisiert gehören. Aber eben auch ein Teil der nicht stellvertretend für eine breite, heterogene und an vielen Stellen diversifizierte Kultur formuliert werden darf. Was sollte jetzt aber passieren?

8 Kommentar unter einer Reddit-Diskussion zum Spiel *Ghost of Yōtei*, Screenshot dokumentiert am 04.12.25, eigene Übersetzung.

Vor allem für Pädagog*innen sollte zentraler Ansatz sein, mit Spieler*innen in den Austausch zu gehen und über Erlebnisse in Gaming-Welten, aber auch über Erfahrungen mit den toxischen Abgründen eben dieser zu sprechen. Dies passiert zu selten und ein differenzierter Austausch über Gaming wird durch Killerspiel-Narrative oder fehlendes Verständnis von Pädagog*innen erschwert. Neben dem Erfahrungsaustausch können Games auch stärker in Vermittlung von Werten und Bildung genutzt werden. Das Modellprojekt ›Games macht Schule‹ der Stiftung digitale Spielekultur hat verdeutlicht, dass der Einsatz von Videospielen in unterschiedlichsten Lehreinheiten einen positiven Effekt auf Motivation und Wissensvermittlung hat. So berichten 71 % der befragten Lehrer*innen das besonders bei passiveren Schüler*innen der Einsatz von Games in der Bildung Motivation steigerte und den Zugang zu Unterrichtsthemen erleichterte (Stiftung digitale Spielekultur 2026). Neben dem Einsatz mit zahlreichen Serious Games gelingt dies auch mit Unterhaltungsgames, wenn beispielsweise über die Darstellung der Hauptfiguren oder die Repräsentation unterschiedlicher Identitäten im Charaktereditor diskutiert wird.

Auch die Spieler*innen selbst können aktiv gegen rechtsextreme Instrumentalisierungsversuche im Gaming vorgehen. Zentral ist an dieser Stelle, dass nicht einfach weggeschaut wird, wenn Menschenfeindlichkeit im Gaming formuliert wird. So können Meldestrukturen auf den jeweiligen Plattformen oder externe Anlaufstellen wie die ›Meldestelle REspect!‹ (vgl. respect im netz e.V.) genutzt werden. Das ist relevant, um Plattformen, aber auch Sicherheitsbehörden auf das Problem hinzuweisen und mögliche Kapazitäten auszubauen. Außerdem können Spieler*innen Gegenrede leisten und z. B. rassistischen, antisemitischen oder queerfeindlichen Äußerungen widersprechen. Die Idee ist dabei weniger, die Person zu überzeugen, die Hassrede geäußert hat, sondern zu unterstreichen, dass gruppenbezogene Menschenfeindlichkeit nicht widerstandslos stehen gelassen wird. Dies ist vor allem auch für Betroffene wichtig, die zusätzlich mit solidarischen Kommentaren unterstützt werden können.

Vor allem Content-Schaffende mit Reichweite sollten dabei stärker ihre Reichweitenverantwortung nutzen. Das bedeutet nicht, dass sie immer über politische Sachverhalte diskutieren müssen, schließlich kann Gaming auch ein Ort des berechtigten Eskapismus sein. Allerdings helfen klare Baselines und Grundsätze, wie beispielsweise eine Nulltoleranz bei rechtsextremen Beleidigungen im eigenen Chat oder auch die Formulierung eigener Guidelines wie ›Sexismus, Transfeindlichkeit und andere Menschenverachtung hat keinen Platz im Twitch Chat und wird nicht geduldet‹. Eine Haltung fungiert

damit auch ohne auf komplexe Debatten und Konflikte im Detail eingehen zu müssen.

Zusätzlich müssen Videospieleplattformen stärker die eigenen Richtlinien durchsetzen und proaktiver hasserfüllten Content moderieren. Plattformen wie Roblox oder Discord haben in den letzten Jahren ihre Bemühungen etwas verstärkt, auch wenn sich hier nach wie vor mit wenigen Klicks toxische Inhalte finden lassen. Auf anderen Plattformen wie Steam oder Twitch wird diese Problemlage noch deutlicher. Es wird Zeit, dass Plattformen hier stärker mit nationalem und europäischem Recht konfrontiert werden und eigene Guidelines effektiv umsetzen. So wird beispielsweise das europäische Gesetz gegen Hass im Internet, der *Digital Services Act (DSA)* im Gaming nur punktuell umgesetzt. Während die Plattform Twitch ihre Meldeverfahren zumindest formell angepasst hat und Meldungen nach dem DSA möglich sind, greift die Gesetzgebung auf anderen Plattformen wie Steam (noch) nicht (Twitch 2026).

Dieser Beitrag hat gezeigt, dass toxische und rechtsextreme Gaming-Communitys auf zu Teilen ähnliche Narrative und Abwertungsgrundlagen zurückgreifen. Die extreme Rechte versucht dabei gezielt Debatten zu bespielen und eine breite Masse an Gamer*innen für die eigene Ideologie zu erreichen. Obwohl Rechtsextremismus im Gaming eine Minderheit ist, unterstreichen Beispiele wie das Attentat am Münchener Olympia Einkaufszentrum und Gamergate, aber auch Diskurse um ›Diversität im Gaming‹ das Gefahrenpotenzial, welches auch vor Videospiele-Welten keinen Halt macht. Es bedarf mehr digitaler Zivilcourage und einem entschiedenen Auftreten unterschiedlichster Akteursgruppen, um die Diversität und Pluralität im Gaming zu erhalten. Das sollten auch jene Spieler*innen verinnerlichen, die auch nicht alle progressiven Errungenschaften im Gaming teilen. Ansonsten wird die laute Minderheit zu einer noch lauterer Mehrheit und die vielfältige Gaming-Kultur zu einem braunen Pixelbrei.

