

Josuweit, Ch. (2026). Should I speak or should I go? Relationship closeness and descriptive norms on reactions to hate memes in private chat groups. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 20(4), Article 14. <https://doi.org/10.5817/CP2026-4-14>

Should I Speak or Should I Go? Relationship Closeness and Descriptive Norms on Reactions to Hate Memes in Private Chat Groups

Christina Josuweit^{1,2}

¹ Faculty of Educational Sciences, University of Duisburg-Essen, Essen, North Rhine-Westphalia, Germany

² Faculty of Applied Social Sciences, TH Köln, Cologne, North Rhine-Westphalia, Germany

Abstract

Visually and humorously packaged hate messages spread quickly on social media and thus may be shared in private chat groups, where they may trigger conflicts. In such situations, individuals must decide whether to speak up, remain passive, and/or leave the group. This study investigates how relational dynamics influence individual responses in such potentially contentious situations. Specifically, we examined the effects of relationship closeness and conformity to descriptive norms on the intention to comment and the intention to leave a group following the posting of a racist meme. A total of 1,090 German-speaking participants (60% female, 40% male, 1% diverse) with a mean age of 49.8 (SD = 13.84, range = 24–87) took part in an online experiment. Participants were randomly assigned to one of six conditions in a 2×3 factorial design, manipulating relationship closeness (low, high) and group response to a racist meme (no reaction, full reaction, mixed reaction). Higher relationship closeness to the group significantly increased the intention to comment on the meme, whereas participants in the high-closeness condition reported a significantly lower intention to leave the group. Contrary to expectations, group response had no significant effect on the intention to comment, challenging assumptions about conformity pressures in online groups. The results suggest that relationship closeness is more influential than mere descriptive norms in private group settings, contributing to a refined understanding of social impact and conformity processes in digital environments.

Keywords: group dynamics; hate memes; social impact; conformity; private chat groups

Editorial Record

First submission received:
October 18, 2025

Revision received:
May 31, 2026

Accepted for publication:
July 28, 2026

Editor in charge:
Lenka Dedkova

Introduction

With the widespread adoption of smartphones and instant messaging applications such as WhatsApp (DataReportal, 2025), private chat groups have become central environments of everyday communication, fostering constant, active, intensive, and intimate exchanges (Harrison & Gilmore, 2012; Knop-Hülß, 2023; Winkler et al., 2022; YouGov, 2024). The present study focuses specifically on everyday private chat groups in which access is typically tied to existing social ties of varying closeness and communication remains largely inaccessible to outsiders and the broader public (Carpenter et al., 2018). These characteristics make private chat groups relevant for studying how social influence operates in digitally mediated communication.

Among the many forms of content shared in private chat groups, hate memes are particularly interesting from a social psychological perspective. Memes in general are a hybrid of text and image designed for quick distribution (Jenkins et al., 2018; Wiggins & Bowers, 2015). While memes are often used to humorously express opinions and ideas (Leiser, 2022; Southern, 2022; Wagener, 2021), they also serve important social functions in everyday communication, especially by fostering social connection and self-disclosure (Peng & Chen, 2025). These social functions may become particularly relevant when the shared content is socially problematic or hostile: The use of hate memes which we understand as a visual representation of online hate speech (OHS; Schmitt et al., 2020) is one example of such content. In German research discourse, the concept is often associated with the notion of group-focused enmity (Zick et al., 2008) and encompasses devaluing and hostile expressions targeting specific groups and minorities, for example in the form of sexism, racism, or anti-Semitism (Sponholz, 2021; Wachs & Wright, 2018). A recent study in Germany found that 39% of young people had encountered hate content in private chats (Medienpädagogischer Forschungsverbund Südwest [mpfs], 2022).

Hate memes have proven effective in propagating hate (Quintal & Collen, 2023; Schmid, 2023; Schmid et al., 2022; Schmid et al., 2025). Their often humorous structure introduces interpretive ambiguity, particularly when conveying hostile or discriminatory content (Yoon, 2016). Compared to purely text-based hate speech, visualized hate may be processed more intuitively (Schmid, 2023), while at the same time allowing for plausible deniability due to its humorous framing (Bogerts & Fielitz, 2019; Quintal & Collen, 2023; Tuters, 2019). The perception and interpretation of the same meme may vary depending on the social context and prior orientations, which can lead to different affective reactions and, consequently, different behavioral responses to identical content, fostering polarization (Rossi et al., 2025).

This ambiguity may be especially relevant in group contexts, as it increases uncertainty about appropriate responses and intervention strategies. While much attention has been paid to hate speech in general and on public platforms (Tontodimamma et al., 2021), less is known about its presence and impact within private chat groups, where communication is more personal and shaped by stronger familiarity and social bonding (Cruz-Cárdenas et al., 2021; Zhao et al., 2021). Existing research on hate speech in general suggests that individuals often react passively and avoid confrontation on public social media platforms (Heger et al., 2022; Landesanstalt für Medien NRW, 2023). In private chat groups, however, reactions to problematic content may be particularly challenging. In contrast to many public online environments, hate content in private groups examined in this study is not shared by fake accounts, chatbots, or anonymous internet users, but by real people from one's immediate social circle. Therefore, individuals must respond not only to the content itself, but also to the social dynamics of the group in question. They must also take into account the real-world consequences, particularly with regard to their relationships with other group members. (Gagrčin, 2022).

In this study, we explore specifically through the lens of social conformity and social ties how group dynamics in private chat groups influence individuals' responses to hate memes. To this end, an experimental vignette design was used to systematically vary (1) relationship closeness (i.e., whether participants imagined being in a group of close friends or strangers), and (2) group behavior (i.e., whether others reacted to the hate meme or remained silent). In the present study, these two dimensions are conceptualized as central sources of group influence: relationship closeness reflects group strength (Latané, 1981) and majority behavior may serve as a descriptive norm (Cialdini et al., 1991; Ravis & Sheeran, 2003).

Understanding responses to hate memes in private chat groups is relevant for both research and practice. As digital communication increasingly takes place within socially embedded groups, reactions to problematic content are shaped not only by the content itself but also by interpersonal relationships and group norms. Examining these dynamics helps to clarify how social influence operates in digital environments and how individuals negotiate disagreement, belonging, and social norms within ongoing relationships.

