

Ferret, Nathan

Article

The realm of digital content regulation as a social space: Sociogenesis of moderation norms and policies on Twitch platform

Internet Policy Review

Provided in Cooperation with:

Alexander von Humboldt Institute for Internet and Society (HIIG), Berlin

Suggested Citation: Ferret, Nathan (2025) : The realm of digital content regulation as a social space: Sociogenesis of moderation norms and policies on Twitch platform, Internet Policy Review, ISSN 2197-6775, Alexander von Humboldt Institute for Internet and Society, Berlin, Vol. 14, Iss. 1, pp. 1-25,
<https://doi.org/10.14763/2025.1.2004>

This Version is available at:

<https://hdl.handle.net/10419/315587>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/3.0/de/deed.en>



RESEARCH
ARTICLE



OPEN
ACCESS



PEER
REVIEWED

The realm of digital content regulation as a social space: Sociogenesis of moderation norms and policies on Twitch platform

Nathan Ferret *ENS de Lyon*

DOI: <https://doi.org/10.14763/2025.1.2004>

Published: 31 March 2025

Received: 5 April 2024 **Accepted:** 11 October 2024

Funding: This research was funded by LabEx "TEPSIS," the "Jeu et société" scientific interest group, and the IRIS laboratory (UMR8156).

Competing Interests: The author has declared that no competing interests exist that have influenced the text.

Licence: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License (Germany) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. <https://creativecommons.org/licenses/by/3.0/de/deed.en>
Copyright remains with the author(s).

Citation: Ferret, N. (2025). The realm of digital content regulation as a social space: Sociogenesis of moderation norms and policies on Twitch platform. *Internet Policy Review*, 14(1). <https://doi.org/10.14763/2025.1.2004>

Keywords: Content moderation, Live streaming, Digital platforms, Norms

Abstract: This article seeks to explore the socio-demographic determinants underlying the engagement of moderators and the production of content moderation norms in the French Twitch scene. Using a mixed-method approach, it highlights gender, politicisation, social class, and social vulnerability mechanisms driving these norms and their diversity across channels. Targeting a precarious and marginalised yet politicised population, Twitch is engaged by moral entrepreneurs who reinforce their social positioning through their moderation practices. The question of gender and politicisation is central, organising the positioning space of moderation practices as a realm of social distinction between two poles: on the one hand, the most prominent channels, characterised by low levels of moderation, masculinity, and depoliticisation and, on the other, the more tightly, feminine and politicised channels, serving as sources of norms embraced by platform policies. This research contributes to explaining sociologically Twitch's distinctive policy as compared to its competitors, inviting policymakers to consider how digital platforms reproduce existing social structures through their regulatory frameworks.

This paper is part of **Content moderation on digital platforms: beyond states and firms**, a special issue of *Internet Policy Review* guest-edited by Romain Badouard and Anne Bellon.

Introduction

Integrated into social media, live streaming has led to the emergence of new cultural producers on the web - the “streamers” - who use this audiovisual broadcasting technology to film themselves in a live stream while performing various leisure activities and interacting with their audience - the “viewers”. By multiplying the possibilities of real-time interactions between internet “content creators” and the public, (live) streaming has also renewed the challenges and questions surrounding content and online relationships moderation performed by internet users. Streams are indeed a very popular yet fragile form of collective spectacles and rituals, where the *interaction order* (Goffman, 1983) is particularly susceptible to malevolent disruptions. The expansion of this format across various social media platforms (from TikTok to Facebook or YouTube) thus encounters limitations tied to the necessity of safeguarding videographers and their communities from waves of hateful comments in their chat (Brewer et al., 2023).

Amazon-owned Twitch, the dominant platform in the sector¹, is regularly criticised by its users for not providing sufficient content moderation, especially concerning the numerous cases of sexist, homophobic, or transphobic harassment prevalent in this predominantly male gaming universe (Nakandala et al., 2016; Coavoux & Roques, 2020). Specialising in the broadcasting of video game streams, Twitch indeed offers a new socio-media stage for expressions of masculinities typical of video game social spaces (Lignon, 2015). This particularly applies to “warrior masculinity,” which associates men with combat, war, and conquest (Kline et al., 2003), and “geek masculinity,” which claims a gendered monopoly over the mastery of computer technologies (Kendall, 2002). In the context of commentary via chats on video games played by others, we observe the same toxic behaviours (trolling, insulting, harassing) that occur in many online games and which, both on Twitch and in games, are only minimally regulated through punitive moderation measures (Ma et al., 2023). Conversely, the platform’s moderation policy can sometimes be deemed too restrictive. Similar to Instagram (Smith et al., 2021), its rules on body visibility are criticised by some as stigmatising racial or gender minorities (Zolides,

1. 72% of the total live-streaming hours watched in 2022 (outside of China) were on Twitch - Stream Hatchet. (2022). *2022 yearly live streaming trends report*. <https://streamhatchet.com/2022-yearly-live-streaming-trends-report/>

2020), while, as on Wikipedia (de Laat, 2012), content control is perceived by others as a form of “bureaucratic” coercion that stifles freedom of expression - in this case that of “white men” (Powell & Williams-Johnson, 2023). This duality of criticism, highlighting the sometimes overly restrictive and sometimes too lax nature of Twitch’s moderation, reflects the complexity of its regulatory model.

Unlike many other self-publishing platforms primarily focused on connecting internet users through text or image in a deferred manner, Twitch can rely on its numerous highly engaged micro-communities operating under a “community-reliant” regulation logic (Caplan, 2018). Unlike platforms like X or Facebook, for example (Gillespie, 2018), Twitch indeed does not rely on commercial partners to moderate content and instead employs a hybrid model between “top-down” regulation (unilateral and automated imposition of centralised norms on the digital public space) and “bottom-up” regulation (adoption of norms produced by user groups in its policy standards). A fractal and hierarchical power structure ensures both orientations. On the one hand, Twitch acts as an intermediary between all its streamers and their audiences: at the platform-wide level, the company defines the overarching principles of content moderation supposed to be applied across all channels (the “Community Guidelines”) and moderates the video contents streamed by its broadcasters through its own employees and algorithms. On the other hand, streamers uphold this principle of intermediation and control vis-à-vis their channel viewers: at this tighter scale, chat moderation is at stake, along with the adaptation of the Community Guidelines to the localised context of each community. Here, the company provides users with various tools (such as commands for banning live viewers, either temporarily or permanently, or for prohibiting certain words in the chat), and it is up to the streamers and, most importantly, their viewers themselves to carry out the chat moderation work. The latter has thus been characterised as unpaid “digital labour” performed by viewers for the socio-economic benefit of videographers and the platform itself (Cocq, 2018; Cardon & Casilli, 2015; Li et al., 2022).