Literaturverzeichnis

- Adams, Philippa R.; Sommers, Maria; Scholl, Ben; Lusoli, Alberto; Chun, Wendy Hui Kyong (2025): The revenge of ›tribal‹ toxic masculinity: hate raids on Twitch. In: *Journal of Visual Culture*. 24, 1. URL: <https://doi.org/10.1177/14704129251327916>.
- Anti-Defamation League (2024): Hate is No Game: Hate and Harassment in Online Games 2023. In: ADL. URL: <https://www.adl.org/resources/>

- report/hate-no-game-hate-and-harassment-online-games-2023 [Zugriff 27.01.26].
- Ballweber, Jana (2024): Zehn Jahre »Gamergate« – Eine Blaupause für rechten Kulturkampf. In: Evangelische Zeitung. URL: <https://evangelische-zeitung.de/zehn-jahre-gamergate-eine-blaupause-fuer-rechten-kulturkampf/> [Zugriff 09.03.26].
- Belltower-News (2026): Lexikon: Woke/Wokies/Wokeness (Code). In: Belltower-News. URL: <https://www.belltower.news/lexikon/woke-wokies-wokeness-code> [Zugriff 29.01.26].
- Belltower-News (2012): Rechtsextremismus – Welche Komponenten gehören zur rechtsextremen Ideologie? In: Belltower-News. URL: <https://www.belltower.news/welche-komponenten-gehoren-zur-rechtsextremen-ideologie-34536/> [Zugriff 09.03.26].
- Bernstein, Martin (2018): Das Netzwerk der Todesschützen. In: Süddeutsche Zeitung [online]. URL: <https://www.sueddeutsche.de/muenchen/oez-attentaeter-amoklaeufer-kontakt-1.3979444> [Zugriff 27.01.26].
- Bitkom e. V. (2025): Frauen zocken täglich knapp 20 Minuten mehr als Männer. In: Bitkom. URL: <https://www.bitkom.org/Presse/Presseinformation/Gaming-Deutschland-2025> [Zugriff 27.01.26].
- Bundesamt für Verfassungsschutz (BfV) (2024): Rechtsextremismus im Internet: Gefahren digitaler Agitation und Radikalisierung. In: BfV. URL: https://www.verfassungsschutz.de/SharedDocs/publikationen/DE/rechtsextremismus/2024-09-rechtsextremismus-im-internet.pdf?__blob=publicationFile&v=2 [Zugriff 20.01.26].
- Bundesministerium des Inneren für Heimat (2024): Gemeinsam für Demokratie und gegen Extremismus – Strategie der Bundesregierung für eine starke, wehrhafte Demokratie und eine offene und vielfältige Gesellschaft. In: BMI. URL: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/ministerium/BMI24021.pdf?__blob=publicationFile&v=8 [Zugriff 09.03.26].
- Cai, Jie; Chowdhury, Sagnik; Zhou, Hongyang; Wohn, Donghee Yvette (2023): Hate raids on Twitch: Understanding real-time human-bot coordinated attacks in live streaming Communitys. In: Proceedings of the ACM on Human-Computer Interaction. 7, 2. S. 1–28. URL: <https://doi.org/10.1145/3610191>.
- Claus, Robert; Virchow, Fabian (2017): The Far Right's Ideological Constructions of »Deviant« Male Sexualities. In: Gender and Far Right Politics in Europe. Hg. v. Michaela Köttig, Renate Bitzan u. Andrea Pető. London: Pal-

- grave Macmillan. S. 305–319. URL: https://doi.org/10.1007/978-3-319-43533-6_20.
- Dietrich, Martin (2022): Die Erben von GamerGate – Gaming-Youtuber beeinflussen den Diskurs über Videospiele. In: Bundeszentrale für politische Bildung (Hg.). URL: <https://www.bpb.de/themen/kultur/digitale-spiele/504985/die-erben-von-gamergate-gaming-youtuber-beeinflussen-den-diskurs-ueber-videospiele/> [Zugriff 29.01.26].
- Dreßler, Antonia (2025): Zwischen Gamergate und Woke: Der interessante Fall von Kingdom Come: Deliverance. In: PC Games. URL: <https://www.pcgames.de/Kingdom-Come-Deliverance-2-Spiel-74644/Kolumnen/Zwischen-Gamergate-und-Woke-Der-interessante-Fall-von-Kingdom-Come-Deliverance-1464690/3/> [Zugriff 09.03.26].
- Ebner, Julia (2019): Radikalisierungsmaschinen: Wie Extremisten die neuen Technologien nutzen und uns manipulieren. Berlin: Suhrkamp Nova.
- Farokhmanesh, Megan (2024): The Small Company at the Center of Gamergate 2.0. In: WIRED. URL: <https://www.wired.com/story/sweet-baby-video-games-harassment-gamergate/> [Zugriff 29.01.26].
- Fieierabend, Sabine; Rathgeb, Thomas; Gerigk, Yvonne; Glöckler, Stephan (2025): JIM-Studie 2025: Jugend, Information, Medien. Basisuntersuchung zum Medienumgang 12- bis 19-Jähriger. In: Medienpädagogischer Forschungsverbund Südwest (Hrsg.). URL: https://mpfs.de/app/uploads/2025/11/JIM_2025_PDF_barrierearm.pdf [Zugriff 29.01.26].
- Heider, Matthias (2025). Rechtsextreme Aktivitäten auf Gaming(-nahen)-Plattformen. In: Handbuch Gaming & Rechtsextremismus. Hg. v. Aurelia Brandenburg, Linda Schlegel u. Felix Zimmermann. Bonn: Bundeszentrale für politische Bildung. S. 175–183.
- Kaiser, Susanne (2020): Politische Männlichkeit. Wie Incels, Fundamentalisten und Autoritäre für das Patriarchat mobilmachen. Frankfurt a.M.: Suhrkamp.
- Kracher, Veronika (2025): Misogynie & Antifeminismus. In: Handbuch Gaming & Rechtsextremismus. Hg. v. Aurelia Brandenburg, Linda Schlegel u. Felix Zimmermann. Bonn: Bundeszentrale für politische Bildung. S. 97–107.
- Lang, Juliane (2015): Familie und Vaterland in der Krise. Der extrem rechte Diskurs um Gender. In: Anti-Genderismus: Sexualität und Geschlecht als Schauplätze aktueller politischer Auseinandersetzungen. Hg. v. Sabine Hark u. Paula-Irene Villa. Bielefeld: transcript. S. 167–181.
- Mattsson, Christer; Johansson, Thomas (2022): The Neo-Nazi Doxa, the Deep State, and COVID-19: Neo-Nazis' Ambivalent Relations to Big Pharma and