Silence or Responding as Conforming Behavior

Although there are various ways to react to hate speech on social media (Bartlett & Krasodonski-Jones, 2015; Jakubowicz et al., 2017), this study focuses on the decision to either speak up in the chat or remain silent. Inaction does not necessarily mean that users ignore a post (Lutz & Hoffmann, 2017; Preece et al., 2004). One possible explanation for the decision to remain passive and observe is the behavior of others and the resulting social influence of the group. Accordingly, we examine the extent to which reactions to hate memes can be attributed to conformity processes in group chats. Research on social influence, norm formation and conformity in groups has a long history (Deutsch & Gerard, 1955; Sherif, 1936). However, while digitally mediated interaction differs

from offline communication in important ways, recent research suggests that conformity processes remain highly relevant in online environments, although often in attenuated form (Wijenayake & Goncalves, 2025).

While such research primarily focuses on opinion congruence, the present study examines behavioral adaptation to observable majority behavior, which may serve as a social cue for responding to hate memes within a specific chat group. According to the Focus Theory of Normative Conduct (Cialdini et al., 1990; Cialdini et al., 1991), individuals use salient social cues to infer appropriate behavior, making norms particularly influential when they are brought to the forefront of attention in a given situation. Descriptive norms are especially relevant in this regard because they inform individuals about how others typically behave in a given situation. For example, when most users in an online forum introduce themselves before posting, this visible pattern may signal that self-introduction is expected and guide newcomers to do the same. As a consequence, descriptive norms can emerge implicitly, and, as a meta-analysis shows (Rivis & Sheeran, 2003), can predict (planned) behavior. Because descriptive norms tend to be particularly influential in uncertain situations, they may also shape responses to hate memes, whose humorous elements can create ambiguity regarding their meaning and appropriate reactions (Schmid, 2023). Therefore, the absence or presence of comments on a hate meme may similarly act as a descriptive norm that subtly signals appropriate behavior, reinforcing either passive inaction or active engagement.

Importantly, the social cues that shape descriptive norms need not consist of overt behavior alone. Silence and inaction can likewise communicate information about what is considered appropriate in a given situation (Aarts & Dijksterhuis, 2003; Sorrels & Kelley, 1984). A similar logic is reflected in work on the bystander effect (for an overview, see Latané & Nida, 1981), particularly in the concept of pluralistic ignorance, where individuals infer appropriate behavior from others' non-responses and may mistakenly assume that no intervention is required, thereby reinforcing inaction (Latané & Darley, 1968; Rendsvig, 2014). While bystander research also discusses diffusion of responsibility which assumes that the mere presence of other bystanders reduces perceived personal responsibility to intervene (Leonhard et al., 2018), the present study focuses on the interpretation of others' visible reactions. The self-reinforcing effect of pluralistic ignorance can lead to the perception that intervening violates an implicitly mediated social norm. Online environments exacerbate this effect, as the "funhouse mirror" of social media can distort group norms, making individuals more prone to misinterpreting others' behaviors and beliefs (Robertson et al., 2024). Studies have confirmed the impact of pluralistic ignorance on inhibiting public interventions for marginalized groups (DeSouza & Schmader, 2022; Halbesleben, 2009; Prentice & Miller, 1996). Additionally, in chat group conversations, a bystander's silence or an interruption of the conversational flow may be interpreted as subtle approval of the expressed opinion (Koudenburg et al., 2013; Koudenburg et al., 2021). The transformation of observed majority behavior into a perceived norm helps explain why individuals, even if they personally disagree with hate speech, fail to intervene when the group remains silent.

Beyond silence, social norms can also shape active engagement in online communication. For instance, Wachs et al. (2022) demonstrated that descriptive hate speech norms (how often participants perceived hate speech in their environment) positively predicted motivations to engage in hate speech, while injunctive anti-hate norms (the extent to which friends and family disapprove of hate speech) reduced motivations. Similarly, Kreuder et al. (2024) experimentally showed that induced group norms in simulated WhatsApp chats directly shaped anticipated reactions: when aggressive peer responses were present, participants were almost twice as likely to conform with cyber-aggressive behavior compared to friendly or humorous norm conditions. Taken together, these findings underscore that group norms can substantially influence individuals' willingness to engage in (or refrain from) discriminatory online communication. Assuming that descriptive norms exert influence particularly when group behavior is consistent and uniform, conformity effects should be strongest when all group members react similarly, rather than when only some members comment or remain silent. In the following hypotheses, this general willingness to act is operationalized as behavioral intention, referring to the reported likelihood of performing a specific behavior in a given situation.

This leads to the first hypothesis:

H1a: The intention to comment on hate memes is higher in chat groups in which all other group members comment than in groups in which none or only half of the members comment.

Moderating Role of Relationship Closeness

In contrast to the theoretical approach, Leonhard et al. (2018) found that intervention from others does not significantly influence action in public settings. In fact, individuals do not necessarily conform to every group norm: According to Social Impact Theory (Latané, 1981; Latané & Wolf, 1981), social influence of a group may increase with the number, strength and immediacy of sources influencing the individual (Mullen, 1985). Since the number and thus the bystander effect have often been investigated in the context of online hate speech (OHS) in an anonymous setting (Fischer et al., 2011; Leonhard et al., 2018; Obermaier et al., 2016), we focus on the other two aspects, namely strength and immediacy, which become more relevant in the context of chat groups examined in this study (Fernández et al., 2025). The present study conceptualizes relationship closeness as a relational configuration reflecting aspects of both strength and immediacy within Social Impact Theory:

Basically, “strength” refers to the social significance of the source, which determines how impactful the influence is for the individual. Groups that are perceived as more important, cohesive, or personally relevant are therefore more likely to elicit conformity, as their reactions carry greater interpersonal consequences and increase the motivation to maintain group belonging and avoid social disapproval (Latané, 1981). Although Latané and Wolf (1981) define ‘strength’ in a relatively broad manner (typically including power, social status, but also attractiveness, entity and group cohesiveness), most studies primarily associate strength with status and prestige (Mullen, 1985). Nevertheless, we use a broader definition of ‘strength’, highlighting the distinction in personal importance between groups with high and low relationship closeness. Groups with low closeness may resemble anonymous settings, where individuals feel less bound by group norms and less influenced by peer expectations. Research on anonymity shows that when group members are not personally connected, individuals disclose more and conform less (Clark-Gordon et al., 2019; Huang & Li, 2016; Wu & Atkin, 2018; You & Lee, 2019).

Immediacy, e.g., due to temporal and local proximity of the group with which the individual communicates via chat, could have a similar effect, even when group members are not physically co-present (Crutchfield, 1955; Latané et al., 1995; Neubaum & Krämer, 2018). Instant messaging nevertheless creates a sense of presence and relationship closeness increases the possibility of offline encounters (Knop-Hülß, 2023), which may further strengthen social impact. Relationship closeness therefore encompasses both strength and immediacy, shaping the extent to which individuals are responsive to group behavior.

The following hypothesis can be formulated from this theoretical framework:

H1b: The conformity effect described in H1a is stronger in chat groups characterized by high rather than low relationship closeness.