The normative articulation between the global and local levels implies a certain autonomy of channel moderation rules, which in turn generates modifications to the Community Guidelines. The definition of these emerging community norms is constructed and stabilised through practice, based on exchanges, sometimes conflictual (Cai & Wohn, 2023), between streamers and the voluntary viewers they have designated as moderators (Seerin et al., 2019), but also between these moderators and other viewers (Matias, 2019). Giving rise to forms of collective collaboration (Cai & Wohn, 2022), this moderation is an integral part of streaming: where-

as regulation work is typically hidden from users (Gillespie, 2018), on Twitch, the acts of moderation themselves are showcased in the stream, highlighting their significance not only instrumentally but also symbolically. Streamers may even conduct "ban court" streams, in which they act as judges to decide on bans imposed on viewers, either lifting or maintaining the sanction (Thach et al., 2022). This dual nature, effective and spectacular, of moderation on Twitch manifests that the practice involves "commissive performances" (Austin, 1962) contributing to the transformation of the passive audience into an active reflexive public (Dayan, 2000), engaged in community building and relationships of mutual recognition with the streamers. Thus, it is through an approach in terms of motivations to help the streamer grow its community and ensure a good atmosphere on the channel that the literature has explained the engagement of voluntary viewer moderators (Wohn, 2019). This investment is described following a triple logic, affective (related to emotional attachment to the channel), normative (related to feelings of obligation towards the community), and rational (guided by the recognition of a cost of disengagement from the activity) (Cai & Wohn, 2023; Bateman et al., 2021).

However, by utilising data that are largely detached from social context and adopting a perspective rooted in social psychology, this motivational approach faces several limitations. These are related to the lack of consideration for the socio-demographic causality of practices and community moderation norms on Twitch. While it allows describing the engagement logic of moderators, the grammar of individual motivation does not account for the socialisation processes that frame this investment. Sticking to online activity masks, in other words, the offline social processes that are reproduced in the digital sphere and ultimately explain the investment logic observed among the platform's moderators. Similarly, speaking in motivational terms certainly describes forms of self-responsibility and internalisation of moderation constraints benefiting the platform but tends to leave in the shadow the collective space within which these individual behaviours make sense². Therefore, following a socio-demographic and relational approach, this article intends to analyse the engagement logic of viewer moderators and the emergence of community norms on Twitch. It is, in other words, about approaching how the platform's regulation policy is determined by its socio-demographically marked population.

The article proceeds in three stages, corresponding to different social scales. First,

2. This collective space encompasses both the intra-platform level (among Twitch communities) and the cross-platforms one (between Twitch and other platforms offering live streaming like TikTok or Kick)

at the platform level, I will present recent developments in Twitch's policy, whose progressiveness offers it a distinctive positioning vis-à-vis its direct competitor in the context of the "platform wars". Next, at the individual level of viewers, I will show that this policy stems from the characteristics of Twitch's audience and the sociodemographic factors that frame the selection of its moderators. Finally, at the level of channels, I will discuss concrete moderation practices, showing their variation from one community to another: this will refine the connection between socio-demography and live streaming moderation norms, highlighting the diverse ways in which Twitch's general policy principles are applied across socially distinct regions of the platform.

Methodology

I rely on a mixed-method research approach, combining both quantitative and qualitative methods, with a focus on the French Twitch scene. The quantitative aspect is based on the Covideo online questionnaire survey³, which explored Twitch and video game usage during the COVID-19 period⁴ (3,455 responses in total, of which 1,814 were related to the section on moderation). This survey, initially aimed at capturing an accurate representation of the platform's audience, employed an innovative distribution method. Financial partnerships were established with streamers (at the standard payment rate for typical promotional campaigns on the platform⁵), who directly shared the questionnaire link, introduced the survey to their audience during streams, and encouraged viewers to complete it. Leveraging the close connection between streamers and their viewers, as well as the platform's technical interface, all of Twitch's visibility outlets were used for the research. An announcement introducing the survey (including objectives, survey duration, and the option not to answer certain personal questions) was accompanied by reminders from streamers over two to three weeks, messages posted on the streamers' Twitter accounts and Discord servers⁶, and bots that regularly posted the survey link in their stream's chat (at an average frequency of every 20 minutes). While distributing the questionnaire directly within the practice environment

3. Ferret N. & Gallinari Safar P. (2021). Covideo Survey. LabEx TEPSIS, "Jeu et société" scientific interest group, IRIS laboratory (UMR8156).

4. The survey was conducted from early April to the end of May 2021, partly during the last lockdown of the French population (implemented from 3 April to 3 May 2021).

5. This payment model, linked to the size of the live audience and the duration of the message broadcast, facilitated the social acceptability of the survey by embedding it within the format of "digital marketing," with which streamers are familiar.

6. Discord is a messaging platform widely used by video game players and the Twitch audience to communicate outside of streams (each channel has its "own" community Discord)

meant an overrepresentation of frequent users, the continuity and variety of completion reminders also allowed us to reach occasional viewers who are less active on the platform.

TABLE 1: Overview of channels contributing to the survey

USERNAME	GENDER	MOST STREAMED CONTENT	TWITCH FOLLOWERS	COVIDEO RESPONSES
Modiie	Woman	Just chatting	30 000	681
Marex	Man	<i>League of Legends</i>	81 000	551
Nat Ali	Woman	<i>Among us</i>	25 000	365
Zulzorander	Man	Just chatting	28 000	335
Desastre	Man	<i>THESO</i>	31 000	302
Dye	Man	<i>League of Legends</i>	50 000	288
Crawling	Man	<i>Darkest Dungeon</i>	16 000	184
Krok	Man	Just chatting	7 000	107
Bino le Dino	Woman	<i>Stardew Valley</i>	10 000	100
Melchior	Man	<i>Fortnite</i>	38 000	74
VitaPvPey	Man	<i>Fortnite</i>	102 000	45
MrChonks	Man	Just chatting	4 000	41
Chatdesbois	Woman	<i>League of Legends</i>	10 000	39
Magicknup	Man	<i>Minecraft</i>	80 000	30
Other (spontaneous sharing of the questionnaire link by viewers)				312

Source (content and followers): sullygnome.com. Information collected at the end of the survey.

The most streamed content was determined based on Twitch categories (game names or "Just Chatting" when streamers film themselves without gaming).

