HAMBURG IMPULSES FOR DEPOLARIZING MEDIA DEBATES

When journalists cover debates on topics such as climate protection, immigration, and the Middle East conflict, they often select aggressive and extreme voices. In digital communities, but also in news reporting, social understanding and a joint search for solutions to social problems hardly seem possible. When public debates hinder political decision-making, instead of critically and constructively considering solutions, then democracy as a whole will suffer. Increasingly, people are expressing their frustration by voting for far-right, anti-democratic parties. The media often contribute to polarization by presenting politics as a ritualistic exchange of blows between extreme voices instead of a joint search for solutions.

This is the problem the Hamburg impulses for depolarizing media debates address. Based on scientific findings and practical experience, scientists, journalists, community moderators and representatives of civil society organizations have developed recommendations on how journalists and digital community moderators can contribute to a better debate culture. These ideas are intended as food for thought, as an invitation to reflect on one's own role and to come to one's own conclusions for personal professional practice. The starting point for this paper is the work of the program "Depolarizing Public Debates: Developing the Tools for Transformative Communication" and a joint workshop at The New Institute in Hamburg in 2024. The aim is to depolarize debates without restricting the diversity and competition of different ideas.

Polarization means the division of society into groups that disagree on the most fundamental social issues and no longer see each other as legitimate participants in a joint dialogue. This is a particular problem when the issues affect everyone and require common solutions. Excessive polarization complicates or blocks political problem-solving, which we can currently observe particularly in debates on climate protection.

Polarization can be a productive part of democracy when it presents clear alternatives represented by specific stakeholders. Reporting on controversial debates and encouraging discussion that includes a wide range of ideas can both be part of a well-functioning democracy. For certain challenges, such as climate protection, solutions do not lie somewhere in the middle between two extremes: A little climate protection won't be enough. Hence, radical ideas also need to be discussed. The fact that emotions can boil over in the debate might also be an unavoidable part of a committed debate.

Yet excessive affective polarization does not promote a healthy democratic culture. Such affective polarization can be seen in increased hostility between groups, the silencing of moderate voices, the stigmatization of opposing positions, or the ignoring of creative solutions beyond a pointed pro and con debate. In such debates, the participants and the audience each choose their side and then ignore or demonize the other. Our proposals address these undesirable developments.

The aim should be a lively exchange of ideas that supports democratic listening, that is an exchange that enables media users to reflect on and engage with opposing positions. Such exchange can lead to partial agreement or to justified dissent. Democratic listening can also promote mutual respect and thus mitigate harmful forms of polarization.

How should journalism deal with highly polarized debates, both in its reporting and in its role as a moderator of debates, for example in the comment columns of online articles? What should platform providers do?

Polarization is a process of fragmentation that can be fueled or contained by mediated communication. Mediated polarization especially includes debates that focus not on the concerns of everyone but on the provocative behavior of their most extreme representatives. Politics and journalism shape debates through framing (what is at stake, what problems and solutions are being discussed?) and narration (which and whose stories are being told?). Operators of digital networks and forums can intervene in debates by ranking, deleting, or moderating. A lot is already happening: in some countries, platforms and media companies are getting involved in deleting illegal content or by identifying and fact-checking false information. Much of what we are calling for here is already being done. However, we are also convinced that more can be done. At the very least, the following suggestions can be read as encouragement that we are already on the right track.

This document is not intended as the ultimate truth, but as a collection of suggestions for constructive mediated discussions and journalistic reporting. We are interested in improving and supplementing the suggestions listed here and ask for constructive criticism. We will produce a revised version of these suggestions in 2025.

How journalism can deal with polarized debates:

Impulse 1: Journalism should take up social conflicts and even radical ideas - if these address relevant problems productively and do not violate the fundamental values of democracy**.**

> Simply ignoring existing conflicts and supposedly too radical positions in debates is not a suitable means of depolarization. After all, truth and solutions do not always lie in the middle (Boudana, 2015). Conflict-prone ethical issues, such as distributive justice or freedom and moral duties, are triggers for polarization (Mau et al., 2024), but they cannot simply be ignored. Political problems cannot be reduced to purely technical questions and solutions (Pepermans & Maesele, 2016). However, the discussion of radical proposed solutions reaches its limits where they violate fundamental values, such as democratic procedures, human dignity and civil rights.