The Direct Effect of Relationship Closeness on the Intention to Comment or Leave the Chat

Beyond increasing conformity to visible group reactions, relationship closeness may also generally increase individuals’ willingness to engage with the group communication itself. From the perspective of Social Impact Theory, groups that are perceived as socially important and personally consequential become more difficult to ignore, making both the group’s reactions and the interaction context itself more behaviorally relevant. If a group is perceived as an important part of one’s identity, members are more inclined to protect its integrity (Tajfel & Turner, 1979). This includes actively addressing problematic behavior such as the spread of hate memes.

Empirical findings support this notion in the digital context of hate content: Studies on cyberbullying (e.g., DeSmet et al., 2016) as well as studies on OHS (Pabian, 2025) indicate that having a personal connection with a victim or identifying with the attacked group significantly increases the likelihood of intervention. This highlights the role of perceived personal involvement in motivating prosocial action. However, little is known about how the relationship to the group of bystanders affects such responses. In the context of OHS, prior qualitative research has shown that individuals are more willing to react directly and openly when they are familiar with the group and its members (Josupeit, 2023; Kromer et al., 2021). On the basis of vignette-induced interviews, Gagrčin (2022) analyzed the fact that uncivil remarks made by friends led to greater irritation and consequently to an increased willingness to take responsibility and intervene. The interviewees reported an educational mission for themselves. Therefore, it can be assumed that familiarity fosters a greater sense of responsibility and may override tendencies toward passive strategies as “critical ignoring” like the “do-not-feed-the-troll-heuristic” (Kozyreva et al., 2022) or “deliberate ignorance” (Hertwig & Engel, 2016), which are more applicable in anonymous or impersonal settings.

Additionally, people tend to leave groups and situations that they deem to be unimportant. While individual mobility facilitates group exit, strong social ties often inhibit withdrawal (Ellemers et al., 1997). To date, no study has examined how variations in relationship closeness within chat groups affect user behavior in reaction to problematic content. This is particularly relevant for actions such as leaving a group, which in private chats can carry symbolic meaning and signal social disengagement. In contrast to public online spaces where an individual's leaving usually goes unnoticed, exiting a private group may itself constitute a form of social expression or subtle dissent.

In sum, relationship closeness does not only amplify conformity to perceived social norms but also increases the willingness to engage with problematic content rather than disengaging from the group. Accordingly, the hypotheses are as follows:

In chat groups characterized by high relationship closeness, the intention to comment on hate memes is higher (**H2a**) and the intention to leave the group is lower (**H2b**) than in chat groups characterized by low relationship closeness.

The preceding considerations primarily conceptualized these effects in terms of the social importance of the group and its members. However, interactions with hate memes occur within chat groups as specific communication environments. To account for this contextual dimension, the present study exploratorily examines whether perceived chat group importance mediates the relationship between relationship closeness and behavioral responses.

Methods

Sample

The sample size needed for the hypotheses tests was determined a priori with G*Power. Due to the online conditions and the hypothetical situation, we expected only a small to medium effect of $\eta^2 = 0.03$ (effect size $f = 0.176$). With a targeted power of .9 and an α -error-probability of .05, a total of 538 valid cases are needed to obtain a significant result with a two-factorial (2×3) ANOVA.

The survey ran via the non-representative SoSci Panel (Leiner, 2025), an established online access panel widely used in the German-speaking research community. This recruitment method was chosen because it provides rapid access to a large pool of participants, enabling sufficient statistical power to detect even subtle effects. While the panel tends to overrepresent individuals with higher educational attainment, it offers considerably more heterogeneity in age and professional background compared to typical student samples, which are still common in experimental social psychology (Roberts et al., 2020).

In October 2023, the questionnaire was sent to 4,000 people. A total of 1,177 people started the survey, and $N = 1,090$ participants completed it and provided explicit consent for the processing of their data following the debriefing. 60% ($n = 652$) women, 40% ($n = 429$) men and 1% ($n = 9$) diverse participants participated in the study. Their age varied between 24 and 87 years old ($M = 49.8$, $SD = 13.84$ years). The participants had an overall high level of education (63% with university degree, $n = 685$). A total of 25 participants (2%) reported discrimination on racial grounds, because of membership of an ethnic group or origin from another country. Accordingly, the sample can be considered WEIRD (Western, Educated, Industrialized, Rich, and Democratic; Henrich, 2021). Detailed sociodemographic information is provided in Table A1 in Appendix A.

Experimental Design

An online experiment was conducted using a 2×3 between-subjects design, manipulating relationship closeness (low, high) and group response (no reaction, full reaction, mixed reaction). The study was embedded in an online questionnaire using SoSci Survey (Leiner, 2024), and participants were randomly assigned to one of six conditions. Vignettes were used to simulate realistic social interactions. Vignettes are short descriptions of hypothetical scenarios that allow researchers to examine how attitudes, emotions, or behavioral intentions vary as a function of systematically manipulated situational characteristics (Aguinis & Bradley, 2014). In this approach, independent variables are varied within the scenario without requiring participants to actually experience the situation. Although vignette studies allow systematic manipulation under controlled conditions, compared to field

experiments, they are often considered to have lower ecological validity, especially when it comes to behavioral measurements (Eifler & Bentrup, 2003; Eifler & Petzold, 2019; Groß & Börensen, 2009; Nisic & Auspurg, 2009). However, as Fahmie et al. (2023) discussed, the relevance of ecological validity depends on the specific research objective. Although the present study addresses a phenomenon relevant to everyday communication, its primary objective is to examine fundamental social-psychological mechanisms under controlled conditions, focusing on internal validity and feasibility.

The vignette design explicitly allowed participants to indicate no action, which was intended to ensure that responses reflected genuine intentions rather than experimental pressure. Additionally, creating high relationship closeness in a real experimental setting requires significant time and effort, which could introduce confounding variables or unwanted group dynamics. By using vignettes, the study avoids these complications, allowing for a more controlled environment (Steiner et al., 2016). Moreover, the vignette-based design facilitated the recruitment of a large number of participants (Auspurg et al., 2009), increasing statistical power and allowing for the detection of subtle effects that might have gone unnoticed in smaller, more constrained group settings.

The study was reviewed and approved by the Ethics Committee of the Institute of Psychology at the University of Duisburg-Essen.

Procedure

At the beginning of the experiment, participants were informed that the study examined behavior in chat groups and that they would encounter potentially disturbing images. However, to avoid priming effects, no explicit reference to hate speech was made before exposure. A debriefing followed at the end of the study.