- the Anti-Vaccine Movement. In: *Social Identities*. 28, 6. S. 714–727. URL: <https://doi.org/10.1080/13504630.2022.2138310>.
- Reddit (2025): If Atsu in Ghost of Yotei actually looked like this... In: Reddit. URL: https://www.reddit.com/r/KotakuInAction/comments/1o99dh7/if_a_tsu_in_ghost_of_yotei_actually_looked_like/ [Zugriff: 13.05.26].
- respect im netz e.V. (o.J.): Respect im Netz (Website). URL: <https://respectimnetz.de/> [Zugriff 29.01.26].
- Rother, Joachim; Gerke, Jessica; Brandenburg, Aurelia; Wimmer, Jeffrey (2025): Spielräume für Demokratie. Potenziale und Spannungsfelder im Gaming. In: Bertelsmann Stiftung (Hg.). URL: <https://www.bertelsmann-stiftung.de/de/themen/aktuelle-meldungen/2025/august/das-demokratische-potenzial-von-gaming-communitys-besser-nutzen> [Zugriff 29.01.26].
- Römer, David; Stumpf, Sören (2022): Verschwörungstheorien – und wie sie sprachlich glaubhaft gemacht werden. In: *Verschwörungstheorien. Eine Publikation von Gegen Vergessen – Für Demokratie e.V.* Hg. v. Klaus Müller u. Christopher Kirchberg. Berlin : Gegen Vergessen – Für Demokratie e.V. S. 60–89.
- Röther, Valentina; Reszke, Paul; Böhm, Felix (2024): #Klima – Mediale Logiken eines Diskurses. In: *Aptum Zeitschrift für Sprachkritik und Sprachkultur*. 20, 3. S. 241–257. URL: https://doi.org/10.46771/9783967694512_1.
- Salzborn, Samuel (2020): Die »Neue Rechte« im bundesdeutschen Rechtsextremismus. In: *Die Neue Rechte. Hintergründe und Hauptelemente neurechten Denkens*. Hg. v. Daniel Engel, Adriana Lanza u. David Meier-Arendt. Darmstadt: TU Prints. S. 21–30.
- Schlegel, Linda (2021): *The Role of Gamification in Radicalization Processes* (Working Paper 1/2021). Berlin: modus|zad.
- Schölzel, Stephan (2025): Jugendarbeit. In: *Handbuch Gaming & Rechtsextremismus*. Hg. v. Aurelia Brandenburg, Linda Schlegel u. Felix Zimmermann. Bonn: Bundeszentrale für politische Bildung. S. 283–290.
- Spiegel (2019): Seehofer will Gamerszene stärker beobachten – und erntet Hohn. In: *Spiegel Netzwelt*. URL: <https://www.spiegel.de/netzwelt/games/horst-seehofer-shitsstorm-fuer-bemerkung-zur-gamerszene-a-1291347.html> [Zugriff 30.03.26].
- Steinlechner, Peter (2020): Entwickler von The Last of Us 2 bekommen Morddrohungen. In: *Golem.de*. URL: <https://www.golem.de/news/naghty-dog-entwickler-von-the-last-of-us-2-bekommen-morddrohungen-2007-149470.html> [Zugriff 27.01.26].

Stiftung Digitale Spielekultur gGmbH (2016): Initiative: Games machen Schule. In: Stiftung für Digitale Spielekultur. URL: <https://www.stiftung-digitale-spielekultur.de/project/games-schule-unterricht/> [Zugriff 29.01.26].

Twitch (2026): Informationen zum Gesetz über digitale Dienste. In: Twitch. URL: <https://safety.twitch.tv/s/article/Digital-Services-Act-Information?language=de> [Zugriff 09.03.26].

Abbildungsverzeichnis

Abbildung 1: Screenshots eines Memes über GTA 6 (angekündigt für 2026) in einem YouTube Short. Der dazugehörige Kommentarbereich voller ›edgy‹ und toxischer Kommentare, erstellt am 14.01.26.

Abbildung 2: Screenshot eines Memes auf Twitter/X zur Darstellung von Atsu in Ghost of Yōtei (2025). URL: <https://x.com/mrphillipchan/status/1976006553049366943>, erstellt am 29.01.26.

»Crunch as Curriculum«

Toxic Learning Culture and the Aesthetics of Overwork in Concept Art Education for the Games Industry

Marion Plank

Abstract Art schools that train concept artists for the video game industry face unique challenges, caused by the inherent complex nature of artistic traineeship. They must teach technical skills and prepare students for the realities of creative work in the video game industry, including tight deadlines, competition and pressure to perform.

A crowded job market and an overflow of graduates have created learning environments that often mirror the industry's own toxic habits, harming students' well-being and creative confidence during and after their studies. In this essay I ask: What would it mean to teach industry-preparedness in the game concept art space without reproducing toxicity?

Based on interviews with educators and students, as well as theories of the hidden curriculum, platformization, and care ethics, I show how ideals of passion and productivity in art schools reflect the game industry's normalization of overwork. The essay outlines five principles for a healthier curriculum: sustainability over performance, structural reform, teachers as protectors, differentiation as philosophy, and empathy as a diagnostic tool.

Keywords Games, Education, Concept Art, Overwork, Pressure

Abstract Kunsthochschulen, die Concept Artists für die Videospielindustrie ausbilden, stehen vor besonderen Herausforderungen, die aus der komplexen Natur künstlerischer Ausbildung resultieren. Sie müssen einerseits technische Fähigkeiten vermitteln und andererseits Studierende auf die Realitäten kreativer Arbeit in der Spielebranche vorbereiten, darunter enge Fristen, Wettbewerb und Leistungsdruck.

Ein überfüllter Arbeitsmarkt und ein Überangebot an Absolvent*innen haben Lernumgebungen hervorgebracht, die häufig die toxischen Praktiken der Branche selbst widerspiegeln und dadurch das Wohlbefinden sowie das kreative Selbstvertrauen der Studierenden während und nach ihrem Studium beeinträchtigen. Vor diesem Hintergrund stellt der Aufsatz die Frage: Was würde es bedeuten, im Bereich der Game Concept Art auf die Anforderungen der Industrie vorzubereiten, ohne dabei toxische Strukturen zu reproduzieren?

Auf Grundlage qualitativer Interviews mit Lehrenden und Studierenden sowie unter Einbezug von Theorien zum Hidden Curriculum, zur Plattformisierung und zur Fürsorgeethik zeigt der Beitrag, wie Ideale von Leidenschaft und Produktivität an Kunsthochschulen die Normalisierung von Überarbeitung in der Spieleindustrie widerspiegeln. Abschließend werden fünf Prinzipien für einen gesünderen Lehrplan skizziert: Nachhaltigkeit statt Leistungsorientierung, strukturelle Reformen, Lehrende als Schutzinstanz, Differenzierung als pädagogische Philosophie sowie Empathie als diagnostisches Werkzeug.

Schlagworte Games, Studium, Concept Art, Überarbeitung, Leistungsdruck

1. Prologue / Enter the Artist

Many young art students dream of a career in game art. Among the most coveted roles is that of the concept artist, a profession that attracts creatively gifted individuals seeking a commercial path in the games industry. Concept artists are responsible for visualizing characters, props, and environments, shaping the first impression of a game long before it is playable. Their work is foundational: it influences not only the game's aesthetics but often its overall creative direction.

