Proceedings of the 5th International Conference on International Law and Legal Policy
DOI: 10.54254/2753-7048/2026.37100

Specific Manifestations of Group Polarization on Social Media and Related Psychological Factors, with a Discussion of Interventions

Xiaoqi Ma

Immersion Academy, Irvine, USA
max758195@gmail.com

Abstract. With the development of society, groups on social media have begun to spread more widely; as such, group polarization has also occurred in a more serious form. Based on the above, this paper will review related studies on the specific forms of group polarization in social media, the psychological reasons for it, and countermeasures. Group polarization on social media is not due to a single reason but is the result of the combined effects of the information environment. A small number of reasons, selective exposure and confirmation bias can cause people to repeatedly encounter and be reinforced by information that supports their own beliefs. Social identity, conformity, and group differences also affect how users are perceived by the public. At the same time, social feedback, such as likes and comments, and emotions can also lead to such changes. Group polarization: Simply increasing exposure to information from the other side will not reduce polarization; it is also affected by the source of the information and group identity. Therefore, the intervention should consider the platform's way of spreading information, social influence in groups, and how the content is presented. polarization can be reduced in many ways, such as expanding the variety of information distribution, having credible and influential group members present balanced views, and improving the presentation of different opinions. Based on the above mechanisms and their interactions, this study can provide a theoretical basis for building a more moderate and diverse environment for public discussion online.

Keywords: Social Media, Group Polarization, Depolarization Interventions

1. Introduction

Group polarization is the phenomenon that, in a group discussion, people tend not to converge toward moderate positions; instead, they move toward more extreme positions [1]. As more people have begun to use social media in their daily lives, both the extent and types of this phenomenon have been expanding considerably. Social media has become a place where people can express their views and communicate with others; at the same time, the echo chamber effect of social platforms causes users to be exposed to similar opinions repeatedly and thus continuously strengthens their own beliefs. This phenomenon existed before the emergence of the Internet. A large amount of information in this communication environment may not be suitable for view integration. It will give

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-7048/2026.37100

some people a sense of belonging and make others feel alienated. Group polarization will occur on social media, and repeated questioning and negative expressions may reduce the scope for reasonable discussion in the public sphere. They may also continuously add more confrontational content to online communication, generate negative emotions, and thus create an extreme communication environment that hinders the collective effort to build a moderate and diverse communication space. Therefore, based on the above literature review, this paper will examine group polarization and related psychological mechanisms in social media environments and compare the effects of different de-polarization methods.

2. Definition of group polarization on social media

Group polarization is a phenomenon in which, before a discussion, members of a group who already hold a certain view on an issue are influenced by the opinions of others with similar views during the group discussion and interaction. Then the original tendency is enhanced, and the views become more extreme [1]. After discussing an issue with people who hold different opinions, the group members may not incorporate their own views and those of the opposition into a neutral position. Defensiveness or pressure related to identity may cause them to remain in their original positions even more stubbornly. Group discussion may therefore lead to the extremes of people's opinions and not narrow the differences. For example, the members of a group who support a certain policy may be strengthened in their initial position after discussing it and reaching an agreement. When facing group members who are opposed to the policy, they may also view the opposing views as a challenge to their own group's position and thus support the policy even more strongly. With the rapid development of social media, people have been exposed to and shared information through social media. Group polarization has moved away from the traditional face-to-face model and is now an online environment that includes algorithmic recommendations, real-time interaction and community clustering. Recommendation algorithms on social media continuously present information according to users' interests, thereby increasing their conviction in their own beliefs and further avoiding others. The traditional form of group interaction has had limited communication due to distance and social circles. Research shows that the pattern of interaction in a group on social media can help them communicate and fosters the development of an echo chamber [2]. Social media has made communication more convenient, and now people with similar interests and ideas can talk to each other all the time and form strong but relatively closed online social groups.

2.1. Limited argument pools

A reason for group polarization is a limited argument pool; that is, the sources of the group's arguments are restricted [1]. A few people are no longer convinced after hearing what the others have said. When the group members are in similar positions, the arguments they present will also be the same. Therefore, arguments that support the members' current positions repeatedly appear in the discussion, and other arguments are relatively rare. At the same time, when members come across more new arguments that support their original ideas through group discussion, they become more convinced of those ideas and thus take a more extreme position.

In the environment of social media, a user's online behavior helps determine which kinds of information are recommended to them and, as a result, how different types of information affect their feelings. Users can freely participate in online discussions, join social groups and follow accounts on the Internet to keep in touch with others who have similar interests at any time. At the same time, the platform's recommendation algorithm will continue to show users similar content

DOI: 10.54254/2753-7048/2026.37100

based on what they have liked in the past. Therefore, the sources of information will be more uniform, and people who are similar will be in closer contact, thus forming an echo chamber. In an echo chamber, one's own ideas are repeatedly confirmed, and other views are rarely heard; thus, there is little exchange of ideas.