Impulse 2: Journalistic reporting should focus on the search for solutions to problems in society, instead of merely emphasizing the potential for conflict.

> Reporting often takes the form of conflict framing: prominent voices confront each other with aggressive statements (Semetko & Valkenburg, 2000; Wagner & Gruszczynski, 2018). "Polarizing figures" (Slater & Arugay, 2018) or "entrepreneurs of polarization" (Mau et al., 2024) are rewarded with media attention. Instead, reporting can specifically give a voice to actors who build bridges and propose solutions to problems. The solution frame primarily discusses the "how and not the whether," in the case of generally shared goals such as climate protection or asylum law. This allows for a pluralistic, controversial, but also constructive debate.

Impulse 3: Journalistic reporting can raise awareness of the dilemmas often present in political decision-making.

> Behind escalating debates are sometimes conflicts of values in which different goods have to be compared (for example, if a railroad line is to run through a nature reserve). Journalism can make these dilemmas clear and thus enable media users to develop a better understanding of the political decision-making (Pörksen & Schulz v. Thun, 2020). The result of political trade-offs can also be critiqued.

Impulse 4: Journalistic reporting should make a clear distinction between verifiable facts and political opinions. False factual claims should be corrected immediately or not quoted at all**.**

> The research process, with its critical discussion of uncertainties, is sometimes portrayed as a failure of science or as a duel between competing scientists (example: Covid-19). Instead, scientific debates can also be framed as a "pluralist search for consensus" (Fahy, 2018). In some cases, facts are also reduced to opinions, in which "the other side" (deniers of facts) are also allowed to have their say. This leads to "false balance" (Boykoff & Boykoff, 2004). Alternatives are contextualizing reporting ("interpretive reporting," Brüggemann & Engesser, 2016) and the explicit naming of misinformation and scientific consensus (Merkley, 2020).

Impulse 5: Journalistic reporting on polarization and conflict should reflect scientific data on controversy and consensus in a society.

> Media coverage sometimes describes society as deeply divided between increasingly extreme groups. However, this does not apply to German society in such general terms (e.g. Mau et al., 2024). Even in more polarized societies, such as the USA, there is the

phenomenon of "false polarization": the perception that there are deep divides between groups when there actually are not (Levendusky & Malhotra, 2016; Wilson et al., 2020). Journalism should therefore also research the facts and can also highlight areas of consensus in public opinion (including in a playful way: Public opinion quiz: what percentage of Republicans do you think deny climate change?, Scacco et al., 2016).

Impulse 6: Journalistic storytelling should depict behaviors and contexts in ways that are representative of the respective group and avoid instead providing misleading representations of the world.

It may be factually correct to report on the fate of a specific individual, but it may imply a false generalization, for instance reporting, "Two climate change activists fly to Bali on vacation - that's how hypocritical they all are." Such false generalizations fuel conflicts (Gottschalk, 2017). Instead of judging and evaluating facts, people evaluate whether they deem specific people as sympathetic or unsympathetic (Bloom, 2016).

Impulse 7: Journalistic storytelling should not only allow readers to identify with specific perspectives, but also to distance themselves from viewpoints.

In conflicts, people tend to quickly take a side and then identify with only one of the groups involved. This leads to one-sided attention and the ignoring of other perspectives (Woodward et al., 2024; Simas et al., 2020; Breithaupt, 2017). It is therefore recommended to tell a story not just from one perspective; to be careful with stereotypical narrative schemas, such as victim narratives or hero stories; and also to present possible commonalities between different perspectives.

Impulse 8: Media organizations can go beyond reporting and develop discussion formats that encourage constructive discussion in their communities.