After informed consent, participants were introduced to their assigned chat group scenario (low or high relationship closeness). Additionally, they were asked about the importance of the chat group and prior experience and imaginability of the hypothetical situation.

In both scenarios, participants were told that they had been invited to a birthday party and to a chat group through which the celebration would be organized. To provide context for the following memes, a transitional sentence was added: "The chat group isn't just used to plan the party, but for other things as well. Today, for example, people are sending various memes back and forth in the birthday chat group."

Next, screenshots from the chat group were shown, depicting two neutral memes being sent, with the third meme being racist. Participants were then presented with the reactions (or lack thereof) of the other group members, depending on which experimental group they were in (see Figure 1 to 3).

After every meme and, in case of the racist one, after viewing the conversation, participants indicated how they would respond to the meme. Afterwards, participants evaluated the group's reactions, before having the opportunity to write their own message in the chat group.

Following the experimental part, they answered questions about their everyday use of chat groups and previous encounters with racist memes. To further assess participants' attitudes towards such content, they evaluated an additional racist and a neutral meme. Additionally, they provided information on age, gender, and previous experiences of discrimination.

Stimulus Material

The study examined individuals' responses to hate memes in chat groups using manipulated WhatsApp-style screenshots. Initially, group members shared harmless memes, followed by a hateful meme posted by one member. A pre-study was conducted to select suitable stimulus material for the main experiment, including the hateful meme and the two neutral memes. In this pre-study, participants evaluated and compared two sexist, anti-Semitic, racist, classist, and neutral memes regarding, among other aspects, the extent to which they conveyed an intentional devaluation of a social group (for more details, see Supplemental Material). The meme selected for the main study was one of the racist memes, depicting an orangutan mother holding two dark-skinned children, captioned "Mother's Day in Africa".

Experimental Manipulations

Relationship Closeness

A chat group for planning a birthday party was chosen as the common scenario. To manipulate the relational context of the group, several relational cues theoretically associated with increased social impact were varied simultaneously, including participants' familiarity with the group members, cues of social connectedness within the group, and the expected immediacy of future interaction.

The exact wording for the high relationship closeness (HRC) condition is: "You've been invited to one of your best friends' birthday parties. To share details about the time and location of the celebration, a chat group was created, and you're now part of it along with five other people. You know almost everyone in the group, and it quickly becomes clear to you that the other group members know each other very well, too. Since you're free on the evening of the birthday party, you'll see the others again in just a few days."

In the low relationship closeness (LRC) condition, the scenario was adapted such that participants had been invited by a new colleague they had not yet met, did not know the other group members, and were unable to attend the party, thereby reducing both relational closeness and expected future interaction.

Group Response

To manipulate group response, participants saw different reactions to the racist meme: No one commented (no reaction), everyone commented (full reaction), half of the group commented (mixed reaction). In order to eliminate the potential influence of emotionality and tonality on the reactions (Josupeit, 2023) and opinion congruence (Neubaum, 2016; Neubaum & Krämer, 2017), care was taken to use neutral wording and to avoid insults when simulating the comments. Figure 1 shows the manipulated screenshot for the no reaction condition, Figure 2 for the mixed reactions (two comments) and Figure 3 for the full reaction (four comments).

Participants were instructed to imagine checking their phone one day after the meme was posted, ensuring they perceived the scenario as complete.

Figure 1. *Manipulated Screenshot With No Group Reaction.*



Note. Translation of information and messages: 1 unread message – Yesterday – meme.

Figure 2. *Manipulated Screenshot in Which Everyone Commented.*



Note. Translation of information and messages: 5 unread messages – Yesterday – „oh, that's pretty harsh, a bit off the mark..." – "Well, as long as it stays in the circle. I don't think it's so bad. It's certainly not meant to be nasty, especially here." – "Just dark humor..." – Today – "To be honest, I think that picture is pretty violent. It's not my kind of humor, but everyone is different."

Figure 3. *Manipulated Screenshot With Mixed Group Reaction.*



Note. Translation of information and messages: 3 unread messages – Yesterday – meme – „oh, that's pretty harsh, a bit off the mark..." – "Well, as long as it stays in the circle. I don't think it's so bad. It's certainly not meant to be nasty, especially here."

Measures

Behavioral Intention

The dependent variable in this study is the participant's behavioral intention in the given scenario, specifically their likelihood of commenting or leaving the group. Since the range of possible responses is broad, even in chat groups, and these responses are not mutually exclusive, participants were asked to indicate their likelihood of engaging in each type of response. Similar to the approach recently used by Meerson et al. (2025), multiple response options were assessed separately. However, rather than combining them into broader intervention categories, the present study analyzed the individual response tendencies directly at the single item level. The response options included: (a) openly commenting in the chat group, (b) reacting with an emoji, (c) posting another image, (d) sending a voice message, (e) not reacting at all, and (f) leaving the group. The inclusion of options (b) through (e) served to illustrate the range of possible actions within a chat environment without being central to the analysis.

Given the hypothetical nature of the situation and the complexity of real-life decision-making, the focus was placed on participants' self-reported likelihood of engaging in certain behaviors rather than forcing a binary choice. To

take this into account, we followed the approach outlined in the study by Leonhard et al. (2018) and asked participants to rate the likelihood of each response option on a 7-point scale (1 = *in no case*, 7 = *in any case*). Descriptive statistics for the central study variables are reported in the Results section.

Chat Group Importance

It should be verified whether the scenarios effectively manipulate perceived chat group importance, serving as a manipulation check to confirm that chat groups with higher relationship closeness are indeed rated as more important. Beyond this, perceived chat group importance may be central as a potential mediator in the relationship between relationship closeness and behavioral responses. To this end, participants rated the personal significance of the group using five items ($\alpha = .85$), for example, "This chat group would be very important to me." A list of items is provided in Table A2 in Appendix A. All items were rated on a 5-point scale (1 = *totally disagree*, 5 = *totally agree*). A mean score was computed for further analysis ($M = 2.40$, $SD = 0.91$; for factor analysis, see Supplemental Material).

Prior Experience and Imaginability of the Hypothetical Situation

To assess whether some scenarios were generally more relatable than others, and to help participants relate to the birthday party chat group scenario, participants indicated to what extent they could recall a similar situation ("I can think of at least one situation that was similar."; $M = 3.38$, $SD = 1.53$) and how well they could imagine the scenario presented ("I can easily imagine the situation."; $M = 4.42$, $SD = 0.88$). Both single items were rated on a 5-point scale (1 = *totally disagree*, 5 = *totally agree*) and were used for manipulation check.

Additional items assessed participants' perceptions of the discussion, their prior experiences with chat groups and an evaluation of a comparable racist meme. These variables were included to capture potentially relevant contextual influences beyond the core theoretical model. As the present study follows a theory-driven approach focusing on specific hypothesized mechanisms, these variables were not included in the main analyses. For transparency, their measures and associations with the main variables are reported in the Supplementary Material.