Other research has also associated echo chambers with selective exposure and confirmation bias. Selective exposure refers to the tendency to select information that aligns with one's own beliefs and values. Confirmation bias is the tendency for people to believe in what they already believe and ignore or disregard evidence that contradicts their beliefs; they may selectively pay attention to such information. Therefore, although there are many sources of information on social media, only a small number of them are chosen by users to follow.

At the same time, spreading the other side's views on social media does not mean one has left the echo chamber. Begby points out that people develop selective epistemic trust in the process of obtaining information [3], and the persuasiveness of the content depends on whether it is provided by a trusted group rather than on the content itself. Group members often see people with similar opinions as peers and are thus more likely to believe their own views. However, when group members disagree with another group, they may view it as an untrustworthy out-group and be even more devoted to their own position in response to the opposition from that group. Therefore, if a group is not included in the circle of trusted groups, other groups may give it a lower epistemic weight. Even if the group has a reasonable cause, it may still be considered biased and unreliable.

2.2. Social influences on behavior

The second way that groups are polarized is through social influence on behaviour. Therefore, the members of the group adjust their positions according to the behaviour of others, evaluation pressure and group norms. Some group members may not express their own views independently during discussion. People are generally aware of what the group expects, and they also want to know whether other members of the group are in agreement with it; thus, they hope that expressions aligned with the group's position will receive approval. When the position of the majority is obvious, some members may change how they express their views to align with the group's direction, seeking approval or avoiding being out of step with the group; thus, they show strong support for the group's opinion [4]. Therefore, if no new or conflicting opinions are added, the group's views may be pushed to an extreme.

According to the theory of social identity, the group's identity affects where and in what way its members behave, and people gain a sense of self through their membership in a particular group [5]. Therefore, in accordance with the theory of social identity, group identity will further promote group polarization. When the group members clearly state their opinions, the individuals may change how they express themselves to fit in with the group and thus feel a sense of belonging, even if this opinion does not align with their own thoughts. Wuestenenk and others have shown that the effect of in-group norms in online environments is not reduced by the absence of offline, face-to-face interaction. At the same time, this influence may become even stronger in anonymous online environments [6].

2.3. Emotional expression and social feedback

The online environment of social media also promotes group polarization. Social media is not face-to-face communication, so the social feedback, such as comments, likes and shares, is less obvious. These indicators can directly show users the level of support for a particular position. Such social

DOI: 10.54254/2753-7048/2026.37100

feedback also informs the group members whether a view has been acknowledged by the group, whether it is worth continuing to support, and whether expressing the same position will receive more positive responses. Highly emotional statements are more likely to attract group attention and receive recognition from others. Due to conformity and the desire for social acceptance, the group members change their behaviour. At the same time, social media recommendation algorithms tend to spread statements that have received more social feedback and are exposed to more users, thus strengthening the original position.

User behaviour is also a reason for group polarization; frequently browsing, commenting, liking and continuing to participate in discussions on similar topics are all examples. Based on research of Facebook, Del Vicario and others have found that users' participation and activity on the platform are positively correlated with their inclination for emotional expression. The more often users interact on the platform, the more likely they are to repeatedly encounter similar positions and emotional content, and thus the probability of expressing negative emotions and producing extreme content may also increase [7]. Thus, as the group interaction gradually increases, the entire content of the discourse may become more extreme. The main type of expression is a negative one, such as questions and objections, and a very negative environment has been created in social media communication. This general pattern of negative expressions will further divide the group's attitude and increase group polarization on social media.

3. Intervention measures for group polarization on social media and their effects

Platform recommendation algorithms are one of the reasons for and a reinforcement of group polarization, as they continuously recommend similar content according to users' interests and behaviour. Discussions of depolarization strategies should first consider changes to the platform's governance system. Algorithmic recommendation is a way that platforms retain users and earn money; therefore, platform operators may be unwilling to change the recommendation mechanism in an effort to achieve decentralization. If the platform over-restricts or regulates users' speech by directly suppressing polarized expression, it will violate their freedom of expression and infringe upon their right to privacy. Therefore, Avin and others have proposed Random Linear Regulation [8], which randomly adds some users outside the original audience selected for content distribution to increase the chance that the content recommendation will reach people in different groups. Thus, the existing recommendation mechanism of the platform will not be disturbed, and it will also meet the requirements for freedom of expression and user privacy.

However, even when users are shown conflicting opinions or when the scope of information disclosure is expanded, the effect of depolarization may be limited if users do not base their judgments solely on the content of the views themselves. Begby points out inter-group belief polarization: after the group members communicate with people who hold different views, they do not change their opinions but are instead more firmly committed to them [3]. More exposure to the other side will not change people's minds. Therefore, the Depolarization Strategy should not be limited to showing users different views.

