

Representation, Resistance, and Moderation: The Politics of Review Bombing in Assassin's Creed Shadows

Alan Tapscott, Jordi Justicia

Department of Communication, Universitat Pompeu Fabra
alan.tapscott@upf.edu, jordi.justicia@upf.edu

Ona Anglada

Department of Library and Information Science and Audiovisual Communication,
Universitat de Barcelona
ona.anglada@ub.edu

Mercè Oliva

Department of Communication, Universitat Pompeu Fabra
merce.oliva@upf.edu

Keywords

Review bombing, historical games, game reviews

This abstract details preliminary research examining how online user reviews transform into spaces for cultural and political discussion. Focusing on the release of *Assassin's Creed Shadows* (Ubisoft 2025), this study analyzes the review bombing campaign that erupted after the game's release. The inclusion of an historical Afro-American samurai and a female ninja, and the option for queer relationships, are instances of typically underrepresented collectives and their social dynamics for a protagonist role. This inclusion provoked an immediate and polarized response, highlighting "authenticity," "wokeness," and the perceived "politicization" of the franchise.

Review bombing, defined as the coordinated posting of negative user reviews often driven by cultural or political opposition, has become a recognized mode of collective resistance in game culture. This "ideology-driven polarization" (Cantone et al. 2021) takes place on sites like Metacritic, where user ratings are frequently more polarized and emotional than professional critic reviews, reflecting differing standards of authority and evaluation (Coronado-Blázquez 2024; Santos et al. 2019). Furthermore, the deletion of reviews represents a mechanism through which toxic technocultures restrict cultural ownership and belonging for certain audiences (Spallaccia 2023). In the absence of prescriptive legal constraints, platforms develop their own bespoke policies, effectively regulating the emerging discourse. Platforms must also contend with threats to integrity, ranging from ideological backlash to review spam and rating manipulation. This research situates the *Shadows* controversy within broader contemporary debates about gaming culture, representational resistance, and the politics of moderation.

Proceedings of DiGRA 2026

© 2026 Authors & Digital Games Research Association DiGRA. Personal and educational classroom use of this paper is allowed, commercial use requires specific permission from the author.

The study employed a mixed-method approach combining quantitative score aggregation, topic modeling, and qualitative thematic analysis (Braun & Clarke 2022). We conducted 4.077 iterations of automated scraping of Metacritic’s *Assassin’s Creed Shadows* (PlayStation 5) user review page from March 21 to April 23, 2025 (the first month after launch). The final dataset includes 3.502 user scores, 1.835 of which included a textual review, and 635 of which were deleted by Metacritic’s moderation system. Additionally, the dataset includes 92 professional site Metacritic scores for the game.

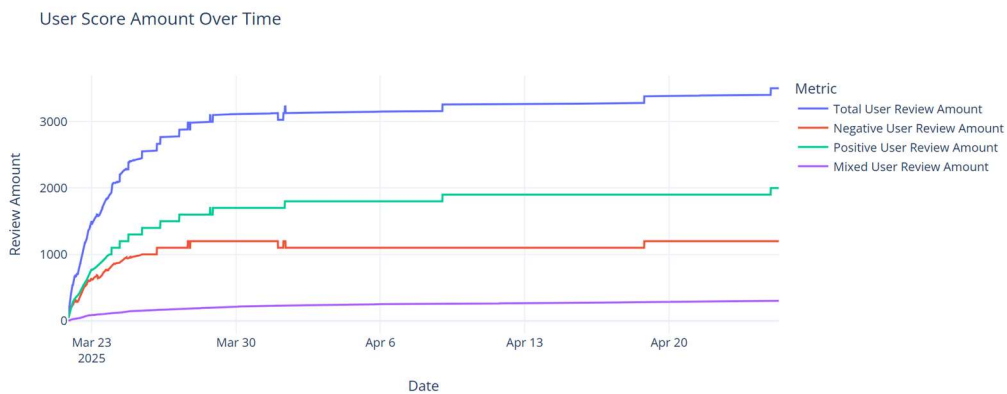


Figure 1: Metacritic user score amount over time.

Most of the parsed reviews were written over the first day. *Shadows’* estimated length is between 34 and 64 hours¹ (Metacritic imposes a 36-hour delay). Professional critic scores had a mean of 84%, whereas user scores had a mean of 51%. However, when computed separately, user reviews showed a distinctly bimodal distribution, with deleted reviews having an average of 27% (and a median of 10%), and accepted reviews having a mean of 63% of (and a median of 80%). Due to the anonymous nature of the reviews, it is not possible to infer the demographic data of the authors.



Figure 2: User review scores.