Today, the concept artist is widely recognized as a key contributor to the creative development of games and other entertainment media. They are expected to master fundamentals such as anatomy, perspective, colour theory, and visual design, while also working within technical constraints to ensure their ideas can be implemented seamlessly into production pipelines (CG Spectrum 2025; Vancouver Film School 2025). To qualify for one of the rare jobs, aspiring artists often attend learning institutions that promise to prepare them, with densely structured curricula comprising skill-focused

modules like sketching, visual communication, 2D and 3D as well as sculpting and product design (e.g. Sawyer 2018).

This essay investigates how toxic or harmful learning cultures within game art education normalize overwork and creative self-sacrifice in aspiring concept artists. Through qualitative interviews with students and educators, it examines the psychological, institutional, and discursive mechanisms that frame overwork as a necessary step toward professional success. By analysing both personal experiences and broader educational practices, the essay aims to expose the implicit curricula that prepare students for a precarious industry shaped by crunch culture, visibility pressures, and fears of obsolescence due to AI. The goal is not only to document these dynamics, but to critically question what it truly means to be 'industry-ready', and to imagine alternatives.

2. Methodology

The interview material is decoded through a qualitative-interpretive lens with a focus on thematic condensation rather than full categorical coding. A total of five interviews were conducted via Discord video calls in German and English, each of which was recorded and fully transcribed and guided by a semi-structured interview framework. The German interviews were translated into English by the author. All interviews cited in this essay were conducted by the author between September and October 2025 with current or former students and educators in game art education. The goal is not statistical saturation, but to illustrate how toxic pedagogical narratives are experienced and reproduced within the educational context of game concept art. Interviews with students aim to explore individual experiences of workload, psychological strain, and institutional expectations, while interviews with educators provide insight into the curricular logic, pedagogical aims, and assumptions about industry readiness.

Seen from both perspectives, it becomes possible to trace how toxic learning cultures form, endure, and occasionally come under critique within today's educational systems.

The following text first outlines key theoretical concepts, before analyzing both structural and individual dimensions of exploitative learning cultures through qualitative interviews with students and educators. It concludes by proposing principles for a healthier approach to teaching in the field of game art.

3. Limitations

This study is subject to several limitations. The small sample size and qualitative design do not allow for generalizable conclusions, but instead aim to provide insight into subjective experiences and recurring patterns within a specific educational context. The interview material is shaped by self-reporting and retrospective reflection, which may be influenced by personal bias, selective memory, or individual positioning within the field. Furthermore, as the interviews were conducted, translated, and interpreted by the author, the analysis is inevitably informed by the researcher's own perspective and disciplinary background. While care was taken to preserve meaning during translation, nuances may have been altered in the process. Finally, the study focuses primarily on European game art education and may not fully account for global variations in institutional structures and industry relations.

4. Theoretical Lense: Crunch as Curriculum

As Jesse Schell notes, the early integration of concept art during game development can provide a unifying vision for the player experience, aligning mechanics, narrative, and design from the outset (2019: 388).

In the early days of video games, art and graphics were typically dictated by technical limitations (e.g., pixel count, colour palette restrictions). Artists at the time were primarily tasked with promotional materials such as box art, rather than in-game visuals (RMCAD 2024; GameDeveloper 2023). This changed with the advent of more powerful consoles and growing consumer expectations for visually immersive experiences. (RMCAD 2024; Hawey 2025: 2). However, the emergence of the specialized role of the Concept Artist is not linear but developed gradually at the intersection of illustration, industrial design, and (film) production. Artists such as Syd Mead, who designed not only individual objects but entire visual worlds for films like *Blade Runner*, exemplify an early form of what would later become standardized as concept art. In parallel, artists like Moebius operated across comics, film, and video games, contributing to visual development processes before the role was formally institutionalized. By the late 1990s, the increasing complexity of digital production pipelines led to the emergence of dedicated “conceptual designers” within game studios, marking the consolidation of concept art as a specialized profession which then further branched into subfields

such as character design, creature design, prop design, vehicle design, and environment concept art. The increasing specialization of the role led to the development of dedicated training programs, online courses, and private schools aimed at preparing young artists for the demands of the industry. No longer was it sufficient to transition into games from adjacent fields like illustration or comics. Industry expectancies shifted: junior artists were now expected to possess not only artistic talent but also a deep understanding of production pipelines, visual storytelling, and software workflows.

By the late 2000s and early 2010s, a wave of specialized schools emerged, portfolio-driven programs explicitly marketed as pathways to ‘industry readiness’ (FZD School of Design 2010; Concept Design Academy 2012). These schools often promised direct access to industry mentors, streamlined workflows, and a competitive edge in the job market. The booming demand for visually appealing and story-driven content created pressure to train faster and more efficiently, leading to high-stakes, fast-paced learning environments.

One of the common justifications for these intensive programs and their excessive workloads is the aestheticization of overwork in the games industry itself, euphemistically referred to as *crunching*, or simply *crunch*.

Months of overtime, 100-hour workweeks, and 12+ hour days do not sound like the characteristics of a ‘dream job’; they sound like red flags that should leave anyone with an alternative heading in the opposite direction. Yet within the video game industry, tales of ‘crunch’- extended periods of drastic overtime-are commonplace (Cote, Harris 2023: 609).

As alleged thorough preparation for this culture of long workdays, the educational programs that prepare the students for their life in the video game industry often reflect these normalized ideas, conditioning them to accept the idea that “it is nearly impossible to produce exciting, memorable games without at least some crunch” (ibid.: 610).

5. The Hidden Syllabus

In a study by Sellers and Villanueva Alarcón (2023) the hidden curriculum is defined as “the unwritten, unofficial, and often unintended assumptions, lessons, values, beliefs, attitudes, and perspectives not openly acknowledged in any given environment” (Sellers, Villanueva Alarcón 2023: 176). They argue that through institutional norms, interpersonal interactions, and structural prac-

tices, educational systems convey powerful messages about what counts as legitimacy, competence, or belonging. These messages are not part of the formal syllabus but are absorbed by students through everyday signals, thus shaping their identity, behaviour, and perceptions of success. In the context of creative education, the hidden curriculum can thus include toxic norms such as glorification of overwork, self-sacrifice, and performative productivity as prerequisites for 'serious' artistry.