The exploratory and non-random approach to sampling the streamers invited to participate in the survey aimed to capture the broadest possible diversity of channels. The 14 selected streamers were distinguished by: gender (5 women and 9 men), age (ranging from 20 to 33 years), experience on the platform (ranging from 4 to 9 years), audience size (from 4,000 to 102,000 followers on Twitch), and the variety and type of content streamed—either video game-related (including types

of games and play styles) or non-video game-related (such as political news reviews, scientific outreach, and artistic practices) (Table 1). A consequence of this diversity, compounded by the fact that viewers indicated they followed a range of channels in addition to those directly involved in the survey⁷, was the uneven distribution of responses according to the streamers' audiences (Table 1). Interest in social science surveys varies with respondents' age and political views, while the use of inclusive language in the questionnaire (as was the case for the Covideo survey) introduces selection effects (Arbogast, 2017). As such, "voluntary respondent panels on the internet are not a random sample of the population" (Lensvelt-Mulders et al., 2009, p. 4): viewers of older streamers or of those who stream content on political news or scientific popularisation are overrepresented here compared to their audience on the platform. Notably, the most represented channel is that of Modiiie, known specifically for her social science outreach streams, even though she ranks only seventh in terms of followers on Twitch among the participating streamers. In contrast, despite having more than twice as many followers as Modiiie, Magicknup - a streamer focused on games with a young audience, such as Minecraft and Fortnite - appears only at the bottom of the contribution to the sample.

While the data set is thus imperfect in its representativeness, it captures a sufficiently diverse population to enable intra-group comparisons, allowing for the identification of the subpopulation of viewers engaged in moderation work ($n=185$, representing 5% of the total). Statistical regression (conducted using *R* software) will be used to identify the social determinants of this involvement and its variations within the platform's relational space. This statistical analysis is complemented by a qualitative approach, drawing from an extended "virtual ethnography" (Hine, 2000) within a variety of Twitch communities ($n=5$) and based on interviews ($n=14$) with moderators from these communities. Lastly, the article incorporates an analysis of the evolution of Twitch's normative content moderation policy texts, available online⁸.

The platform war as a socio-political struggle through moderation

In the context of intense competition between social media platforms (referred to in the media as the "platforms war"⁹), moderation norms become a central eco-

7. And particularly those of "star streamers" with several hundred thousand viewers (such as Kamet0, Squeezie, AntoineDaniel, Mister MV, or Maghla), who remained unreachable for the survey.

8. [https://safety.twitch.tv/s?language=en_US], accessed on 1 April 2024.

nomic issue. Their direction and level of strictness socially and politically shape the virtual space they regulate, thereby contributing to attract or repel specific audience segments. As in any media or expression space, such as the press (Degand & Simonson, 2011), television (Bourdieu, 1996), or forums (Wojcik, 2007), the different rules governing what can be said on live streaming platforms structure their competition and distinctive strategies. These rules are indeed at the heart of the social construction of respectability or stigma associated with expressive spaces, separating "safe" and "violent," "serious" and "frivolous," or "official" and "conspiratorial" ones. By regulating *what* can or cannot be said, and *how* one can express themselves, they also tend to select *who* gets to speak, thereby reproducing socio-demographic opposition structures in the media field. In this way, platform policies are heavily communicated by these companies. They become a marketing tool to attract and retain an audience and streamers, by playing on social and political distinctions between *populations*.

As a market leader, Twitch has historically implemented an "inclusive" moderation policy aimed at attracting a wide audience within the "Twitch community". For this reason, its policy is paradoxically based on a narrowing of expressive possibilities. In its attempt to include as many people as possible, the platform not only excludes expressions of hate or intolerance, but also activities socially associated with marginality or immorality, such as gambling, drug use, or pornography. In addition to imposing legal restrictions on the platform, allowing content that promotes these elements can damage Twitch's respectability (and, by association, that of its streamers) and its ability to maintain a "mainstream" image. Its Community Guidelines thus prohibit the use of drugs during live streams, as well as nudity. Similarly, although these streams were particularly profitable, Twitch announced in October 2022 that it will ban the broadcasting of gambling games from "gambling sites that include slots, roulettes, or dice games that aren't licensed", accusing them of "exposing the community to potential harm"¹⁰. Furthermore, like the broader media space (Plottu & Macé, 2024), Twitch has faced in recent years the rise of far-right, anti-feminist, and racist discourse emanating from influencers of the "fascosphere" (Bouron, 2017) and the "manosphere" (Ironwood, 2012). On its

9. Twitch X et TikTok 2024-2025 : La guerre des plateformes continue. (2024, September). *PC-Tablet France*. <https://pc-tablet.fr/2024/09/twitch-x-et-tiktok-2024-2025-la-guerre-des-plateformes-continue/>

Twitch vs. Kick : La guerre des plateformes de streaming est déclarée et tous les coups sont vraiment permis. (2025, January). *Jeuxvideo.com*. <https://www.jeuxvideo.com/news/1770253/twitch-vs-kick-la-guerre-des-plateformes-de-streaming-est-declaree-et-tous-les-coups-sont-vraiment-permis.htm>

10. Twitch. (2022, October 18). *Prohibiting unsafe slots, roulette, and dice gambling sites*. https://safety.twitch.tv/s/article/Prohibiting-Unsafe-Slots-Roulette-and-Dice-Gambling-Sites?language=en_US

French scene, its response was to ban almost all of these streamers for “repeated hate speech” during a massive wave of channel closures in 2022. All these restrictions, by reducing the socio-political base of its audience, directly enabled the emergence of a segment of the live streaming market not captured by Twitch. It was then claimed by the platform Kick, a new competitor with an opposing position in the realm of online moderation rules.

In the same year, 2022, Kick indeed adopted an aggressive competitive strategy by establishing financial partnerships with some of Twitch's star streamers, such as Ninja (for USD \$100m), and announced offering more attractive moderation policies for content creators. Seizing the announcement of Twitch's ban on slot machine streams, Kick positioned itself as a competitor with more flexible moderation rules, allowing gambling and being more permissive towards offensive expressions. By promoting individual agency against “blanket punitive measures” and “freedom of expression” against “knee-jerk reactions often associated with cancel culture”¹¹, this libertarian stance highlights how moderation policies are inherently political, and how market segmentation strategies also function as social segmentation strategies. Kick officially claims to provide a “mature”, “rude”, and “free” space for expression, positioning itself against Twitch which it implicitly portrays as overly protective and infantilising.

“Kick can get rowdy, reflecting the lively and spontaneous nature of our live-streaming community. If you insist on having total control over your environment at all times, you will likely have a tough time on Kick (...). Some Creators may be too loud, annoying, or even offensive to your tastes. In such instances, we encourage you to exercise your agency by navigating away from their links.”

In France, many former Twitch streamers claiming allegiance to the far right and anti-feminism, such as Kroc Blanc and Psychodelik, have found refuge on this platform, bringing to the forefront the implicitly gendered and conservative nature of Kick's moderation rules.