Data Analysis

The data were analyzed using SPSS 29. The hypotheses were tested using two-way ANOVA and t-test for independent samples. For mediation analysis, we used the PROCESS macro Version 4.3 by Hayes (2023) for SPSS.

Results

Manipulation Checks

Participants in HRC-condition ($M = 2.81$, $SD = .74$) reported significantly higher levels of perceived chat group importance than those in LRC-condition ($M = 1.98$, $SD = .74$), $t(1064.039) = -16.88$, $p_{\text{two-tailed}} < .001$, $d = -1.02$. This is an indication that the scenario described has the desired effect on the perceived importance of the chat group, but that nevertheless the significance is relatively low in both groups.

In addition, participants in the HRC-condition ($M = 3.96$, $SD = 1.315$) reported higher levels of prior experience than those in LRC-condition ($M = 2.79$, $SD = 1.50$), $t(1065.44) = -13.72$, $p_{\text{two-tailed}} < .001$, $d = -0.83$. Furthermore, imaginability was higher in HRC ($M = 4.65$, $SD = .70$) than in LRC condition ($M = 4.19$, $SD = .99$), $t(970.176) = -8.71$, $p_{\text{two-tailed}} < .001$, $d = -0.53$. These experimental differences must be taken into account when interpreting the data.

To ensure that the observed reactions were specific to the hate meme rather than reflecting a general response tendency, behavioral intentions were descriptively compared across the two preceding neutral memes and the critical hate meme. Mean intention ratings and standard errors were plotted in a trend chart (see Supplemental Material). The graphical comparison suggested noticeable shifts in behavioral intentions for the hate meme.

Preliminary Statistical Analyses

Across all experimental conditions, participants reported the highest behavioral intentions for leaving the group ($M = 4.55$, $SD = 2.19$) and openly commenting in the chat group ($M = 4.02$, $SD = 2.51$), followed by not reacting at all ($M = 3.48$, $SD = 2.41$). The other reactions such as reacting with an emoji ($M = 1.90$, $SD = 1.71$), posting another image ($M = 1.72$, $SD = 0.84$), or sending a voice message ($M = 1.62$, $SD = 1.51$) received comparatively low endorsement.

Despite the violation of the normal distribution assumption in all six groups for the intention to comment and to leave (results of separated Shapiro-Wilk-Tests, $p < .001$ in all experimental groups, for details and histograms, see Supplemental Material), even for their residuals, a parametric procedure was carried out due to the sample size being large enough and the robustness of ANOVA (Blanca et al., 2017; Glass et al., 1972). A boxplot analysis of the dependent variables intention to comment and intention to leave, separated by experimental group, did not identify any outliers.

Additional analysis of non-experimental factors influencing behavioral intentions are presented in the Supplemental Material.

Analysis of the Intention to Comment

A two-way ANOVA was conducted to examine the effects of relationship closeness and group reaction on the intention to comment, including their interaction to assess whether relationship closeness moderates this effect.

The means and standard deviations for the intention to comment are presented in Table 1.

Table 1. Means and Standard Deviations for the Intention to Comment (Relationship Closeness by Group's Reaction).

Group's Reaction	Low relationship closeness		High relationship closeness	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
No reaction	3.46	2.44	4.44	2.53
Full reaction	3.62	2.47	4.63	2.39
Mixed reaction	3.47	2.53	4.45	2.41

Note. 7-point-scale from *in no case* (1) to *in any case* (7).

The results indicated no significant main effect for group's reaction, $F(2, 1084) = 0.568$, $p = .567$, and no significant interaction effect between relationship closeness and group's reaction, $F(2, 1084) = 0.005$, $p = .995$. Thus, hypotheses H1a and H1b are rejected. In line with hypothesis H2a, the statistical test reveals a significant main effect for relationship closeness, $F(1, 1084) = 44.004$, $p \leq .001$, partial $\eta^2 = .039$. In total, the model is significant, $F(5, 1084) = 9.057$, adjusted $R^2 = .036$, $n = 1,090$. Since the only differences found are between two groups, no post hoc test is required.

Following the confirmation of the main effect of relationship closeness on behavioral intentions, a mediation analysis by Hayes (2023) was conducted to examine whether this effect was statistically explained by the perceived importance of the group. This method was chosen because it allows for bootstrapping, which is particularly advantageous given the non-normal distribution of the data. As a result, a direct effect of relationship closeness on the intention to comment was found ($B = .991$, $\beta = .395$, $p < .001$). By entering the group's importance as mediator, relationship closeness predicted the group's importance ($B = .827$, $\beta = .910$, $p < .001$, 95% CI [.733, .926]¹) and the group's importance predicted the intention to comment ($B = .312$, $\beta = .113$, $p = .001$, 95% CI [.131, .498]). Nevertheless, the relationship between relationship closeness and intention to comment was still significant ($B' = .733$, $\beta' = .293$, $p < .001$, 95% CI [.395, 1.058]), so that we have to assume that the effect was only partially mediated by the group's importance. Overall, there was a partially standardized significant indirect effect of .103, 95% CI [.043, .167].²

Analysis of the Intention to Leave the Group

The means and standard deviations for the intention to leave the chat group are presented in Table 2.

The one-tailed t -test³ for independent samples shows that the intention to leave the chat group is higher in the LRC-condition than in the HRC-condition, $t(1076.37) = 11.65, p < .001, 95\% \text{ CI } [1.21, 1.70], d = 0.71$. Thus, hypothesis H2b is confirmed.

We also tested whether the personal importance of the group mediates the effect of relationship closeness on the intention to leave. As a result, a total effect of relationship closeness on the intention to leave was found ($B = -1.454, \beta = -.665, p < .001$). By entering the group's importance as mediator, the group's importance predicted the intention to leave ($B = -0.622, \beta = -.259, p < .001, 95\% \text{ CI } [-.769, -.468]$). Again, the relationship between relationship closeness and intention to leave was still significant ($B' = -.940, \beta' = -.430, p < .001, 95\% \text{ CI } [-1.222, -.657]$), so that the effect is only partially mediated by the perceived importance. Overall, there was a partially standardized significant indirect effect of $-.235, 95\% \text{ CI } [-.300, -.173]$.

Table 2. Means and Standard Deviations for the Intention to Leave (Relationship Closeness by Group's Reaction).

Group's Reaction	Low relationship closeness		High relationship closeness	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
No reaction	5.17	1.98	4.00	2.19
Full reaction	5.40	1.93	3.76	2.15
Mixed reaction	5.29	1.90	3.74	2.20

Note. 7-point-scale from *in no case* (1) to *in any case* (7).