In the interview with student V., they state:

There are no breaks during your studies – as soon as submissions are ready, the next project phase begins. The reality of the job is more flexible and human. But crunch was often glorified during studies. One lecturer openly said that crunch is part of the job. Others who didn't come from industry never addressed it. (V-01, 2025)¹

Student L. reports a hidden curriculum in their studies: "The requirements are not very high, but if you only meet the requirements, then it's often too little, I'd say, you always have to add a little more." (L-01, 2025)

5.1 Unsustainable Learning & Burnout

The success of any educational institution depends not only on academic standards but also on students' motivation, fulfilment, and emotional well-being. Yet research consistently shows that higher education environments expose students to significant psychological strain, with many experiencing stress, emotional difficulties, and declining well-being during their studies (Douwes et al. 2023; Majerová, Sokolová 2025). University study often coincides with major life transitions such as relocation, social detachment, and the sudden demand for autonomous decision-making, all of which heighten psychological strain (Worsley et al. 2021). When programs prove more demanding, competitive, or misaligned with students' expectations, the result is a pattern of exhaustion, reduced vitality, and declining educational satisfaction (Jin, Ye 2022: 4). Within creative education, these stressors are amplified by aesthetic ideals that conflate productivity with passion and endurance

1 To preserve anonymity, participants are identified by coded pseudonyms and neutral pronouns, e.g. V-01, E-01.

with excellence. These are conditions under which burnout becomes not an exception but a normalized mode of learning.

'Toxic learning cultures' in educational contexts refer to the systemic elevation of competitive productivity, meritocratic values, and perpetual performance pressure at the expense of psychological well being and creative autonomy. Recent literature on educational performance cultures highlights how school environments can normalize 'toxic productivity', linking meritocracy and chronic stress to negative mental health outcomes and reduced learning satisfaction (Sukmana et al. 2025). Conceptually related research on toxic organisational climates shows that environments characterized by harassment, bullying, and toxic leadership contribute to increased stress, burnout, and diminished emotional well-being in institutional settings (Rasool et al. 2021).

"In the first three semesters, I never slept on the last day of project submission deadlines" reports L. in the interview, a student with highly positive feedback from peers and educators for their work (L-01, 2025). "[A]fter that come two weeks where I can't do anything at all. There is just nothing left," they continue, indicating reduced ability to create and a need to recover from the strenuous studying.

"Many have become mentally or physically ill" explains V. (V-01, 2025), while another interviewee recalled that what once felt like "stress-free passion" became a voice in their head repeating external expectations (J-01, 2025). These narratives align with research defining art block as the affective endpoint of chronic performance pressure, when creation, once joyful, turns into a site of exhaustion and self-doubt (Głaziewicz, Golonka 2024). Educator S. states in the interview: "[...] they are very inhibited to just try things out, to make mistakes; making mistakes is generally not mentally possible for many" (S-01, 2025).

As many programs offer only short breaks (Jin, Ye 2022) or, due to the timing of the student projects, no breaks at all during the term, this can quickly turn into a vicious cycle of blocked creativity and output. "Despite the desire/need to create something or express one's emotions, one cannot muster enough energy or space to achieve it" (Głaziewicz, Golonka 2024).

Drawing on sociological theories of practice, toxicity in creative education can be understood as a feedback loop between institutional structures and student subjectivity, in which external expectations are internalized and reproduced through everyday practices (Smith 2020). Hidden expectations and the fear of failure create a cycle in which external pressure becomes self-discipline, self-discipline becomes overwork, and overwork leads to burnout and creative paralysis, only for the cycle to begin again.

While the present study focuses on education, it resonates with broader industry dynamics. The Gamergate movement, for example, brought to light not only the misogyny in gaming communities but also the denial and resistance toward discussions of structural toxicity (Chatzakou et al. 2017). Similarly, global discourses around burnout, gendered labor, and precarious work have exposed how affective exploitation is both denied and glorified, especially in industries that frame passion as privilege. These public debates contextualize the institutional dynamics explored here, situating the classroom as an early site of normalization for broader cultural toxicity.

5.2 Performativity & Visibility in Creative Education

The value of creative education is not only defined by what students produce, but by how visibly they perform their commitment. Within studios and online spaces, artistic identity becomes a public display of endurance, where fatigue and self-sacrifice signify authenticity. As one interviewee (E-01, 2025) recalled, “We tried to show it like, ‘Look at me. I ruined my health.’” Such expressions of suffering act as moral signals in a culture defined by constant visibility. In this sense, the overworked art student is not merely responding to external evaluation but enacting a learned mode of self-policing in which visibility itself becomes a moral demand.

Within today’s platform-driven creative culture, visibility has become synonymous with professional survival. Sites like ArtStation and Instagram celebrate relentless productivity, turning creative practice into an ongoing performance of activity. As one interviewee noted, “Even though I had a full-time job, I stayed active on Discord, LinkedIn, and Twitter [...]. I kept working after hours to maintain my portfolio” (V-01, 2025). Another reflected that “You see what others are doing and how far you are from that” (J-01, 2025), describing how constant exposure fuels comparison and self-doubt. These accounts reveal how the demand to remain visible transforms artistic identity into a form of unpaid, perpetual self-promotion.

For example, ArtStation’s and other platform’s feed designs can pressure artists to post continuously to remain discoverable. In this context, the act of resting or experimenting invisibly becomes risky: absence can mean irrelevance. In the interview one student says: “You constantly compare yourself. Some people post new stuff every day, and you think: Why am I so slow?” (L-01 2025).

Similarly, David Nieborg and Thomas Poell (2018) describe this condition as part of the broader platformization of cultural production, where visibility metrics and algorithmic ranking mechanisms reconfigure how creative labour is valued. What was once an internal drive toward artistic excellence is now mediated by engagement data and social proof. The interviewees' sense of constant comparison mirrors this shift: productivity and worth are measured not by creative insight but by public visibility. "I noticed that I no longer create for myself, but so that it looks good when I show it." (L-01 2025)

Building on this, T. L. Taylor (2018) argues that live streaming platforms transform play into a form of performative and affective labour, in which creators are expected to display authenticity, emotional engagement, and passion as part of their work. For students and early-career artists, who show or livestream their artwork, this 'performance of care' can become indistinguishable from self-exploitation: the display of dedication feels like it must be continuous, emotionally charged, and socially validated.

Together, these dynamics illustrate how digital platforms extend and amplify the pressures of art education. The online gallery becomes a stage for competitive self-branding, where exhaustion, productivity, and visibility converge into a single, moralized image of success.

5.3 The Aesthetics of Overwork & Industry Mythmaking

The culture of performative visibility is reinforced by educators who in some cases, knowingly or not, model exhaustion as a professional virtue. Yet this modeling is often based on a mythology of industry experience rather than its lived reality. Student interviewees noted that many teachers were no longer active in production or had left the industry precisely because of its unsustainable conditions. One student recalled: "Teachers said: 'When I was your age, I worked harder. If you want to be cooler, you need to work harder'," (E-01, 2025) while another added that "[o]nly one lecturer had actually worked in the industry – he was the one who said crunch is just part of it" (V-01, 2025). These accounts suggest that the educator's authority frequently rests on a secondhand image of the professional world, that is even inherited from a system they may have escaped. In this sense, teachers become false mirrors: they reproduce the rhetoric of endurance without the corresponding context of labour relations, risk, or remuneration. One student marked in the interview (L-01 2025): "It's something that gets glorified. People say, 'I worked so hard,' and next semester everyone does it again."

