Supporting Respectful and Constructive Online Discussion using Moral Reframing

HYUNWOO KIM, School of Computing, KAIST, Republic of Korea
VEDAD MISIRLIĆ, Graz University of Technology, Austria
JUHO KIM, School of Computing, KAIST, Korea

Online discussion platforms like Reddit, X(formerly Twitter), and Facebook are some of the most common spaces for people to discuss and exchange their opinions. However, online discussion spaces can often be filled with hostility, negativity, and toxicity towards people with different opinions, hindering civil and constructive discussion. In this position paper, we aim to explore the applicability of *moral reframing*, a communication strategy of reframing one's opinion to be aligned with the moral convictions of the conversation partner, in the context of online discussion. We conducted an interview study with 13 occasional online discussion participants to understand their perception of moral reframing and the potential challenges of applying the strategy in online discussions. Based on the findings, we propose that technical support for moral reframing should support understanding the moral frames of other comments and building one's own argumentation that aligns with the moral frames identified while maintaining the ownership of the user in shaping their own ideas. We plan to implement the idea as an interactive system and explore how such a system would contribute to more civil and constructive online discussions.

ACM Reference Format:
Hyunwoo Kim, Vedad Misirlić, and Juho Kim. 2024. Supporting Respectful and Constructive Online Discussion using Moral Reframing. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX)*. ACM, New York, NY, USA, 5 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Online discussions are one of the most common spaces for people to discuss and exchange their opinions on social issues. By sharing and exchanging opinions, people learn the thoughts and opinions of fellow citizens and form their own informed opinions on the social issue [7].

However, it is often challenging to have a civil and constructive conversation in online discussions. Instead of acknowledging others and finding common ground between different perspectives, people usually have polarized views while expressing hostility towards others. Some existing works from human-computer interaction focused on designs for exposing diverse viewpoints to the user [8, 13, 14], pointing out the lack of exposure to different opinions as a leading cause of polarization. On the other hand, recent research has shown that simply increasing exposure to different viewpoints may backfire to increase polarization [4], which would hinder civil and constructive conversation. To mitigate that, recent research points out that people's receptiveness towards different opinions needs to be improved [11]. For instance, some recent work tried re-expressing the opinions of others using large language

Authors' Contact Information: Hyunwoo Kim, khw0726@kaist.ac.kr, School of Computing, KAIST, Daejeon, Republic of Korea; Vedad Misirlić, vedad.misirlic@student.tugraz.at, Graz University of Technology, Graz, Austria; Juho Kim, juhokim@kaist.ac.kr, School of Computing, KAIST, Daejeon, Korea.

Manuscript submitted to ACM

models to add signals of receptiveness, such as adding some words of acknowledgment, which contributed to more constructive conversation [2, 10].

In an effort to improve receptiveness towards different opinions, we propose **moral reframing** as a solution to support more civil and respectful online discussion. Moral reframing refers to the persuasive strategy in which people reframe their arguments to be aligned with the moral convictions of individuals with differing opinions [6]. By doing so, moral reframing acknowledges the difference in moral convictions and the importance of moral values that underlie people's diverse opinions and leverages the differences to understand others better and persuade others [6]. A series of works has shown its effectiveness for changing opinions of the others, especially on politicized issues [1, 3, 5]. However, there has been less research on supporting people to use moral reframing strategies in online discussion setting.

To investigate how to design technical support for moral reframing to improve online discussion, we first interviewed 13 frequent online discussion participants. In the interview, we explained the moral reframing strategy and asked about their perceptions of the strategy, how it aligns with users' motivation for participating in online discussions, and the right level of technical support in online discussions regarding maintaining ownership over the original comment. Participants commonly valued moral reframing as a tool to facilitate civil and respectful discussions by enabling them to think from others' perspectives, which was the primary motivation for participating in online discussions. At the same time, they had concerns about how to understand the moral convictions of others from short comments with some reluctance to give up their moral convictions. They also agreed that the technical support should remain at the level of assistance while preserving their autonomy in expressing their own opinion. Based on the findings from the formative study, we propose a set of design considerations to technically support moral reframing in online discussions to facilitate constructive discussion. We envision that our approach of technically supporting moral reframing could contribute to positive social technology as a potential approach to support civil and respectful communication between people with different viewpoints and backgrounds by understanding each other better.