Many editorial teams see "community engagement" as getting news users to superficially click, comment, share, or otherwise digitally promote news articles and the news brand. However, newsrooms can also engage their communities by getting news users and other community members to talk to each other on an individual or local level (e.g. My Country Talks/Germany Speaks). This can facilitate understanding *and* strengthen loyalty to the journalistic brand.

How digital community management can deal with polarized debates:

Impulse 1: Discussions must be actively moderated so that they do not lead to excessive polarization.

> Dedicated and consistent community moderation is needed to counteract harmful polarization tendencies in discussions. Moderators should establish clear guidelines that are adapted to the communication habits of the respective target group and platform and that are then consistently implemented by moderators. Such community moderation can also counteract the trend that people with moderate opinions increasingly withdraw from online discussions (Das NETTZ, 2024; Ziegele et al., 2019).

Impulse 2: Polarization in comment sections and on social media is characterized less by echo chambers than by trench warfare between hostile parties. Moderation should try to defuse trench warfare.

> Polarized media users often confirm their opinions by connecting with like-minded people, but also do not completely isolate themselves from opposing information. They are aware of and hostile toward the extreme voices of the other side that are prominent online (Brüggemann et al., 2020; Bail, 2021). In such trench warfare situations, it is not enough to feed polarized media users information with opposing opinions. Due to the hostile atmosphere in trench warfare, information from the respective opposing side can even reinforce the polarization of users instead of weakening it (Karlsen et al., 2017; Bail et al., 2018; Törnberg, 2022).

Impulse 3: Moderation should be multifaceted: regulating, supportive, encouraging and listening-focused.

> Good moderation can do more than just remove illegal statements. Studies have shown that diverse moderation styles (Ksiazek & Springer, 2020; Ziegele et al., 2018/2019) can reinforce desirable behavior in digital communities. A shift in focus away from moderation that primarily regulates problematic statements towards moderation that supports and encourages constructive contributions has a positive effect on the perceived quality of discourse (Masullo et al., 2022; Ziegele et al., 2021). Moderators can also experiment with moderation styles that promote democratic listening, for example by accompanying opposing information with incentives to listen; encouraging users to discuss content; and adopting perspectives and sharing personal experiences even in heated debates.

Impulse 4: Moderation requires professionalized moderators and resources.

> Community management is not a task that can be done on the fly. It is a clearly defined task (Pein, 2016) that requires appropriate skills and resources (Riedl, 2023). Employers must ensure that people in these roles have at least basic training in moderation and receive regular ongoing training. In addition, the psychological risks of moderating hateful and harmful posts (see e.g. Newton, 2019 or BVCM study, 2023) must be taken into account and considered within the legal frames of protection and safety for employees.

Impulse 5: The success of the moderation should be evaluated. In addition to the reach, metrics should capture the quality of the discussions generated.

> The reach of a post is not in itself an adequate measure to assess the success of moderation or the quality of online discussions. Rather, metrics must be developed that demonstrate the contribution of community moderation to a constructive discourse and to enable a

qualitative assessment of the discussions. These include, for example, the degree of polarization, the level of respectful exchange and mutual understanding. Suggestions from journalism and deliberation research can be the starting point for collaborations between media organizations and academia that enable a more comprehensive evaluation of online discussions (Risch et al., 2020; Behrendt et al., 2024).

Impulse 6: Human moderation has its limits. This is why the supportive use of AI makes sense - as long as it is based on fundamental social and ethical values of the digital discussion space.

Artificial intelligence can support community management in scaling moderation by pre-sorting comments, hiding or reporting clearly illegal comments, suggesting response options and recognizing conspicuous users and behaviors (Wojcieszak et al., 2021; Wilms et al., 2024). A critical approach to AI is necessary for both technical and ethical reasons. For example, training data and algorithms must be known and AI interventions must be regularly checked by humans. Furthermore, a transparent approach to the use of AI vis-à-vis the community is desirable.