Within creative education, the rhetoric of being industry ready functions as an invisible moral framework that extends beyond the acquisition of technical skills. Under the guise of professional preparation, it transmits a set of unspoken expectations: constant availability, emotional resilience, and the willingness to endure hardship as proof of commitment. One student described this internalized pressure bluntly: “I must now perform well... otherwise I won’t get a job later,” (L-01, 2025). These narratives illustrate how the ideology of readiness legitimizes exhaustion as a necessary precondition for employability. As Farrier and Whittaker argue, this pervasive mindset

so strongly influences the possibilities for students, they can be anxious about being ‘industry ready’ before they’ve had the chance to explore or define fully what kind of practitioners they want to become (2025).

Likewise, a recent Times Higher Education commentary notes that the emphasis on employability success has narrowed creative learning into “a competitive race for visibility and performance” (Coffield, Markham 2024), masking structural inequities behind personal responsibility. Even studies in design pedagogy warn against this reduction, arguing that creative education must preserve space for “risk-taking and productive failure” (Henriksen et al. 2021). The pursuit of industry readiness thus becomes part of the hidden curriculum (Sellers, Villanueva Alarcón 2023): a regime of affective labour and self-discipline that is never listed in the syllabus yet silently governs what it means to be a ‘successful’ concept art student.

6. Outlook: What a Non-Toxic Curriculum Could Look Like

When reflecting on the systemic harmful traits that permeate higher concept art education, it becomes essential to propose tangible responses. This final section outlines five key starting points, supported by concrete methods, that educators can integrate into their daily teaching. Small shifts in individual practice can begin to repair systemic failures and model alternative, sustainable ways of learning.

As a professor and curriculum designer, I ask: What would it mean to teach industry-preparedness without reproducing harm? The following principles are directly derived from the patterns identified in the interview material and theoretical analysis above.

6.1 From Performance to Sustainability

A sustainable approach reframes success in professional concept art as longevity: the capacity to maintain creative health across years, not just semesters. It values recovery as part of the learning cycle. For students and teachers alike, managing energy rather than time becomes central. Students who understand how to renew physical, emotional, and cognitive energy are more likely to sustain meaningful practice beyond graduation.

Shifting from performance to sustainability means recalibrating assessment criteria, which in some schools is already reflected in the grading system. Educator S. states in the interview: “I think more exercises would be quite good, [...], so that students realize, okay, making mistakes is perfectly fine.” (S-01, 2025) Reflection, iteration, and rest are recognized as forms of progress. Actively asking for help or even extension of deadlines in this scenario are not treated as mistakes and punished but praised as it takes courage and foresight from the student.

6.2 Structural change

Lasting reform in concept art education requires structural change. The systemic issues identified throughout this essay (compressed timelines, lack of industry alignment) cannot be solved by better intentions alone.

One promising direction lies in dual study programs that integrate sustained studio work, not as an add-on or short internship, but as an embedded component of the curriculum. Time spent within a functioning game studio offers students insight into real production cycles, communication hierarchies, and workflows while allowing them to apply academic reflection to practical challenges. Such programs bridge the gap between academic ideals and industry realities, reducing the shock of transition after graduation.

A second structural lever is extended program duration. As interviewee S. (S-01, 2025) noted, “more time would be the biggest factor I would change.” Longer study periods enable depth over acceleration, giving students the temporal space to iterate, recover from setbacks, and build genuine creative maturity.

Finally, increasing touchpoints with active industry artists, through guest lectures, mentoring partnerships, or collaborative modules, exposes students to diverse professional realities rather than a single idealized trajectory. These interactions recalibrate ambition toward authenticity, preparing students to

navigate the industry (and its online representations on platforms) more critically. Structural reform, then, is the groundwork upon which all pedagogical ideals depend.

6.3 Teachers as Protectors, Not Gatekeepers

Within the hidden hierarchies of art education, teachers often inherit the industry's mythology of toughness: the idea that critique must hurt to matter, or that pressure reveals true talent. Yet this stance reproduces the very harm it claims to prepare students for. One interviewee, E., states:

Later in the industry, I saw similar patterns: producers saying, 'You mean nothing, just do your job.' But I also see improvements. New generations are less toxic. Still, some people think overwork made them strong, so they repeat it. (E-01, 2025)

Reimagining the teacher as a protector means defending students' right to sustainable growth. Protection does not imply overprotection, it involves boundary-setting and transparency. The teacher models how to navigate systems without capitulating to them.

In this role, the educator shifts from judging performance to stewarding potential. S. states in the interview: "So I think it's very important that as a teacher you follow through on most of what we tell the students" (S-01, 2025). The goal is no longer to filter who is 'industry-ready', but to prepare every student to enter the industry without losing themselves to it. S. further elaborates: "For me, industry readiness means developing adaptability, being able to prepare for situations, adapt, and think in terms of problem solving. It's extremely important." (S-01, 2025)

6.4 Differentiation as Philosophy

A non-toxic curriculum must move beyond the expectation of uniform output. As McKee argues, "the duty of the lecturer is not to lift every student to a set bar – but rather to ensure that each student learns as much as they are individually capable of achieving" (2020: 77). Building on this, adaptive structures such as flexible deadlines, modular assignments, and reflective checkpoints can help distribute intensity more evenly across a term. What McKee calls "choice of

task” (ibid.: 80) might, in this context, become a broader framework for agency in the students’ academic and professional lives. As one interviewee put it:

I dream of building an environment where education doesn't end when classes end – continuous mentorship, focus on the person, not just skills. Education should round you as a person, not as a tool. (E-01, 2025)

6.5 Empathy as diagnostic literacy

Toxicity in educational contexts often emerges through the misinterpretation of student behaviour, where signs of exhaustion, withdrawal, or reduced engagement are attributed to a lack of discipline or motivation rather than structural or psychological strain (D’Arcy 2025). These misreadings reinforce a hidden curriculum in which endurance and constant productivity are valued over well-being and sustainable creative practice. Empathy offers a corrective lens for these distortions, it can be seen as a form of diagnostic literacy that helps educators take a different perspective, as framed by Warren (2018: 173). It enables teachers to ask not what’s wrong with the student, but what in the environment is producing this reaction. In their 2025 study on the role of teacher empathy in influencing student mental health outcomes, Ampofo et al. show:

Higher levels of perceived teacher empathy correlate with reduced stress, anxiety, and depression while enhancing student engagement in learning activities. Furthermore, student engagement mediates the relationship between teacher empathy and mental health outcomes.