2 Online Discussion Participants Interview

To understand how moral reframing could be applied in the current online discussion space, we conducted formative interviews. We specifically aimed to understand two factors: a) people's perceptions of moral reframing and potential challenges, and b) perceptions of the right level of technical support in online discussions.

We recruited 13 participants who participated in online discussions on social issues at least three times per week for a 1-hour interview. We recruited the participants using the local university board, X (formerly Twitter), Reddit, Upwork, and Prolific. Participants' ages ranged from 19 to 61 and had diverse nationalities, including the US, Kenya, and India. Participants mainly used Reddit, X, and Facebook as their main channels for participating in online discussions.

2.1 Findings

Here, we report the findings from our interview.

2.1.1 Perceptions on Moral Reframing. P2, P3, P4, P6, P8, P9, P11, P12 and P13 all stated that their main goal is to find common ground in online discussions, with many of them stating that the online discussion place can be a very toxic and unconstructive environment. During the interview, we introduced moral reframing as a way of understanding how other individuals think about the issue and delivering their arguments better to others to reach a consensus. While the participants were unaware of the term itself, almost all participants feel familiar with the concept, often considering it as a way of "putting yourself in other people's shoes" P11. Extending on the experience of taking others' perspectives,

all participants but P5 considered moral reframing as an effective method of understanding others and finding common ground with them.

However, participants raised some practical concerns about applying the strategy. P10, P12, and P13 pointed out the difficulty of grasping other people's morals from a small amount of information available from online discussions, which, in the case of Reddit, is often limited to the comments and the profile information . In the words of P13, "Connecting moral frames with arguments might be hard. Even though there are some commonalities, it might be hard to start the discussion with them.", we can see that there can be issues in understanding the moral foundations on a deeper, semantic level.

Furthermore, some participants expressed concern about aligning their beliefs with other people, which might make them feel like they have lost ownership of the argument they were making. P11 mentioned that using the strategy may feel like giving up their beliefs too much by saying that they are "Opening a bit too much to the other side", as well as saying that "Many times there would be a certain degree to which I have to agree with the other person".

2.1.2 Perceptions on Ideal Level of System Support. Participants valued the online discussion space as a space for people to openly and respectfully exchange their views. In that sense, participants expected that technical support, especially with AI, could positively contribute to online discussion by helping users form more logical arguments with better-quality writing. For example, people considered AI intervention as useful in the context of refining grammar and the tone of text (P1, P2, P10, P11), playing devil's advocate (P7, P11, P13), giving unbiased opinions on an issue (P4, P9), and providing background knowledge on a topic (P8). However, at the same time, most participants noted that the opinions shared in the discussion should be the thoughts of the discussion participants themselves and argumentation support tools should not violate people's agency in making their arguments.

3 Design Considerations

Based on the interview findings, we propose necessary features and potential design considerations for technically supporting the moral reframing strategy in online discussions to facilitate more constructive and respectful discussions in online spaces.

C1. Support identifying and understanding the moral convictions of other discussion participants. When applying moral reframing, our participants were concerned about how to identify and understand the moral values of others from a short interaction from the comments. Therefore, we propose that the system should be designed to support identifying and understanding the moral convictions of others. For instance, the system may present labels of moral convictions in each comment, along with explanations on how the comment relates to the conviction, to support the user in thinking about the moral conviction of others. We expect that a technical pipeline could support the identification and analysis of moral convictions by leveraging existing NLP work for detecting moral values of text [12, 16, 17]. At the same time, we acknowledge that identifying moral frames requires subjective interpretation, which may lead to presenting erroneous moral frames to the users. Therefore, the system should also consider potential interactions for the users to mitigate errors, such as explaining the detection results to enable users to make their own decision on the accuracy.

C2: Provide active support for reshaping users' arguments into the moral convictions of other discussion participants. When using moral reframing, we expect that the user will experience challenges in writing appropriate claims and logic to argue their stance, as they need to develop a new set of claims to align with the moral convictions of others. Therefore, to support users, we expect the system to support ideating various potential claims to argue their stance.