To examine whether the observed effects were robust across relevant person-level variables, additional ANCOVAs for both dependent variables were conducted controlling separately for age, gender, prior discrimination experiences, and participants' evaluation of a comparable racist meme. Across all models, the overall pattern of results remained unchanged, with relationship closeness remaining a significant predictor of behavioral intentions. Detailed results are reported in the Supplementary Material.

Discussion

The present study examined under which group-dynamic conditions individuals respond to hate memes in private chat groups by either commenting or leaving the group. Using a 2×3 vignette-based experimental design, the study tested the role of descriptive norms (Cialdini et al., 1991, operationalized via group behavior) and relationship closeness as a source of social influence (Latané, 1981). While no evidence was found for conformity to descriptive norms or their interaction with relationship closeness, the findings show that individuals were more likely to respond and less likely to disengage when the chat group consisted of close others. Both effects were partially mediated by the perceived importance of the chat group.

Relationship Closeness Affects the Intention to Comment and to Leave

The study shows that relationship closeness to the group increases the intention to comment on hate memes and reduces the likelihood of leaving the chat group. This is consistent with qualitative findings of Josupeit (2023), Kromer et al. (2021) and Gagrčin (2022) about the role of social bonds on the willingness to intervene, and it supports the social impact perspective that strong ties foster engagement and in-group responsibility. In addition, the tendency to leave less close groups and engage less with their content shows that critical ignoring (Hertwig & Engel, 2016; Kozyreva et al., 2022) may be more likely in less socially embedded communication settings.

The mediation analysis further revealed that the perceived importance of the chat group partially mediated the effect of relationship closeness on behavioral intentions. This suggests that close social ties may increase engagement not only directly, but also indirectly by increasing the perceived importance of the communication setting. This mechanism may partly account for the silent majority observed in public online spaces, where participation is low precisely because the setting holds little personal meaning. Importantly, the findings highlight that relationship closeness may foster conditions under which shared understandings of acceptable behavior can be negotiated, even in digital contexts where avoidance is the easier option.

No Conformity Effects Were Found

However, the study did not provide evidence that responses to hate memes are driven by group-conforming behavior when social norms are only indirectly conveyed through uniform group behavior. This finding challenges

assumptions that visible group behavior alone guides individual responses in private chat groups. Studies demonstrating the effect of pluralistic ignorance in hate-speech related fields (DeSouza & Schmader, 2022; Halbesleben, 2009; Prentice & Miller, 1996) do not explicitly refer to online contexts. It may be due to the online context that the uniform group behavior did not become sufficiently salient as a descriptive norm and therefore had no impact (Robertson et al., 2024). At the same time, the visible reactions of others may also have reduced participants' perceived need to intervene themselves (Josupeit, 2023). Different social influence processes may therefore have operated simultaneously and partially offset each other.

Contrary to Social Impact Theory (Latané, 1981), high relationship closeness did not increase conformity. As described above, it may be that the respondents simply did not perceive the behavioral norm intended in the experiment. Further, relationship closeness may represent a more complex form of social influence than assumed in the present application of Social Impact Theory. Unlike more formal sources of influence, friendships involve shared experiences and pre-existing group norms that may shape responses independently of visible behavioral cues. It is therefore possible that other, unobserved group norms—such as injunctive expectations within politically engaged or anti-racist friend groups—shaped participants' reactions in ways that the design could not capture.

Therefore, the absence of conformity effects should not be interpreted as a lack of normative influence, but rather as an indication that in digital communication spaces, it seems to matter little whether the group responds or not (in line with Leonhard et al., 2018), suggesting that visible group behavior may not function as a primary reference point for individual decision-making in such contexts. Instead, the findings point to the importance of relational embeddedness: Individuals' responses appear to be shaped less by observable group behavior and more by the social relevance of the group.

Broader Implications

The group-dynamic mechanisms examined here have broader implications for digitally mediated communication: In contexts where social ties are weak or absent, individuals may be less inclined to respond, even when they disagree with the content. As a result, a small number of visible reactions may disproportionately shape perceived norms, potentially creating an impression of consensus or lack of opposition. This, in turn, may contribute to situations in which problematic content remains unchallenged or where those who engage in counter speech appear isolated.

At the same time, many online environments are structured in ways that bring together like-minded, but not relationally close individuals, for example in semi-public groups or algorithmically curated spaces. In such contexts, reduced motivation to intervene, combined with a greater tendency to disengage from unwanted interactions (e.g., by leaving groups or avoiding discussions), may further reinforce homogeneous viewpoints and contribute to the amplification of a small number of highly visible voices.

Beyond concerns about silence and polarization, the findings also point to the potential of socially embedded chat groups as spaces for addressing and discussing problematic content. Compared to interactions among strangers, ongoing relationships may facilitate engagement and disagreement without immediately ending the interaction. While the specific direction of such discussions depends on the norms and attitudes prevailing within a group, the findings suggest that close social ties do not necessarily inhibit responses to problematic content. Rather, they may provide a social context in which disagreement can be expressed and negotiated.

The present study, along with the questions it leaves open, highlights the need for further research on how descriptive, injunctive, and personal norms interact in private digital contexts, and on which additional factors may strengthen conforming behavior.

Limitations

To examine the causal influence of conformity-related factors on responses to hate memes in private, actually unobservable settings, an experimental vignette-design was chosen. This approach allows for large samples and controlled manipulation (Steiner et al., 2016). To increase visual and contextual realism (Aguinis & Bradley, 2014) authentic WhatsApp screenshots were used and participants were encouraged to complete the study on mobile devices. Nevertheless, vignette studies cannot fully capture real interactional pressures and actual behavior, which limits ecological validity (Eifler & Bentrup, 2003; Groß & Börensen, 2009; Nisic & Auspurg, 2009). Even if actual

reactions in real group interactions may differ, the findings remain informative with regard to the effects of the investigated factors, as the theory of planned behavior (Ajzen & Madden, 1986) and prior vignette research (Groß & Börensen, 2009; Nisic & Auspurg, 2009) suggest that factors that influence behavioral intention also influence actual behavior.

Another limitation of this study is that perceived behavioral, potentially injunctive norms were not directly assessed. Consequently, it is unclear which behaviors participants considered normative, and the Theory of Normative Conduct (Cialdini et al., 1990; Cialdini et al., 1991), especially the salience of norms, has not been fully tested. Future research should include explicit norm assessments to clarify their role alongside relationship closeness and group's behavior.