Impulse 7: Moderation is only one design feature. In addition, editorial teams should experiment with other interventions that promote thoughtful and factual rather than just quick emotional reactions.

Moderators can get creative to try out new forms of discussion. For example, classic, emotionally charged reaction buttons could be replaced by buttons that are geared towards listening. Buttons can be used to express respect for a different opinion (Stroud et al., 2017) or gratitude for a valuable suggestion. This makes listening practices more visible in online discussions. This approach of marking "listening" rather than the expression of quick emotions could also be a criterion for ranking posts.

References

Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The "nasty effect": Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication, 19*, 373–387.

Argyle et al. L.P., Bail, C.A., Busby, E.A., Gublera, J.R., Howe, T., Ryttingc, C., Sorensen, T., & Wingate, D. (2023). Leveraging AI for democratic discourse: Chat interventions can improve online political conversations at scale. *PNAS, 120*(41). https://doi.org/10.1073/pnas.2311627120

Bail, C.A., Argyle, L.P., Brown, T.W., Bumpus, J.P, Chen, H., Fallin Hunzaker, M.B., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *PNAS, 115*(37), 9216-9221. https://doi.org/10.1073/pnas.1804840115

Behrendt, M., Wagner, S. S., Ziegele, M., Wilms, L., Stoll, A., Heinbach, D., & Harmeling, S. (2024). AQuA--Combining experts' and non-experts' views to assess deliberation quality in online discussions using LLMs. *arXiv preprint arXiv:2404.02761*.

Bloom, P. (2017). *Against empathy: The case for rational compassion*. Random House.

Bode, L., & Vraga, E. K. (2015). In related news, that was wrong: The correction of misinformation through related stories functionality in social media. *Journal of Communication, 65*, 619–638. https://doi.org/10.1111/jcom.12166

Boudana, S. (2015). Impartiality is not fair: Toward an alternative approach to the evaluation of content bias in news stories. *Journalism*. https://doi.org/10.1177/1464884915571295

Boykoff, M. T., & Boykoff, J. M. (2004). Balance as bias: Global warming and the US prestige press. *Global Environmental Change, 14*(2), 125–136. https://doi.org/10.1016/j.gloenvcha.2003.10.001

Breithaupt, F. (2017). *Die dunklen Seiten der Empathie*. Suhrkamp.

Breithaupt, F. (2022). *Das narrative Gehirn: Was unsere Neuronen erzählen*. Suhrkamp.

Breithaupt, F., Li, B., & Kruschke, J. K. (2022). Serial reproduction of narratives preserves emotional appraisals. *Cognition & Emotion, 36*(4), 581–601.

Brüggemann, M., & Engesser, S. (2017). Beyond false balance: How interpretive journalism shapes media coverage of climate change. *Global Environmental Change, 42*, 58–67. https://doi.org/10.1016/j.gloenvcha.2016.11.004

Evertz, K., & Evertz, S. (2023). Social Media und Community Management – BVCM-Studie 2023. Nordkirchen: Bundesverband Community Management e. V. – für Social Media und digitale Kommunikation.

Fahy, D. (2018). Objectivity as trained judgment: How environmental reporters pioneered journalism for a "post-truth" era. *Environmental Communication: A Journal of Nature and Culture, 12*(7), 855–861. https://doi.org/10.1080/17524032.2018.1495093

Farjam, M., & Bravo, G. (2024). Do you really believe that? The effect of economic incentives on the acceptance of real-world data in a polarized context. *Royal Society Open Science, 11*(4), 240252.

Gottschall, J. (2021). *The story paradox: How our love of storytelling builds societies and tears them down*. Basic Books.

Jackson, S. J., & Kreiss, D. (2023). Recentering power: Conceptualizing counterpublics and defensive publics. *Communication Theory, 33*(2-3), 102-111.