For example, the system could provide suggestions for the users on how to claim their stance on the issue while considering the moral convictions of their conversation partners, which was shown to be effective in existing work such as AngleKindling [15]. Furthermore, the system may provide feedback on whether the user's draft aligns with the target moral frame to support users iteratively refining the arguments with the moral reframing strategy.

C3: Support in ensuring logical coherence of arguments. Participants commonly considered logically concrete arguments with factual support essential for high-quality arguments, and they also tried their best to provide logical and factual arguments when they participated in discussions. When writing arguments with moral reframing, we expect the users to have more needs and challenges towards developing logically concrete arguments, as users need to think in terms of moral foundations that are less familiar or even opposed to their own moral convictions. Therefore, to support the users to produce high-quality arguments, we expect that the system should be able to provide feedback and potential improvements on the user's argument in terms of logical coherence, which was shown to be an effective way to improve argument quality in previous work [18].

C4: Preserve users' autonomy in shaping arguments. Participants valued online discussion as a space to express their own opinions and listen to the opinions of other people, expecting the system not to violate the users' agency to express their views. Furthermore, when the system tries to align users' arguments with unfamiliar or even disagreeing moral values, it may harm their moral integrity, making them reluctant to use the moral reframing strategy [9]. In that sense, we propose that the support from the system should remain at the level of presenting suggestions and feedback for reframing their arguments, while the user should take the responsibility of choosing and adopting them to form their arguments. For example, the system may present suggestions for arguing against the selected comment on demand and enable the user to adopt tolerable suggestions, rather than directly editing user's comment such as autocomplete or automatic fixes.

4 Future Work

We introduced a set of design considerations to support users applying moral reframing to contribute to healthier online discussions. In future work, we aim to implement this idea into an interactive system and explore the implications and scalability of this technique on making the online discussion space a more friendly, constructive, and safe environment. Through our tool, we aim to analyze and quantify the effect that moral reframing as a communication technique has on the constructiveness and politeness of online discussions. We further envision that the lessons from designing technical supports for moral reframing could be generalizable to facilitate respectful and constructive communication between people with different viewpoints and backgrounds. In turn, this would foster a safer online space, which would be a fundamental step towards leveraging social technologies with the common goal of supporting and augmenting human flourishing.

References

[1] Marie Altgilbers Roweton and Scott Eidelman. 2023. Ideology, moral reframing, and persuasion in the context of COVID-19 vaccines. *Social and Personality Psychology Compass* 17, 9 (2023), e12802. https://doi.org/10.1111/spc3.12802 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/spc3.12802.

[2] Lisa P. Argyle, Christopher A. Bail, Ethan C. Busby, Joshua R. Gubler, Thomas Howe, Christopher Rytting, Taylor Sorensen, and David Wingate. 2023. Leveraging AI for democratic discourse: Chat interventions can improve online political conversations at scale. *Proceedings of the National Academy of Sciences* 120, 41 (Oct. 2023), e2311627120. https://doi.org/10.1073/pnas.2311627120 Publisher: Proceedings of the National Academy of Sciences.

[3] Isaac Baca. 2021. Moral Reframing: A Perspective-Taking Approach to Political Arguments. *Scientia et Humanitas* 11 (June 2021), 39–58. https://libjournals.mtsu.edu/index.php/scientia/article/view/2073

[4] Christopher A. Bail, Lisa P. Argyle, Taylor W. Brown, John P. Bumpus, Haohan Chen, M. B. Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. 2018. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences* 115, 37 (Sept. 2018), 9216–9221. https://doi.org/10.1073/pnas.1804840115 Publisher: Proceedings of the National Academy of Sciences.

[5] Matthew Feinberg and Robb Willer. 2015. From Gulf to Bridge: When Do Moral Arguments Facilitate Political Influence? *Personality and Social Psychology Bulletin* 41, 12 (Dec. 2015), 1665–1681. https://doi.org/10.1177/0146167215607842 Publisher: SAGE Publications Inc.