The dialectic of distinction between Twitch and Kick continues then, with the former reacting by emphasising its policy of defending and promoting minorities. With an update of December 2023, the company evolved its community guidelines, notably by enhancing its algorithm for automatically detecting “hateful” ban

11. Kick. (2024, August 20). *Kick community guidelines*. <https://kick.com/community-guidelines>

words, made available to viewer moderators. It also relaxed its norms related to sexual content, allowing "Content that 'deliberately highlighted breasts, buttocks or pelvic region,' even when fully clothed", "Body writing on female-presenting breasts and/or buttocks regardless of gender", "Erotic dances that involve disrobing or disrobing gestures, such as strip teases" and "popular dances, such as twerking, grinding, and pole dancing". It explicitly justified these policy changes, arguing that the previous rules "resulted in female-presenting streamers being disproportionately penalized". Its assertion of a platform prioritising "safety first" and aiming to foster "empowered communities" in the sense of defending minorities (especially in terms of gender, but also in terms of race) is also showcased during the bi-annual Twitch Con event, where drag shows and panel discussions targeting transgender or racialised communities are organised (titled, for example, "Transgender Day of Visibility," "LGBTQIA+ streamers and online inclusivity," or "Being a Black Streamer on Twitch"). It is worth noting that, some Kick streamers then harassed and physically assaulted participants during the latest edition of the event in San Diego¹².

Taken together, the progressive framing of Twitch's policy can be understood by re-contextualising it within the competitive space of platforms and the socio-political distinctions that emerge from differences in moderation rules. The live-streaming space is agonistic and dialectical, homologous to the structure of the broader social space: porous to the evolutions of the political and media fields, platforms reproduce these fault lines through their moderation rules. Perhaps even more directly than recommendation algorithms, whose effects on "filter bubbles" remain to be empirically proven (Farchy & Tallec, 2023), these rules tend to compartmentalise and polarise online user groups by filtering the populations and discourses they can access.

However, we should not conclude that Twitch has absolute agency, capable of ensuring a media-political market segmentation solely through the power of its moderation rules which it would adapt on its own to societal changes. The strength of these rules lies in the fact that they are influenced and performed daily by its users, ensuring the seemingly spontaneous alignment between the sociodemographics of the audience and the Community Guidelines that govern it.

12. TwitchCon Kick Streamers Incidents: Twitch Response. (2025, January 15). *Game Rant*. <https://gamerant.com/twitchcon-kick-streamers-incidents-twitch-response/>

The socio-demographic origins of Twitch's moderation rules

In alignment with its bottom-up moderation model, the recent policy changes on Twitch outlined earlier are directly driven by its user base, particularly through digital activism campaigns. The ban on gambling was achieved through the efforts of several prominent streamers on the platform, such as Pokimane (9.3 million followers), who threatened to boycott Twitch and called on their viewers to pressure the company (Figure 1). Similarly, the tightening of hate speech controls and the relaxation of Community Guidelines regarding sexual content were responses to the demands of several marginalised communities, following the “#twitchdobetter” movement initiated in 2021 by queer Black streamer Rekitraven, who had been harassed with homophobic and racist messages during her streams. Once again using Twitter to amplify the movement, several thousand users - mainly streamers and moderators - denounced the lack of protection against this type of harassment. A petition titled “Twitch Do Better: Stop Hate Raids Against People of Colour and Marginalised Creators” gathered 18,000 signatures, while the discussion thread also became a space for streamers from gender minorities to share their experiences, including bans for nudity violations. Even more directly, the large-scale removal of French far-right influencer channels in 2022 was the result of waves of reports filed by viewers using Twitch’s built-in reporting mechanism (Figure 2). Twitch’s progressive policy stance thus reflects the activist orientation of a significant portion of its community, which, as we will see, is itself shaped by the sociodemographic characteristics of its audience.



FIGURE 1: A tweet from Pokimane calling on her audience to pressure Twitch to ban gambling.

Report Live Stream ×

You are reporting deste059 for the current live stream

Live Stream

Why are you submitting this report?

☐ Terrorism

☐ Nudity or Sexually Explicit

☐ Hateful Conduct

☐ Bullying or Harassment

☐ Violence or Gore

☐ Self-Harm

☐ Spam, Scams, or Bots

☐ Miscategorized Content

☐ Missing or Incorrect Content Classification Label

☐ Search 🔍

FIGURE 2: Channel report form available through Twitch's interface

While it is not to suggest that all Twitch users share the same social attributes or activist tendencies regarding moderation, the general sociodemographics of the platform's audience stand out for characteristics that lean towards the left of the political spectrum. Objectively distinct from populations gathered on other platforms, Twitch's general audience transforms this objective social distinction into a normative one through the moderation rules they tend to advocate for: as in the broader social world, one's position within the space of relational positions shapes political and cultural stances (Bourdieu, 1987). Compared to platforms like TikTok or Kick, Twitch viewers are distinguished not only by their gender characteristics but also by their levels of education and economic capital (Ferret, 2023). The platform over-represents non-binary individuals compared to the French population - they constitute 3% of the Covideo survey sample compared to 1% in general popu-

lation surveys (Rault & Trachman, 2023) - as well as highly educated young people with relatively limited economic resources. On the one hand, among viewers aged 25 to 34 in the sample (excluding students who have not yet completed their studies), 64% have attained a degree higher than a bachelor's level, compared to 34% of the French population in the same age group¹³. On the other hand, 36% of viewers aged 25 to 29 fall below the poverty line¹⁴, compared to 16% of the general French population in the same age group. This particular combination of a pronounced high level of educational cultural capital and a low level of economic capital shapes the political orientation of this population and the general norms it fosters. In a classical sense (Bourdieu, 1987), such a structure of capitals is indeed typically associated with a left-leaning political stance, as opposed to structures where economic capital outweighs institutionalised cultural capital. This alignment explains the high proportion of Twitch viewers identifying as "left-wing" (23.5%) or "very left-wing" (36.5%) politically.

But behind this overarching view lie disparities that must be addressed to more precisely understand the origins of Twitch's progressive policy shift. As previously mentioned, it is primarily the channel moderators who carry out the work of moderation and act as "moral entrepreneurs" (Becker, 1966), enforcing or imposing new norms¹⁵. In other words, two levels of sociodemographic characteristics are at play in the process of norm emergence: the first at the platform audience level and the second at the more granular level of individual investment in moderation. By shedding light on the factors driving the social selection of moderators, we gain a clearer understanding of the forces shaping Twitch's moderation rules, while emphasising the inherently social nature of these norms and their origins.

TABLE 2: Logistic regression modeling the likelihood of being a moderator within at least one Twitch channel

	NORMALISED COEFFICIENT	STANDARD ERROR	CRITICAL PROBABILITY
Socio-demographic variables			
Age	0,075	0,098	<i>Non-significant</i>

13. INSEE. (2019). *Employment survey*. Institut National de la Statistique et des Études Économiques. <https://www.insee.fr/fr/statistiques/4809583>

14. Using the calculative definition of INSEE, the poverty line is set at 60% of the median standard of living in France (i.e., €1,218 in 2021).