A key methodological challenge was isolating the descriptive norm to comment from group members' opinions on the meme. Since agreement or disagreement with the meme or with the discussion could also influence responses (Neubaum & Krämer, 2018), it remains unclear whether participants conformed to the behavioral pattern or react to the perceived stance of the group. Indeed, according to the explorative analysis, a significant negative correlation for the evaluation of others' reactions as acceptable and the intention to comment was found (see Supplemental Material). Further, the simulated discussion was intended to be neutral, but it could also be interpreted as an attempt at justification. Although studies have shown that such justifications do not influence participants' behavior (Pabian, 2025), future research could address this issue by designing group reactions that clearly signal a behavioral norm without expressing an explicit opinion. Conducting a preliminary study could help ensure that these responses are consistently perceived as neutral.

In addition, some limitations relate to the operationalization and measurement of participants' responses. This study examined decision-making in asynchronous communication, allowing for interruptions and no time pressure. Participants could express action intentions in gradations, accommodating uncertainty but complicating statistical analysis due to non-normal distributions (e.g., a U-shaped 'Intention to Comment' variable). This suggests greater decision certainty than expected. Future studies could require binary choices.

Furthermore, the measure of group importance requires conceptual and psychometric refinement. A validated instrument, such as the Multidimensional and Multicomponent Measure of Social Identification (Roth & Mazziotta, 2015), could strengthen future operationalizations and better capture the affective and evaluative dimensions of digital group significance.

Finally, the present study focused primarily on situational group dynamics and did not systematically assess broader person-related characteristics that may shape responses to hate memes, such as prior intervention behavior, ideological orientations, prejudice-related attitudes, or habitual exposure to online hate speech. Future research should investigate how such individual predispositions interact with group-level dynamics in private digital communication settings.

Constraints on Generality (COG) Statement (Simons et al., 2017)

The findings reported here are based on a WEIRD sample of German-speaking adults with relatively high educational attainment, based on SoSci Panel (Leiner, 2025). Thus, the results are most likely to generalize to populations with similar socio-demographic and cultural contexts. Given that cultural and identity-related experiences shape reactions to hate speech (Ping et al., 2025), replications in more diverse samples are needed to assess the generality of these findings.

Conclusion

The present study highlights the importance of relational context in shaping responses to hate memes in private chat groups. While the experimentally induced descriptive norms conveyed through others' behavior did not influence participants' reactions, relationship closeness emerged as a central factor, increasing engagement and reducing disengagement. These findings suggest that observable group behavior alone may be insufficient to explain responses in digitally mediated communication. Instead, individuals appear to orient their behavior toward the social relevance of the group and the potential relational consequences of their actions.

Although the findings should be interpreted in light of the non-representative sample, the experimental design and its inherent limitations, they point to broader implications for understanding digital communication dynamics:

In contexts where social ties are weak or absent, individuals may be less likely to intervene, even when they disagree with problematic content. As a result, visible reactions may not accurately reflect underlying attitudes, potentially contributing to skewed perceptions of social norms. By emphasizing the role of relational embeddedness, this study contributes to a more nuanced understanding of how social influence operates in private digital environments.

Footnotes

¹ The reported confidence interval refers to bootstrap results for the unstandardized regression model parameters.

² All parameters mentioned in the text are presented in tables in Supplemental Material.

³ Since variance heterogeneity was found ($F = 20.75, p < .001$), results of the Welch-test are presented.

Conflict of Interest

The author does not have any conflicts of interest to report.

Use of AI Services

The author declares she has used AI services, specifically ChatGPT and DeepL Write, to translate parts of the manuscript from German into English and to revise the language. She carefully reviewed the outputs to ensure correctness.

Data Availability Statement

The dataset generated during the current study is not publicly available because participants did not consent to the sharing of their data beyond the present research project.

Author's Contribution

The author of this manuscript was solely responsible for all aspects of the research and manuscript preparation. This includes the conception, literature review and investigation, methodology development, project management, provision of resources, data analysis, visualization of results, and writing of the original draft and its revision and editing. No other individuals contributed to any part of the study.

Acknowledgment

This research was supported by the Institute of Psychology at the University of Duisburg-Essen. The study was approved by the Ethics Committee of the Institute of Psychology, Faculty of Educational Sciences, University of Duisburg-Essen.

References

- Aarts, H., & Dijksterhuis, A. (2003). The silence of the library: Environment, situational norm, and social behavior. *Journal of Personality and Social Psychology*, 84(1), 18–28. <https://doi.org/10.1037/0022-3514.84.1.18>
- Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, 17(4), 351–371. <https://doi.org/10.1177/1094428114547952>
- Ajzen, I., & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology*, 22(5), 453–474. [https://doi.org/10.1016/0022-1031\(86\)90045-4](https://doi.org/10.1016/0022-1031(86)90045-4)