[6] Matthew Feinberg and Robb Willer. 2019. Moral reframing: A technique for effective and persuasive communication across political divides. *Social and Personality Psychology Compass* 13, 12 (2019), e12501. https://doi.org/10.1111/spc3.12501 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/spc3.12501.

[7] James S Fishkin and Robert C Luskin. 2005. Experimenting with a Democratic Ideal: Deliberative Polling and Public Opinion. *Acta Politica* 40, 3 (Sept. 2005), 284–298. https://doi.org/10.1057/palgrave.ap.5500121

[8] Mingkun Gao, Hyo Jin Do, and Wai-Tat Fu. 2018. Burst Your Bubble! An Intelligent System for Improving Awareness of Diverse Social Opinions. In *23rd International Conference on Intelligent User Interfaces*. ACM, Tokyo Japan, 371–383. https://doi.org/10.1145/3172944.3172970

[9] Aaron Isiminger and Roger Giner-Sorolla. 2024. Willingness to use moral reframing: Support comes from perceived effectiveness, opposition comes from integrity concerns. *Social Psychological Bulletin* (March 2024). https://kar.kent.ac.uk/105295/ Accepted: 2024-03-11 Publisher: Psychopen.

[10] Gauri Kambhatla, Matthew Lease, and Ashwin Rajadesingan. 2024. Promoting Constructive Deliberation: Reframing for Receptiveness. http://arxiv.org/abs/2405.15067 arXiv:2405.15067 [cs].

[11] Julia A. Minson and Frances S. Chen. 2022. Receptiveness to Opposing Views: Conceptualization and Integrative Review. *Personality and Social Psychology Review* 26, 2 (May 2022), 93–111. https://doi.org/10.1177/10888683211061037 Publisher: SAGE Publications Inc.

[12] Negar Mokhberian, Andrés Abeliuk, Patrick Cummings, and Kristina Lerman. 2020. Moral Framing and Ideological Bias of News. Vol. 12467. 206–219. https://doi.org/10.1007/978-3-030-60975-7_16 arXiv:2009.12979 [cs].

[13] SA Munson, SY Lee, and Paul Resnick. 2013. Encouraging Reading of Diverse Political Viewpoints with a Browser Widget.. In *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*. 419–428. http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/viewFile/6119/6381

[14] Matti Nelimarkka, Salla-Maaria Laaksonen, and Bryan Semaan. 2018. Social Media Is Polarized, Social Media Is Polarized: Towards a New Design Agenda for Mitigating Polarization. In *Proceedings of the 2018 Designing Interactive Systems Conference*. ACM, Hong Kong China, 957–970. https://doi.org/10.1145/3196709.3196764

[15] Savvas Petridis, Nicholas Diakopoulos, Kevin Crowston, Mark Hansen, Keren Henderson, Stan Jastrzebski, Jeffrey V Nickerson, and Lydia B Chilton. 2023. AngleKindling: Supporting Journalistic Angle Ideation with Large Language Models. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3544548.3580907

[16] Vjosa Preniqi, Iacopo Ghinassi, Julia Ive, Charalampos Saitis, and Kyriaki Kalimeri. 2024. MoralBERT: A Fine-Tuned Language Model for Capturing Moral Values in Social Discussions. In *Proceedings of the 2024 International Conference on Information Technology for Social Good*. 433–442. https://doi.org/10.1145/3677525.3678694 arXiv:2403.07678 [cs].

[17] Markus Reiter-Haas, Simone Kopeinik, and Elisabeth Lex. 2021. Studying Moral-based Differences in the Framing of Political Tweets. *Proceedings of the International AAAI Conference on Web and Social Media* 15 (May 2021), 1085–1089. https://doi.org/10.1609/icwsm.v15i1.18135

[18] Zheng Zhang, Jie Gao, Ranjodh Singh Dhaliwal, and Toby Jia-Jun Li. 2023. VISAR: A Human-AI Argumentative Writing Assistant with Visual Programming and Rapid Draft Prototyping. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23)*. Association for Computing Machinery, New York, NY, USA, 1–30. https://doi.org/10.1145/3586183.3606800

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009