15. Given that, without their social base of viewers (of whom moderators are the most engaged, as we will see), streamers would obviously not be able to effectively conduct their pressure campaigns against Twitch.

	NORMALISED COEFFICIENT	STANDARD ERROR	CRITICAL PROBABILITY
Level of education (quantitative coding)	-0,041	0,066	<i>Non-significant</i>
Monthly income	-0,173	0,101	9%
Gender male (vs female)	-0,054	0,423	<i>Non-significant</i>
Non-binary gender (vs binary)	0,087	0,053	10%
Declares a political orientation (vs no)	0,121	0,060	4%
Does not declare oneself socially isolated (vs yes)	0,091	0,055	10%
Mental health variables			
Declares worry about oneself future (vs no)	0,170	0,077	3%
Discusses "personal mental health issues" on Twitch (vs no)	0,162	0,070	2%
Sociability on Twitch			
Does not participate in chats (vs yes)	-0,129	0,086	<i>Non-significant</i>
Discusses "video games" on Twitch (vs no)	-0,080	0,079	<i>Non-significant</i>
Discusses "political subjects" on Twitch (vs non)	0,161	0,082	5%
Has not met anyone to regularly discuss with on Twitch (vs yes)	-0,238	0,072	>1%
Has not made at least one "friend" on Twitch (vs yes)	-0,196	0,054	>1%
Twitch usage			
Weekly time spent on Twitch	0,072	0,089	<i>Non-significant</i>
Tenure on the platform (quantitative coding)	-0,009	0,075	<i>Non-significant</i>
Is not present from the start of streams (vs yes)	0,017	0,076	<i>Non-significant</i>
Tends to choose streams watched by few people live	0,087	0,099	<i>Non-significant</i>
Amount of monetary donations to Twitch streamers	0,040	0,041	<i>Non-significant</i>
Does not use Twitch upon waking up or to fall asleep (vs yes)	-0,114	0,070	10%

Source: Ferret & Gallinari (2021). Covideo Survey. LabEx TEPSIS, "Jeu et société" scien-

tific interest group, IRIS laboratory (UMR8156).

Sample: 1,814 viewers from the Covideo database who responded to all corresponding questions. $R^2(\text{McFadden}) = 0.254$

Interpretation: with parameters of the model fixed, not participating in Twitch chats significantly decreases (at the 10% threshold) the probability of being a moderator.

The non-significance of many variables in the model (10 out of 20) is itself a notable result, as these primarily pertain to the surveyed participants' usage patterns, characterised simply as users. When compared with variables situating individuals in offline structures of gender, class, and politics, they seem to have a less significant effect than that emphasised by the literature on the subject, which particularly stresses the socially localised dimension of moderation within channel groupings.

In terms of streaming usage patterns, moderators indeed do not seem to significantly differ from other highly engaged viewers. All else being equal, the weekly time spent on Twitch and the tenure on the platform do not have a significant effect on the propensity to become a moderator (Table 2). The same applies to streaming habits and relationships with streamers. The act of donating money to streamers and being present at the start of their streams, in a logic of daily and scheduled appointments with the "content creator" and their other viewers, also do not have a significant effect on the likelihood of investing in moderation on Twitch (Table 2). While the platform is characterised by its multitude of micro-communities purportedly fostering engagement in moderation (Wohn, 2019), preferring to watch channels with fewer viewers live (rather than many viewers or having no preference) actually has no significant effect in itself either (Table 2).

Alongside the intensity of involvement in social interactions on channels, the content of these community exchanges is crucial in entering the moderation career. All else being equal, meeting people, or even "friends," to regularly discuss on the platform has a positive and significant effect (Table 2). However, when we look at the content of these discussions, whereas conversations related to video games have no significant effect in the model, it is different for those about "political subjects" and especially "personal mental health issues" (Table 2). Contrary to a vision that would characterise Twitch primarily as a platform for video game entertainment forming "communities of practice" (Wenger, 1998) and more precisely "gamer communities" (Pearce & Artemesia, 2011), moderators recruited by streamers to support communities are brought together by shared offline life experiences and attributes.

These results have significant implications for understanding moderation activity and the socialisation mechanisms that underpin it. The fact that viewers' sociodemographic variables have an effect suggests that the interpretation should focus on the nature of the social forces shaping streaming as a tool for regulating online discourse. While, all else being equal, age, level of education, and binary gender (male or female) do not have a significant effect on the likelihood of becoming a moderator, the situation is different for viewers who declare a non-binary gender identity, a political orientation (as opposed to no declared political orientation), social isolation, or concerns about their future (Table 2). Similarly, unemployed or inactive individuals represent the most overrepresented activity group among all moderator viewers (Chart 1).