Jennings, J., Stroud, N. J., & Van Duyn, E. (2018). Chicago news landscape. *Center for Media Engagement*. https://mediaengagement.org/research/chicago-news-landscape/

Das NETTZ, GMK, HateAid, NdM; Kompetenznetzwerk gegen Hass im Netz (Hrsg.) (2024). Lauter Hass – leiser Rückzug. Wie Hass im Netz den demokratischen Diskurs bedroht. Berlin. https://kompetenznetzwerk-hass-im-netz.de/download_lauterhass.php

Karlsen, R., Steen-Johnsen, K., Wollebæk D., & Enjolras, B. (2017). Echo chamber and trench warfare dynamics in online debates. *European Journal of Communication, 32*(3), 257–273. https://doi.org/10.1177/0267323117695734

Kashima, Y. (2000). Maintaining cultural stereotypes in the serial reproduction of narratives. *Personality and Social Psychology Bulletin, 26*(5), 594-604.

Ksiazek, T. B., & Springer, N. (2020). *User comments and moderation in digital journalism: Disruptive engagement*. Routledge.

Levendusky, M., & Malhotra, N. (2016). Does media coverage of partisan polarization affect political attitudes? *Political Communication, 33*(2), 283–301. https://doi.org/10.1080/10584609.2015.1038455

Masullo, G. M., Riedl, M. J., & Huang, Q. E. (2022). Engagement moderation: What journalists should say to improve online discussions. *Journalism Practice, 16*(4), 738-754.

Masullo, G. M., Tenenboim, O., & Lu, S. (2021). "Toxic atmosphere effect": Uncivil online comments cue negative audience perceptions of news outlet credibility. *Journalism, 24*, 101–119. https://doi.org/10.1177/14648849211064001

Masullo, G. M., Gonçalves, J., Weber, I., Laban, A., Torres da Silva, M. & Hofhuis, J. (Juni, 2021). How the public views deletion of offensive comments. *Center for Media Engagement*. https://mediaengagement.org/research/how-the-public-views-deletion-of-offensive-comments

Mau, S., Westheuser, L., & Lux, T. (2024). Triggerpunkte: Konsens und Konflikt in der Gegenwartsgesellschaft. *Suhrkamp*.

Meyer, H., Rauxloh, H., Farjam, M., & Brüggemann, M. (2023). From disruptive protests to discursive polarization? Comparing German news on Fridays for Future and Letzte Generation. *OSF Preprints*. https://doi.org/10.31219/osf.io/jkaw8

Muddiman, A. (2024). Robust comment sections need robust resources. *Journal of Media Ethics, 39*(2), 135–136. https://doi.org/10.1080/23736992.2021.1976646

Newton, C. (2019). The trauma floor: The secret lives of Facebook moderators in America. *The Verge*. https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona

Pein, V. (2016). Bundesverband für Community Management e.V. Stellenprofil für Community Manager. https://www.bvcm.org/wp-content/uploads/2021/11/160226-Whitepaper-Stellenprofil-Community-Manager.pdf

Pepermans, Y., & Maeseele, P. (2016). The politicization of climate change: Problem or solution? *Wiley Interdisciplinary Reviews: Climate Change, 7*(4), 478–485. https://doi.org/10.1002/wcc.405

Pradel, F., Zilinsky, J., Kosmidis, S., & Theocharis, Y. (2024). Toxic Speech and Limited Demand for Content Moderation on Social Media. *American Political Science Review*. Published online 2024:1-18. https://doi.org/10.1017/S000305542300134X

Quiz Creator. *Center for Media Engagement*. https://mediaengagement.org/quiz-creator/

Relevanz von Community Management: Wie moderierende Publisher das Engagement von Nutzenden steigern – dfV Marktstudie 2024. https://dfv-strapi.japi.dev/uploads/240613_dfv_Marktstudie_Relevanz_von_Community_Management_ba3367448a.pdf

Riedl, M. J. (2023). Journalism as a profession of conditional permeability: A case study of boundaries in a participatory online news setting. *Journalism, 24*(4), 691-708.