In creative disciplines, this form of attentiveness is already familiar: artists and designers routinely shift perspectives to understand the context from and in which they have to create. Applied to teaching, it becomes a method for recognizing when something is wrong systemically. Empathy, then, does not soften academic rigor. It replaces judgment with interpretation and transforms misunderstanding into data for better design, both of artworks and of learning environments.

7. Conclusion

While the findings presented here offer insight into recurring patterns within concept art education, they are not statistically representative. The interviews serve as qualitative case studies that show tendencies rather than general laws. Further research, both quantitative and comparative, will be necessary to examine how widespread these issues are and to test which pedagogical strategies can meaningfully counteract them.

As a next step, further research will explore how these toxic learning dynamics can be countered through pedagogical experimentation. In 2026, I plan to conduct a practice-based workshop with students and educators from creative disciplines to critically reflect on their experiences and co-develop alternative teaching models. This workshop will serve as both an intervention and a data-gathering opportunity, expanding the empirical basis for understanding how toxic narratives can be challenged within the classroom itself.

Bibliography

- Ampofo, Joshua; Bentum-Micah, Geoffrey; Xusheng, Qian; Sun, Binghai; Asumang, Rita Menshah (2025): Exploring the role of teacher empathy in student mental health outcomes: a comparative SEM approach to understanding the complexities of emotional support in educational settings. In: *Frontiers in Psychology*. 16. URL: <https://doi.org/10.3389/fpsyg.2025.1503258>.
- Concept Design Academy (2012): About Concept Design Academy. In: Concept Design Academy. URL: <https://conceptdesignacad.com/about> [accessed 10.10.25].
- CG Spectrum (2025): Concept Artist Job Description, Salary, Skills & Software. In: CG Spectrum College of Digital Art & Animation. URL: <https://www.cgspectrum.com/career-pathways/concept-artist> [accessed 10.10.25].
- Chatzakou, Dimitra; Kourtellis, Nicolas; Blackburn, Jason; De Cristofaro, Emiliano; Stringhini, Gianluca; Vakali, Athena (2017): Measuring #GamerGate: A Tale of Hate, Sexism, and Bullying. In: arXiv. URL: <https://doi.org/10.48550/ARXIV.1702.07784>.
- Coffield, Emma; Markham, Katie (2024): It's time to rethink creative industries employability. In: THE Campus Learn, Share, Connect. URL: <https://www.timeshighereducation.com/campus/its-time-rethink-creative-industries-s-employability> [accessed 10.10.25].

- Cote, Amanda C.; Harris, Brandon C. (2023): The cruel optimism of “good crunch”: How game industry discourses perpetuate unsustainable labor practices. In: *New Media & Society*. 25, 3. P. 609–627. URL: <https://doi.org/10.1177/14614448211014213>.
- D'Arcy, Sean (2025): Students Are Disengaging to the Point of Skipping Exams. This Is How We Can Help. In: *University Business*. URL: <https://universitybusiness.com/students-are-disengaging-to-the-point-of-skipping-exams-this-is-how-we-can-help/> [accessed 17.03.26].
- Douwes, Rynke; Metselaar, Janneke; Pijnenborg, Gerdina Hendrika Maria; Boonstra, Nynke (2023): Well-being of students in higher education: The importance of a student perspective. In: *Cogent Education*. 10, 1. URL: <https://doi.org/10.1080/2331186X.2023.2190697>.
- Farrier, Stephen; Whittaker, Randall (2025): Creative Higher Education Isn't a Skills Pipeline, It's a Cultural Force. In: *Wonkhe*. URL: <https://wonkhe.com/blogs/creative-higher-education-isnt-a-skills-pipeline-its-a-cultural-force/> [accessed 10.10.25].
- FZD School of Design (2010): FZD School of Design – About the Program. In: *FZD School of Design*. URL: <https://fzdschool.com/the-program> [accessed 10.10.25].
- GameDeveloper (2022): The Evolution of Concept Art in Video Games. In: *GameDeveloper.com*. URL: <https://www.gamedeveloper.com/art/the-evolution-of-concept-art-in-video-games> [accessed 10.10.25].
- GameDeveloper (2023): Prince of Packaging. A Tale of 1990s Box Art. In: *GameDeveloper.com*. URL: <https://www.gamedeveloper.com/art/prince-of-packaging-a-tale-of-1990s-box-art> [accessed 10.10.25].
- Głaziewicz, Karolina; Golonka, Krystyna (2024): When the creative well dries up – burnout syndrome and art block in artists' sample. In: *Thinking Skills and Creativity*. 54. URL: <https://doi.org/10.1016/j.tsc.2024.101692>.
- Hawey, Dave (2025): Understanding Game Art Practice Beyond Technical Expertise: A Qualitative Study. In: *Proceedings of the 20th International Conference on the Foundations of Digital Games*. New York: Association for Computing Machinery (FDG '25), P. 1–12. URL: <https://doi.org/10.1145/3723498.3723724>.
- Henriksen, Danah; Mishra, Punya; Creely, Edwin; Henderson, Michael (2021): The Role of Creative Risk Taking and Productive Failure in Education and Technology Futures. In: *TechTrends*. 65, 4. P. 602–605. URL: <https://doi.org/10.1007/s11528-021-00622-8>.

- Jin, Xuguang; Ye, Yuan (2022): Impact of fine arts education on psychological wellbeing of higher education students through moderating role of creativity and self-efficacy. In: *Frontiers in Psychology*. 13. URL: <https://doi.org/10.3389/fpsyg.2022.957578>.
- Majerová, Nina; Sokolová, Lenka (2025): Well-being and academic stress in higher education students: interview and diary study. In: *International Journal of Qualitative Studies on Health and Well-being*. 20, 1. URL: <https://doi.org/10.1080/17482631.2025.2602369>.
- McKee, Connor (2020): Differentiation Techniques and Their Effectiveness for Video Game Art and Design Lectures in Higher Education. In: *The Computer Games Journal*. 9, 1. P. 75–90. URL: <https://doi.org/10.1007/s40869-020-00096-3> [accessed 10.10.25].
- Nieborg, David B.; Poell, Thomas (2018): The Platformization of Cultural Production. In: *New Media & Society*. 20, 11. P. 4275–4292. URL: <https://doi.org/10.1177/1461444818769694> [accessed 10.10.25].
- Rasool, Samma Faiz; Wang, Mansi; Tang, Minze; Saeed, Amir; Iqbal, Javed (2021): How Toxic Workplace Environment Effects the Employee Engagement: The Mediating Role of Organizational Support and Employee Well-being. In: *International Journal of Environmental Research and Public Health*. 18, 5. URL: <https://doi.org/10.3390/ijerph18052294>.
- RMCAD (Rocky Mountain College of Art + Design) (2024): Pixels to Reality. The Evolution of Video Game Art. In: RMCAD Blog. URL: <https://www.rmcad.edu/blog/the-evolution-of-video-game-art> [accessed 10.10.25].
- Sawyer, McKella (2018): Best Concept Art Schools & Colleges For Top-Notch Training. In: *Concept Art Empire*. URL: <https://conceptartempire.com/concept-art-schools/> [accessed 10.10.25].
- Schell, Jesse (2019): *The Art of Game Design: A Book of Lenses*, Third Edition. New York: A K Peters/CRC Press. URL: <https://doi.org/10.1201/b22101>.
- Sellers, Victoria; Villanueva Alarcón, Idalis (2023): From Message to Strategy: A Pathways Approach to Characterize the Hidden Curriculum in Engineering Education. In: *Studies in Engineering Education*. 4, 2. P. 176–200. URL: <https://doi.org/10.21061/see.113>.
- Smith, Mark K. (2020): Pierre Bourdieu on education: Habitus, capital, and field. Reproduction in the practice of education. In: *infed.org* URL: <https://infed.org/dir/pierre-bourdieu-habitus-capital-and-field-exploring-reproduction-in-the-practice-of-education> [accessed 17.03.26].
- Sukmana, Oman; Saleh, Amin; Kangiden, Fahmi Rahmat; Mokoagow, Moh Syahrudin (2025): SCHOOL CULTURE AND THE NORMALIZATION OF