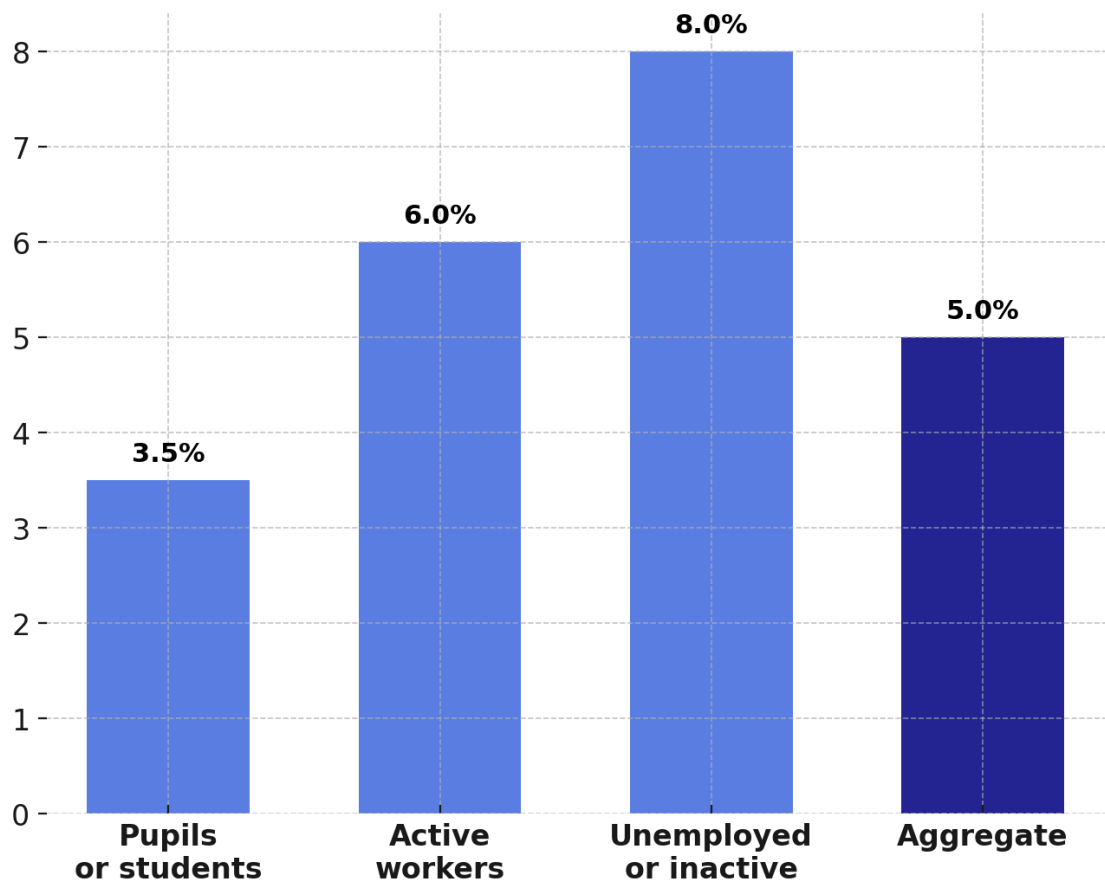


CHART 1: Rates of moderators among viewer groups defined by activity status

Source: Ferret & Gallinari (2021). Covideo Survey. LabEx TEPSIS, "Jeu et société" scientific interest group, IRIS laboratory (UMR8156).

Sample: 1,380 Twitch viewers who provided information on the corresponding variables.

Thus, when adopting a more detailed scale of analysis beyond income and education levels, we uncover the complex interplay of domination and vulnerability, as well as politicisation and mutual aid, that shapes moderation activity on Twitch. By bringing together gender minorities, socially isolated individuals, and psychologically vulnerable youth, the platform's moderation system filters a population inclined to transform it into a space for sharing experiences, fostering politicisation, and collectively resisting forms of domination. This population then tends to seize the tools of expression and silencing offered by the internet to shape conversations on the platform. Through collective actions, it forms the population at the source of the platform's policy orientation towards progressive inclusivity.

However, this particularly engaged population does not encompass all Twitch moderators. While it leaves its general mark on the Community Guidelines, it conceals a plurality of channels and approaches to applying these rules within the specific contexts of individual communities. Concluding this exploration of the sociogenesis of moderation, it becomes essential to consider the platform itself as a space of struggles and distinctions. Similar to the dynamics observed at the inter-platform level, these tensions manifest in the control over what can or cannot be said live.

The spatialisation of moderation norms across different regions of the platform

The practical application of Twitch's policy by moderators varies according to a structure of opposition that is less stark but analogous to the divide between Twitch and Kick. In an ideal-typical sense, we observe, on the one hand, a set of channels that apply the Community Guidelines in a loose and minimal way (or, at times, not at all), and on the other, channels characterised by strict and politically engaged moderation¹⁶. The first group can be described as "mainstream" channels - predominantly male-oriented, with large audiences, focused on gaming content, and formally depoliticised. These channels, often porous to masculinist expressions in gaming culture and with chats that are hard to manage due to message volume, are those most affected by a lack of moderation norms (Cai et al., 2021). The second group, by contrast, comprises channels that can be termed "safe spaces" - tending towards more feminised audiences, smaller and more cohesive viewerships, extra-gaming content, and explicit political engagement. This dichotomy is directly implicated in harassment raids in chats and the reporting of

16. For a detailed presentation of the sociodemographic characteristics of channels of these two regions, see (Ferret & Gallinari-Safar, 2024).

channels that drive Twitch's policy evolution: due to their characteristics, the "mainstream" region tends to generate waves of harassment, while the "safe space" region is more likely to initiate reports. This opposition is then widely recognised by the interviewed viewer-moderators, who perceive it differently depending on their regional affiliation. A mainstream moderator like Ewen downplays the violence in chat exchanges, seeing them as mere "jokes," and labels safe space channels as "woke." In contrast, Léo, a safe space moderator, uses similarly political categories of perception, but focuses on uncovering the violence behind the apparent frivolity of online exchanges.

"(...) We don't overthink it, really. Unless someone says something totally out of line, you know, like insults or something, I'm not gonna start making drama for no reason or stop people from saying what they want to say. (...) Especially because, honestly, it's mostly just jokes. That's what those woke communities don't get, really—it's mostly just humor"¹⁷. (Ewen, 23 years old, student, moderator on a mainstream channel with 81,000 followers)

"When you're browsing around like that to check out other channels, what's the first thing you look at to form an opinion?

Well, naturally, when I'm discovering something new, I try, I try to see how the streamer interacts with their community. Spontaneously, I ask myself, 'What's their community like? What's the vibe they give off?' I think you can tell pretty quickly when you land in a Twitch chat. Just by looking at the amount of trolling in the chat, you can already guess - it's probably mostly young people or right-wing. So, yeah, that gives you a bit of an impression. The way people interact, the kind of reactions they have, like, 'Oh look, she's doing hot tub streams again, blah blah blah.' Yeah, okay, we get it, you hate women. Moving on to another channel." (Léo, 22 years old, moderator on a safe space channel with 900 followers)

The two modes of moderation corresponding to the opposing poles of the previously described structure thus become more defined. The activity of mainstream moderators tends to be not only minimal, but also primarily entertainment-focused. In their testimonies, their chat moderation efforts are particularly geared towards ensuring the smooth running of the video game entertainment spectacle, helping the streamers deliver their content. They pay close attention to preventing messages that might spoil the progression of the game's storyline being played

17. The interview excerpts are translated by the author.

live, as well as promotional messages from smaller competing streamers. Temporary bans can even be used humorously, as seen when, at the request of streamer Kamet0, a moderator banned a viewer arguing with him about the best playable characters in *League of Legends*. These moderators often justify their involvement by highlighting the pride of being recognised and endorsed by well-known streamers, as well as the privileged social status they enjoy within the community—symbolised, for instance, by specific moderation badges attached to their username in the chat. In contrast, *safe space* moderators primarily describe their motivation as activist and serious, their investment in moderation being directed towards combating the discrimination experienced by minorities online:

“Anything related to feminism, respect for human rights, the LGBTQ+ community - yes, if someone crosses the line, is openly transphobic, homophobic, sexist, or anything like that, my role as a moderator means I’ll tell that person, ‘Your words aren’t welcome in this community. Either you change, or you’re out’.”(Luna, 25 years old, engineer, moderator on a safe space channel with 900 followers)

The case of moderating “backseat gaming” (i.e. the act of criticising every in-game decision made by the streamer) illustrates well how the same moderation practice can carry different social significance depending on the region of Twitch. Regardless of the pole to which the channel belongs, backseat gaming is socially prohibited in both cases. However, where a mainstream moderator like Tarek justifies this prohibition as necessary to preserve the entertainment value of the stream from “immature” behaviour, Marianne, a safe space moderator, interprets this practice as a form of misogyny, originating not from “kids” but from “masculinists”:

“What’s backseat?
It happens all the time. It’s when people come in to show off, telling the streamer, ‘Do this, do that in the game.’ It’s not the thing to do, especially when it makes the streamer stop every two seconds to read their tips in chat, even though they’re actually really good at the game. It’s kind of childish”. (Tarek, 17 years old, high school student, moderator on a mainstream channel with 120,000 followers)

“[Backseat] definitely has something ‘mascu’ behind it, that’s for sure. It’s these little guys coming in to tell you how you should play, basically trying to control

what you're doing. Women in gaming are still pretty new, and there's still this stereotype of the gamer girl being bad at video games (...)" (Marianne, 25 years old, local government employee, moderator on a safe space channel with 14,000 followers)

Of course, this opposition should not be overstated. The politicised or non-politicised nature of a channel unfolds along a spectrum, in both directions. Mainstream channels, while presenting themselves as outside the realm of politics (as exemplified by *Joueur du Grenier*, whose "chat rules" on his Twitch page explicitly ban "political discourse"), can nonetheless perpetuate highly political masculinist tropes. Similarly, not all moderators of safe space channels invest in their activity with the same level of political intensity and automaticity. For instance, Léo notes that he only moderates political discussions on the community Discord at the explicit request of the streamer:

"If it goes beyond my convictions, I'll talk to the person privately rather than publicly, because the Discord isn't supposed to be a political Discord. It's originally a Discord for art and gaming. We don't make political topics a focus unless the streamer specifically tells us, 'Hey, this issue matters to me a lot, so be careful if it comes up on the Discord—it bothers me.' In that case, yeah, as a moderator, I'll step in. Otherwise, no." (Léo)

Another example of intra-region variation, the political dimension of moderation is far more explicit and pronounced in a small set of channels highly engaged in Marxist ideology (such as *Bolchegeek* or *Usul*), which are also less feminised. Thomas, a 27-year-old moderator on one of these channels, adopts a more agonistic approach to moderation. He does not hesitate to intervene directly to censor harassment raid messages, which he describes as "fascist," and to "educate" viewers in response to messages he perceives as reflecting a lack of politicisation. Finally, it is worth noting that video games do not inherently equate to depoliticisation, safe space channels, such as *Bolchegeek* (Figure 3), often intertwining gaming culture with activism (Ferret & Gallinari-Safar, 2024).



FIGURE 3: Bolchegeek channel stream background (Twitch, 2024)

Conclusion

The content moderation policies on digital platforms like Twitch are deeply intertwined with the socio-demographic determinants shaping processes of marginalisation, politicisation, and digital social distinction. This study demonstrates that the logic of moderation cannot be disentangled from the broader gender and political structures of offline social life, which manifest and reproduce digitally across three levels: the platform-wide guidelines shaped by competitive and socio-political distinctions, the individual engagement of moderators and the variation in moderation norms between channels.

At the platform level, Twitch's position as a leader in the "platform wars" exemplifies how moderation policies serve as a market segmentation tool. By responding to user-driven activism and aligning itself with progressive values, Twitch has positioned itself against competitors like Kick, which embrace a libertarian and masculinist ethos. This dialectic of inclusivity versus permissiveness highlights the role of platforms in reproducing broader socio-political divisions. At the individual level, the socio-demographic characteristics of moderators - such as their gender identity, political orientation, mental health and socio-economic status - play a crucial role in shaping their investment into moderation practices. This highlights the way Twitch policy emerges from shared offline experiences and sociodemographic dynamics rather than merely from platform-specific dynamics. Then, between channels, the contrast between "mainstream" and "safe space" regions demonstrates how socio-political distinctions influence the application of modera-

tion norms within the platform itself. Mainstream channels (typically more masculine) tend towards depoliticised, entertainment-focused moderation, while safe spaces ones (typically more feminine) align more closely with activist and protective approaches. These variations underscore the pluralism within Twitch's normative and sociodemographic ecosystem, challenging the notion of uniform platform regulation and audience.

By addressing social isolation, economic precarity, and psychological vulnerability of youth, platforms like Twitch play an outsized role in shaping digital social integration of new generations. Policymakers must therefore engage with the implications of community-based moderation models, recognising their dual potential to empower and exclude. For researchers, this study offers a framework to analyze moderation as a multi-scalar phenomenon - bridging platform policies, sociodemographic characteristics, and community dynamics. Future research should explore how these dynamics evolve across different cultural contexts and platforms, particularly as competitive pressures and socio-political landscapes continue to shift.

References

- Arbogast, M. (2017). *La rédaction non-sexiste et inclusive dans la recherche: Enjeux et modalités pratiques [Non-sexist and inclusive writing in research: Issues and practical modalities]*. 231.
- Austin, J. L. (1962). *How to do things with words*. Clarendon Press.
- Bateman, P. J., Gray, P. H., & Butler, B. S. (2011). Research Note – The Impact of Community Commitment on Participation in Online Communities. *Information Systems Research*, 22(4), 841–854. <https://doi.org/10.1287/isre.1090.0265>
- Becker, H. S. (1966). *Outsiders. Studies in the sociology of deviance*. Free Press.
- Bourdieu, P. (1987). *Distinction. A social critique of the judgement of taste*. Harvard University Press.
- Bourdieu, P. (1996). *Sur la télévision [On television]*. Raisons d'Agir.
- Bouron, S. (2017). Des « fachos » dans les rues aux « héros » sur le web: La formation des militants identitaires [From “fascists” in the streets to “heroes” on the web: The training of identity activists]. *Réseaux*, 202–203(2), 187–211. <https://doi.org/10.3917/res.202.0187>
- Brewer, J., Ruberg, B., Cullen, A. L. L., & Persaud, C. J. (2023). *Real life in real time. Live streaming culture*. The MIT Press.
- Cai, J., Guanlao, C., & Wohn, D. Y. (2021). Understanding rules in live streaming micro communities on Twitch. *ACM International Conference on Interactive Media Experiences*, 290–295. <https://doi.org/10.1145/3452918.3465491>

Cai, J., & Wohn, D. Y. (2022). Coordination and collaboration: How do volunteer moderators work as a team in live streaming communities? *CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3491102.3517628>

Cai, J., & Wohn, D. Y. (2023). Understanding Moderators' Conflict and Conflict Management Strategies with Streamers in Live Streaming Communities. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–12. <https://doi.org/10.1145/3544548.3580982>

Caplan, R. (2018). *Content or context moderation? Artisanal, community-reliant, and industrial approaches*. Data & Society Research Institute. <https://datasociety.net/library/content-or-context-moderation/>