Risch, J., Stoll, A., Wilms, L., & Wiegand, M. (September, 2021). Overview of the GermEval 2021 shared task on the identification of toxic, engaging, and fact-claiming comments. In *Proceedings of the GermEval 2021 shared task on the identification of toxic, engaging, and fact-claiming comments*, 1-12.

Scacco, J. M., Muddiman, A., & Stroud, N. J. (2016). The influence of online quizzes on the acquisition of public affairs knowledge. *13*(4), 311–325. https://doi.org/10.1080/19331681.2016.1230920

Semetko, H. A., & Valkenburg, P. M. (2000). Framing European politics: A content analysis of press and television news. *Journal of Communication, 50*(2), 93–109. https://doi.org/10.1111/j.1460-2466.2000.tb02843.x

Simas, E. N., Clifford, S., & Kirkland, J. H. (2020). How empathic concern fuels political polarization. *American Political Science Review, 114*(1), 258-269.

Stroud, N. J., Muddiman, A., & Scacco, J. M. (2017). Like, recommend, or respect? Altering political behavior in news comment sections. *New Media & Society, 19*, 1727–1743. https://doi.org/10.1177/1461444816642420

Stroud, N. J., Scacco, J. M., Muddiman, A., & Curry, A. L. (2015). Changing deliberative norms on news organizations' Facebook sites. *Journal of Computer-Mediated Communication, 20*, 188–203. https://doi.org/10.1111/jcc4.12104

Thier, K., & Lin, T. (2022). How solutions journalism shapes support for collective climate change adaptation. *Environmental Communication, 16*(8), 1027–1045. https://doi.org/10.1080/17524032.2022.2143842
Törnberg, P. (2022). How digital media drive affective polarization through partisan sorting. *Proceedings of the National Academy of Sciences, 119*(42), e2207159119. https://doi.org/10.1073/pnas.2207159119

Van Duyn, E., & Collier, J. (2019). Priming and Fake News: The Effects of Elite Discourse on Evaluations of News Media. *Mass Communication and Society, 22*(1), 29–48. https://doi.org/10.1080/15205436.2018.1511807

Wagner, M. W., & Gruszczynski, M. (2018). Who gets covered? Ideological extremity and news coverage of members of the U.S. congress, 1993 to 2013. *Journalism & Mass Communication Quarterly, 95*(3), 670–690. https://doi.org/10.1177/1077699017702836

Wegner, D. M., & Gray, K. (2017). *The mind club: Who thinks, what feels, and why it matters*. Penguin.

Wilms, L. K., Gerl, K., Stoll, A., & Ziegele, M. (2024). Technology acceptance and transparency demands for toxic language classification—interviews with moderators of public online discussion fora. *Human–Computer Interaction,* 1-26.

Wilson, A. E., Parker, V. A., & Feinberg, M. (2020). Polarization in the contemporary political and media landscape. *Current Opinion in Behavioral Sciences, 34*, 223-228.

Wojcieszak, M., Thakur, A., Ferreira Gonçalves, J. F., Casas, A., Menchen-Trevino, E., & Boon, M. (2021). Can AI enhance people's support for online moderation and their openness to dissimilar political views? *Journal of Computer-Mediated Communication, 26*(4), 223-243.

Woodward, C., Hiskes, B., & Breithaupt, F. (2024). Spontaneous side-taking drives memory, empathy, and author attribution in conflict narratives. *Discover Psychology, 4*(1), 1-13.

Ziegele, M., Heimbach, D. (2021). Konstruktive Online-Debatten fördern durch bestärkende Moderation.

Ziegele, M., Jost, P., Bormann, M., & Heinbach, D. (2018). Journalistic counter-voices in comment sections: Patterns, determinants, and potential consequences of interactive moderation of uncivil user comments. *SCM Studies in Communication and Media, 7*(4), 525-554.

Ziegele, M., Jost, P., Frieß, D., & Naab, T. (2019). Aufräumen im Trollhaus: Zum Einfluss von Community-Managern und Aktionsgruppen in Kommentarspalten. *DIID-Précis*.