- TOXIC PRODUCTIVITY: A REVIEW OF THE SOCIOLOGICAL LITERATURE ON PERFORMANCE PRESSURE AMONG STUDENTS. In: *Sosioedukasi. Jurnal Ilmiah Ilmu Pendidikan dan Sosial*. 14, 2. P. 907–916. URL: <https://doi.org/10.36526/sosioedukasi.v14i1.5697>.
- Taylor, T. L. (2018): *Watch Me Play: Twitch and the Rise of Game Live Streaming*. Princeton: Princeton University Press.
- Vancouver Film School (2025): *Concept Artist*. In: Vancouver Film School – Program Overview. URL: <https://vfs.edu/content/concept-artist> [accessed 10.10.25].
- Warren, Chezare A. (2018): *Empathy, Teacher Dispositions, and Preparation for Culturally Responsive Pedagogy*. In: *Journal of Teacher Education*. 69, 2. P. 169–183. URL: <https://doi.org/10.1177/0022487117712487>.
- Worsley, Joanne Deborah; Harrison, Paula; Corcoran, Rhiannin (2021): *Bridging the Gap: Exploring the Unique Transition From Home, School or College Into University*. In: *Frontiers in Public Health*. 9. URL: [doi:10.3389/fpubh.2021.634285](https://doi.org/10.3389/fpubh.2021.634285).

List of Contributors

Ashhab Ahmed studied economics at Fraser International College and psychology at Douglas College in Vancouver and worked as a gaming director in an Indie Game Studio that specialized on VR games. He gained experience both as a streamer and moderator in DotA2 communities on Twitch.

Patricia Bau ist wissenschaftliche Mitarbeiterin am Institut für Germanistik an der Universität Koblenz. Sie promoviert zu Tabuisierungs- und Enttabuisierungsstrategien in den Diskursen über die Wechseljahre und über die Midlife-Crisis. Ihre Forschungsschwerpunkte sind: Diskurslinguistik, linguistische Tabuforschung, Internetkommunikation und Gender Studies.

Tamara Boddén, Prof. Dr., is a Professor for Linguistics at the Institute of German Studies at Koblenz University. Her research areas focus on toxicity in on-line gaming communities, Game Studies, gender and AI, intertextuality and discourse analysis.

Toke Hoffmeister, Dr., is a research associate (Postdoc) at Marburg University (Research Center Deutscher Sprachatlas and Institute for German Linguistics). He is a member of the Research Training Group 2700 Dynamics and Stability of Linguistic Representations. His main areas of work include variation and change, linguistic theory, semantics and pragmatics, cognitive linguistics, and folk linguistics.

Christina Liemann, Dr., currently works as a visiting scholar at the Waseda University in Tokyo. She wrote her linguistic doctoral thesis on communicative strategies of right-wing influencers at the University of Kassel. Her research is

focused on political language, discourse analysis, digital and multimodal communication as well as gender studies.

Andreas Niegl is a research associate at University Kassel at the institute for art and economies and a member of the documenta Institute. He is currently writing his doctoral thesis on the history of the concept of the political in art and aesthetic theory. Besides this his main areas of research are: the aesthetics of the New Right, the legacy of modernism in philosophy and art, the history and philosophy of European esoteric fascism, anti- and post-liberalism.

Marion Plank is Professor of Game Art at the Game Design program at IU International University of Applied Sciences. In her years of work as a lecturer and tutor in the field of concept and game art, she accompanied many students in their careers as artists. Her research focuses on: AI and creativity, posthumanism and empathy in game development.

Mick Prinz ist Projektleiter des Innovationsprojekts „Good Gaming Support“ der Amadeu Antonio Stiftung. Er hat Sozialwissenschaften an der Universität Siegen studiert und arbeitet zum Thema Rechtsextremismus und digitale Medien. Seine Forschungsschwerpunkte sind: Digitaler Rechtsextremismus, demokratische- und antidemokratische Gaming-Communities, rechtsextreme Propagandagames. Er hält Vorträge und gibt Workshops zu dem Thema Chancen und Herausforderungen von Gaming-Welten. Dafür arbeitet er auch mit Content-Creator*innen aus der Videospiele-Branche zusammen.

Lars Sörries-Vorberger, Prof. Dr., is a Junior Professor of German Linguistics at the University of Hamburg (UHH). He currently substitutes the Chair of German Linguistics (Substitute Professor) at Kiel University (CAU). His research focuses on regional language studies, (socio-)phonetics, gender and queer linguistics.

Julia Reinhold ist Projektmitarbeiterin des Innovationsprojekts „Good Gaming Support“ der Amadeu Antonio Stiftung. Sie hat Soziologie und Politikwissenschaft an der Julius-Maximilians-Universität Würzburg studiert und ihren Master in Politikwissenschaft an der Universität Münster abgeschlossen. In ihrer Masterarbeit beschäftigte sie sich mit Anti-Genderismus als Brückenschlag zwischen konservativen und rechtsextremen Milieus. Sie arbeitet zu den Themen digitaler Rechtsextremismus, Antifeminismus und rechtsextre-

me Einflussnahme im Gaming. Im Projekt ist sie unter anderem für das Monitoring rechtsextremer und toxischer Gaming-Communities sowie für Bildungs- und Beratungsarbeit zuständig.