Cardon, D., & Casilli, A. (2015). *Qu'est-ce que le digital labor? [What is digital labor?]*. INA éditions. <https://larevuedesmedias.ina.fr/quest-ce-que-le-digital-labor>

Coavoux, S., & Roques, N. (2020). Une profession de l'authenticité: Le régime de proximité des intermédiaires du jeu vidéo sur Twitch et YouTube [A profession of authenticity: The proximity regime of video game intermediaries on Twitch and YouTube]. *Réseaux*, N° 224(6), 169–196. <http://doi.org/10.3917/res.224.0169>

Cocq, M. (2018). Constitution et exploitation du capital communautaire: Le travail des streamers sur la plateforme Twitch [Building and exploiting community capital: The work of streamers on the Twitch platform]. *La Nouvelle Revue Du Travail*, 13. <https://doi.org/10.4000/nrt.3911>

Dayan, D. (2000). Télévision: Le presque-public [Television: The almost-public]. *Réseaux*, 100, 427–456. <https://shs.cairn.info/revue-reseaux1-2000-2-page-427?lang=fr>

De Laat, P. B. (2012). Coercion or empowerment? Moderation of content in Wikipedia as 'essentially contested' bureaucratic rules. *Ethics and Information Technology*, 14(2), 123–135. <https://doi.org/10.1007/s10676-012-9289-7>

Degand, A., & Simonson, M. (2011). La modération des fils de discussion dans la presse en ligne [Moderation of discussion threads in the online press]. *Les Cahiers du Journalisme*, 22(23), 56–73.

Farchy, J., & Tallec, S. (2023). De l'information aux industries culturelles, l'hypothèse chahutée de la bulle de filtre [From information to cultural industries, the troubled filter bubble hypothesis]. *Questions de Communication*, 43, 241–268. <https://doi.org/10.4000/questionsdecommunication.31474>

Ferret, N. (2023). *Le capitalisme autonarratif. Production, réception et valorisation du récit de soi sur Twitch [Self-narrative capitalism: Production, reception, and valorisation of self-narratives on Twitch]* [PhD Thesis]. EHESS.

Ferret, N., & Gallinari Safar, P. (2024). « Ce soir, c'est le stream de la dépression »: Sociologie de l'appropriation du live streaming comme ressource de santé mentale chez les jeunes [“Tonight is the depression stream”: Sociology of the appropriation of live streaming as a mental health resource among young people]. *Agora Débats/Jeunesses*, N° 97(2), 56–72. <https://doi.org/10.3917/agora.097.0056>

Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.

Goffman, E. (1983). The interaction order: American Sociological Association, 1982 presidential address. *American Sociological Review*, 48(1), 1. <https://doi.org/10.2307/2095141>

Hine, C. (2000). *Virtual ethnography*. SAGE Publications.

- Ironwood, I. (2012). *The manosphere: A new hope for masculinity*. Red Pill Press.
- Kendall, L. (2002). *Hanging out in the virtual pub: Masculinities and relationships online*. University of California Press. <https://doi.org/10.1525/california/9780520230361.001.0001>
- Kline, S., Dyer-Witheford, N., & Peuter, G. (2003). *Digital play. The interaction of technology, culture, and marketing*. McGill-Queen's University Press.
- Lensvelt-Mulders, G. J. L. M., Lugtig, P. J., & Hubregtse, M. (2009). Separating selection bias and non-coverage in internet panels using propensity matching. *Survey Practice*, 2(6), 1–6. <https://doi.org/10.29115/SP-2009-0026>
- Li, H., Hecht, B., & Chancellor, S. (2022). Measuring the monetary value of online volunteer work. *Proceedings of the International AAAI Conference on Web and Social Media*, 16, 596–606. <https://doi.org/10.1609/icwsm.v16i1.19318>
- Lignon, F. (Ed.). (2015). *Genre et jeux vidéo [Gender and video games]*. Presses Universitaires du Midi.
- Ma, R., Li, Y., & Kou, Y. (2023). Transparency, fairness, and coping: How players experience moderation in multiplayer online games. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–21. <https://doi.org/10.1145/3544548.3581097>
- Matias, J. N. (2019). The civic labor of volunteer moderators online. *Social Media + Society*, 5(2), 2056305119836778. <https://doi.org/10.1177/2056305119836778>
- Nakandala, S., Ciampaglia, G. L., Su, N. M., & Ahn, Y.-Y. (2016). *Gendered conversation in a social game-streaming platform* (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.1611.06459>
- Pearce, C. & Artemesia. (2011). *Communities of play: Emergent cultures in multiplayer games and virtual world*. The MIT Press.
- Plottu, P., & Macé, M. (2024). *Pop fascisme. Comment l'extrême droite a gagné la bataille culturelle sur Internet [Pop fascism: How the far right won the cultural battle on the internet]*. Éditions Divergences.
- Powell, A., & Williams-Johnson, D. (2025). “You dumb cracker b*tch”: The legitimizing of white supremacy during a Twitch ban of HasanAbi. *New Media & Society*, 27(3), 1318–1335. <https://doi.org/10.1177/14614448231191776>
- Rault, W., & Trachman, M. (Eds.). (2023). *Minorités de genre et de sexualité. Objectivation, catégorisations et pratiques d'enquête [Gender and sexual minorities: Objectification, categorisations, and research practices]*. INED Éditions.
- Seering, J., Wang, T., Yoon, J., & Kaufman, G. (2019). Moderator engagement and community development in the age of algorithms. *New Media & Society*, 21(7), 1417–1443. <https://doi.org/10.1177/1461444818821316>
- Smith, S. L., Haimson, O. L., Fitzsimmons, C., & Brown, N. E. (2021). *Censorship of marginalized communities on Instagram* (Algorithmic Bias Report). Salty. <https://saltyworld.net/exclusive-report-censorship-of-marginalized-communities-on-instagram-2021-pdf-download/>
- Thach, H., Mayworm, S., Delmonaco, D., & Haimson, O. (2024). (In)visible moderation: A digital ethnography of marginalized users and content moderation on Twitch and Reddit. *New Media & Society*, 26(7), 4034–4055. <https://doi.org/10.1177/14614448221109804>
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity* (1st ed.). Cambridge University Press. <https://www.cambridge.org/core/product/identifier/9780511803932/type/book>

Wohn, D. Y. (2019). Volunteer moderators in Twitch micro communities: How they get involved, the roles they play, and the emotional labor they experience. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3290605.3300390>

Published by



ALEXANDER VON HUMBOLDT
INSTITUTE FOR INTERNET
AND SOCIETY

in cooperation with



CREATE



centre
— internet
et — société



R&I IN3
Internet
interdisciplinary
Institute
Universitat Oberta de Catalunya



UNIVERSITY OF TARTU
Johan Skytte Institute of
Political Studies