The polarized bimodal distribution of user scores suggests distinct regimes of cultural authority, mirroring findings on discrepancies between expert and amateur reviews (2019 Coronado-Blázquez 2024; Santos et al.). The concentration of user scores in the first hours (for a game that is meant to last dozens of hours) suggests that the rigor of the review is secondary to the need to project either early praise or criticism to reporting media. These results postulate the possibility that ratings were positioned as ideological statements rather than nuanced assessments of quality

To analyze the ideological opposition, the BERTopic² topic modelling technique was applied, resulting in a total of 42 topic clusters. We then used Gemma-2-2b-it LLM³ to write natural language titles for each cluster using the contained words and the designated most representative reviews. The following LLM BERTopic cluster titles from the deleted review sample are ranked first, second, and thirteenth in size, respectively:

- Black Yasuke Controversy (including 107 reviews): Criticism on the Afro-American protagonist and appeals to "historical accuracy".
- Turbocharged Inclusivity/Woke Agenda (including 99 reviews): Rejecting non-heteronormative representations and denouncing a hidden political agenda.
- Manipulated Assassin's Creed Reviews (including 19 reviews): Claims that Metacritic censors dissent and inflates positive scores to favor Afro-American and female protagonists, as well as queer representation, reflecting user distrust of platform governance.

A complementary qualitative thematic analysis of the most representative deleted reviews of the clusters created through BERTopic is being conducted. Out of the eight code groups identified by one researcher, the three most prevalent are *Illegitimate black samurai*, *Metacritic dynamics discussion*, and *inclusion of non-traditional relationships or gender identities*. Discourses of "historical accuracy" seem to be policing representation. Backlash against the Black samurai or queer romance suggests that "realism" serves as a cultural boundary designed to maintain whiteness, masculinity, and heteronormativity as the unmarked norms of historical imagination within the medium. Users instrumentalize "cultural respect" toward Japanese heritage to legitimize exclusion, using authenticity as a rhetorical weapon. This rejection of inclusion aligns with research connecting backlash to hegemonic masculinity and toxic technocultures (Spallaccia 2023) and the rejection of established conventions of whiteness, masculinity, and heteronormativity (Ruberg 2020; Shaw 2015). Through the lens of the 3A framework for historical narrative analysis in games (Donald 2023), the game's account, the narrative centering on Yasuke, is grounded in historical accuracy, yet it is perceived as lacking authenticity and a rhetorical weapon for a political agenda.

The deletion of hundreds of discriminatory reviews underscores the political function of platform moderation. This process, which can be seen as "algorithmic forgetting" (Chun 2016), obscures the true extent of toxicity while leaving the underlying ideological sentiment intact. Moderation, therefore, manages the visibility of hate rather than addressing its root causes. The opacity of Metacritic's rules makes it difficult for users to know how moderation is occurring, breeding distrust and the conspiratorial claims seen in the deleted reviews.

Ultimately, the *Shadows* controversy reveals a polarized reception for non-heteronormative white male representation, especially in a protagonist role. Online rating systems in the context of review bombing are forms of discursive participation, where scoring translates social tensions and resistance into quantifiable metrics. Understanding review bombing not as noise but as discourse allows observing how audiences articulate and resist competing visions of what games should represent and how. It also represents a chance to elucidate the discourse that is removed from the desired sanitized criticism, excluding hateful and discriminatory comments against the representation of marginalized collectives and questioning the moderation policies.

Braun, V., & Clarke, V. 2021. *Thematic analysis: A practical guide*. SAGE.

Cantone, G. G., Tomaselli, V., & Mazzeo, V. 2021. *Ideology-driven polarization in online ratings: The review bombing of *The Last of Us Part II**. arXiv preprint arXiv:2104.01140.

Chun, W. H. K. 2016. *Updating to Remain the Same: Habitual New Media*. MIT Press.

Coronado-Blázquez, J. (2024). *A NLP Approach to Review Bombing in Metacritic PC Videogames User Ratings*. arXiv preprint arXiv:2405.06306.

Donald, I. and Reid, A. 2023. *Account, accuracy, and authenticity: A framework for analyzing historical narrative in games*. In (Not) in the Game: History, Paratexts, and Games, edited by R. Seiwald and E. Vollans, 57–80. Berlin/Boston, De Gruyter.

Ruberg, B. 2020. *The Queer Games Avant-Garde*. Duke University Press.

Santos, T., Lemmerich, F., Strohmaier, M. & Helic, D. 2019. *What's in a review: Discrepancies between expert and amateur reviews of video games on Metacritic*. ACM HCI, 3(CSCW), 1–22.

Shaw, A. 2015. *Gaming at the Edge: Sexuality and Gender at the Margins of Gamer Culture*. University of Minnesota Press.

Spallaccia, B. 2023. *Hegemonic Masculinity and Toxic Technocultures: Discourse in the Review Bombing of HBO's *The Last of Us**. Anglistica AION, 27(2), 177–190.

ENDNOTES

¹ <https://howlongtobeat.com/game/145003>

² <https://maartengr.github.io/BERTopic>

³ <https://huggingface.co/google/gemma-2-2b>