Another way of decentralization is to have influential people in a group spread the neutral view among the rest of the group. Wu and others therefore believe that depolarization interventions should also consider who is expressing these views.Compared with rebuttals from outside the group, neutral views proposed by members within the group are more likely to be accepted by other members because these individuals are generally regarded as members of the "peer group". Their opinions are more likely to be trusted, and thus may gradually lead group members to change their

DOI: 10.54254/2753-7048/2026.37100

original opinions [9]. In addition, Neutral positions communicated by socially influential group members are likely to be accepted more readily by other members.

Nelimarkka and others have studied ten Finnish Facebook groups focused on migration issues and found that even when groups with different views are exposed to the same content, they still interpret it according to their own positions [10]. For example, when the same news article is seen by both pro-migration and anti-migration groups, the pro-migration group may believe that the article proves migrants should be welcomed, while the anti-migration group may believe that it shows migrants are problematic. Hostile, aggressive and uncivil language also frequently appear in the comment sections of such content. It can be inferred that the two sides' discussions of the same content may only be an occasion to show hostility towards each other. Therefore, Nelimarkka and others propose that intervention can be introduced at the level of content publishers' edited posts [10]. For example, instead of only reporting news that supports one side, a post can also present all sides.

4. Discussion and suggestions

A substantial body of research has shown that group polarization on social media is the result of multiple social factors and psychological mechanisms, such as the information environment, individual cognition, group relationships, and platform recommendation mechanisms. Echo chambers and selective exposure cause users to repeatedly encounter information in line with their own views for a long time, and group identification continuously reinforces these existing beliefs. According to Begby, being exposed to the other side can increase one's own conviction [3]. This is because people's assessments of a place are affected by both the content of the information and its origin. When the information comes from an out-group source, users are more likely to believe that it is not true and do not change their own beliefs. At the same time, social identity and group norms also influence how people present themselves; thus, they may change the way they put forward their ideas to conform to the group. Modes of social media interaction, such as comments, likes and follows, can further spread these social norms and cause all the mechanisms to reinforce each other.

Therefore, the methods to reduce group polarization on social media cannot be used in isolation. Without affecting the profitability of the platform, freedom of expression or user privacy, it may be beneficial to use platform recommendation mechanisms to increase the quantity of content representing positions different from those held by group users or to expand the audience for content expressing a particular view, but this will not be suitable for all users. Therefore, the impact of an in-group can be used to spread neutral ideas through members with higher social influence [9]. Intervention can also modify the form of the content. Nelimarkka and others have also shown that groups with different opinions can interpret the same content differently [10]. Therefore, platforms can encourage users to use more neutral language in their posts and reduce the likelihood of aggressive comments during discussion.

5. Conclusion

This paper examined group polarization on social media and its psychological and social mechanisms. As people increasingly obtain information and participate in group communication through social media, group polarization has been changing. Several factors, such as the information environment, individual cognition, group relationships and platform recommendation mechanisms, all promote group polarization. Depolarization interventions for group polarization on social media should therefore not be used in isolation. Several ways will be used in combination to reduce the

DOI: 10.54254/2753-7048/2026.37100

amount. Research on group polarization in social media helps researchers understand the behaviour of groups and why members of an online group may hold different views. At the same time, such research can also help curb group polarization on social media. to reduce excessive confrontation and aggression in online discussions, lower the risk of an extreme communication environment, and support the construction of a more moderate and diverse online public space.

References

[1] Sunstein, C. R. (2002). The law of group polarization. *Journal of Political Philosophy, 10*(2), 175–195.
[2] Pranesh, S., & Gupta, S. (2024). The impact of social media on polarization in the society. arXiv preprint arXiv:2408.12877.
[3] Begby, E. (2024). From belief polarization to echo chambers: A rationalizing account. *Episteme, 21*(2), 519–539.
[4] Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology, 55*(1), 591–621.
[5] Hogg, M. A. (2016). Social identity theory. In *Understanding peace and conflict through social identity theory: Contemporary global perspectives* (pp. 3–17). Springer International Publishing.
[6] Wuestenenk, N., van Tubergen, F., & Stark, T. H. (2025). The influence of group membership on online expressions and polarization on a discussion platform: An experimental study. *New Media & Society, 27*(1), 225-245.
[7] Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Echo chambers: Emotional contagion and group polarization on Facebook. *Scientific Reports, 6*(1), 37825.
[8] Avin, C., Daltrophe, H., & Lotker, Z. (2024). On the impossibility of breaking the echo chamber effect in social media using regulation. *Scientific Reports, 14*(1), 1107.
[9] Wu, Y., Li, L., Yu, Q., Gan, J., & Zhang, Y. (2023). Strategies for reducing polarization in social networks. *Chaos, Solitons & Fractals, 167*, 113095.
[10] Nelimarkka, M., Laaksonen, S. M., & Semaan, B. (2018, June). 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* (pp. 957-970).