- Auspurg, K., Hinz, T., & Liebig, S. (2009). Komplexität von Vignetten, Lerneffekte und Plausibilität im faktoriellen Survey [Complexity of vignettes, learning effects, and plausibility in a factorial survey]. *Methoden, Daten, Analysen (mda)* [Methods, data, analyses (mda)], 3(1), 59–96. <https://nbn-resolving.org/urn:nbn:de:0168-ssaoar-126659>
- Bartlett, J., & Krasodonski-Jones, A. (2015). *Counter-speech: Examining content that challenges extremism online*. Demos. <http://www.demos.co.uk/wp-content/uploads/2015/10/Counter-speech.pdf>
- Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, 29(4), 552–557. <https://doi.org/10.7334/psicothema2016.383>
- Bogerts, L., & Fielitz, M. (2019). "Do you want meme war?" Understanding the visual memes of the German far right. In M. Fielitz & N. Thurston (Eds.), *Post-digital cultures of the far right: Online actions and offline consequences in Europe and the US* (pp. 137–154). <https://doi.org/10.1515/9783839446706-010>
- Carpenter, J., Green, M., & Laflam, J. (2018). Just between us: Exclusive communications in online social networks. *The Journal of Social Psychology*, 158(4), 405–420. <https://doi.org/10.1080/00224545.2018.1431603>
- Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 201–234). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60330-5](https://doi.org/10.1016/S0065-2601(08)60330-5)
- Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58(6), 1015–1026. <https://doi.org/10.1037/0022-3514.58.6.1015>
- Clark-Gordon, C. V., Bowman, N. D., Goodboy, A. K., & Wright, A. (2019). Anonymity and online self-disclosure: A meta-analysis. *Communication Reports*, 32(2), 98–111. <https://doi.org/10.1080/08934215.2019.1607516>
- Crutchfield, R. S. (1955). Conformity and character. *American Psychologist*, 10(5), 191–198. <https://doi.org/10.1037/h0040237>
- Cruz-Cárdenas, J., Guadalupe-Lanas, J., Ramos-Galarza, C., & Palacio-Fierro, A. (2021). Factors influencing the consumer tendency to use mobile instant messaging (MIM): A mixed method study. *Contemporary Economics*, 16(1), 32–50. <https://doi.org/10.5709/ce.1897-9254.467>
- DataReportal. (2025). *Digital 2025: Global overview report*. <https://datareportal.com/reports/digital-2025-global-overview-report>
- DeSmet, A., Bastiaensens, S., van Cleemput, K., Poels, K., Vandebosch, H., Cardon, G., & de Bourdeaudhuij, I. (2016). Deciding whether to look after them, to like it, or leave it: A multidimensional analysis of predictors of positive and negative bystander behavior in cyberbullying among adolescents. *Computers in Human Behavior*, 57, 398–415. <https://doi.org/10.1016/j.chb.2015.12.051>
- DeSouza, L., & Schmader, T. (2022). The misjudgment of men: Does pluralistic ignorance inhibit allyship? *Journal of Personality and Social Psychology*, 122(2), 265–285. <https://doi.org/10.1037/pspi0000362>
- Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgement. *Journal of Abnormal Psychology*, 51(3), 629–636. <https://doi.org/10.1037/h0046408>
- Eifler, S., & Bentrup, C. (2003). *Zur Validität von Selbstberichten abweichenden und hilfreichen Verhaltens mit der Vignettenanalyse* [On the validity of self-reports of deviant and helpful behavior using vignette analysis] (Bielefelder Arbeiten zur Sozialpsychologie No. 208). Universität Bielefeld, Fak. für Soziologie. <https://pub.uni-bielefeld.de/record/1862507>
- Eifler, S., & Petzold, K. (2019). Validity aspects of vignette experiments: Expected "what-if" differences between reports of behavioral intentions and actual behavior. In P. J. Lavrakas, M. W. Traugott, C. Kennedy, A. L. Holbrook, E. D. de Leeuw, & B. T. West (Eds.), *Experimental methods in survey research: Techniques that combine random sampling with random assignment* (pp. 393–416). Wiley. <https://doi.org/10.1002/9781119083771.ch20>
- Ellemers, N., Spears, R., & Doosje, B. (1997). Sticking together or falling apart: In-group identification as a psychological determinant of group commitment versus individual mobility. *Journal of Personality and Social Psychology*, 72(3), 617–626. <https://doi.org/10.1037/0022-3514.72.3.617>

- Fahmie, T. A., Rodriguez, N. M., Luczynski, K. C., Rahaman, J. A., Charles, B. M., & Zangrillo, A. N. (2023). Toward an explicit technology of ecological validity. *Journal of Applied Behavior Analysis*, 56(2), 302–322. <https://doi.org/10.1002/jaba.972>
- Fernández, A., Elmer, T., Sádaba, C., García-Manglano, J., & Vanden Abeele, M. (2025). The quality of face-to-face and digitally mediated social interactions: Two experience sampling studies exploring the moderating role of physical location, interaction partner familiarity, and interaction purpose. *Journal of Computer-Mediated Communication*, 30(2), Article zmaf004. <https://doi.org/10.1093/jcmc/zmaf004>
- Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., Heene, M., Wicher, M., & Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin*, 137(4), 517–537. <https://doi.org/10.1037/a0023304>
- Gagrčin, E. (2022). Your social ties, your personal public sphere, your responsibility: How users construe a sense of personal responsibility for intervention against uncivil comments on Facebook. *New Media & Society*, 26(8), 4299–4316. <https://doi.org/10.1177/14614448221117499>
- Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research*, 42(3), 237–288. <https://doi.org/10.3102/00346543042003237>
- Groß, J., & Börensen, C. (2009). Wie valide sind Verhaltensmessungen mittels Vignetten? Ein methodischer Vergleich von faktoriellem Survey und Verhaltensbeobachtung. [How valid are behavioral measurements using vignettes? A methodological comparison of factorial surveys and behavioral observation] In P. Kriwy & C. Gross (Eds.), *Klein aber fein! Quantitative empirische Sozialforschung mit kleinen Fallzahlen [Small but mighty! Quantitative empirical social research with small sample sizes]* (pp. 149–178). VS Verl. für Sozialwiss; GWV Fachverlage. https://doi.org/10.1007/978-3-531-91380-3_7
- Halbesleben, J. R. B. (2009). The role of pluralistic ignorance in the reporting of sexual harassment. *Basic and Applied Social Psychology*, 31(3), 210–217. <https://doi.org/10.1080/01973530903058284>
- Harrison, M. A., & Gilmore, A. L. (2012). U txt WHEN? College students' social contexts of text messaging. *The Social Science Journal*, 49(4), 513–518. <https://doi.org/10.1016/j.soscij.2012.05.003>
- Hayes, A. F. (2023). *PROCESS Macro for SPSS* (Version 4.3) [Computer software]. <https://processmacro.org>
- Heger, K., Leißner, L., Emmer, M., & Strippel, C. (2022). *Weizenbaum Report 2022*. <https://doi.org/10.34669/wi.wr/3>
- Henrich, J. P. (2021). *The WEIRD people in the world: how the West became psychologically peculiar and particularly prosperous*. Farrar, Straus and Giroux.
- Hertwig, R., & Engel, C. (2016). Homo ignorans: deliberately choosing not to know. *Perspectives on Psychological Science*, 11(3), 359–372. <https://doi.org/10.1177/1745691616635594>
- Huang, G., & Li, K. (2016). The effect of anonymity on conformity to group norms in online contexts: A meta-analysis. *International Journal of Communication*, 10, 398–415. <https://ijoc.org/index.php/ijoc/article/view/4037>
- Jakubowicz, A., Dunn, K., Mason, G., Paradies, Y., Bliuc, A.-M., Bahfen, N., Oboler, A., Atie, R., & Connelly, K. (2017). *Cyber racism and community resilience: Strategies for combating online race hate*. Palgrave Macmillan Cham. <https://doi.org/10.1007/978-3-319-64388-5>
- Jenkins, H., Ford, S., & Green, J. (2018). *Spreadable media: Creating value and meaning in a networked culture*. New York University Press. <https://www.jstor.org/stable/j.ctt9qfk6w>
- Josuweit, C. (2023). Was beeinflusst unseren Umgang mit online hate speech? [What influences the way we deal with online hate speech?]. *MedienPädagogik: Zeitschrift Für Theorie Und Praxis Der Medienbildung*, 19, 87–114. <https://doi.org/10.21240/mpaed/jb19/2023.03.04.X>
- Knop-Hülß, K. (2023). *The permanently connected group (PeCoG)*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-43238-6>
- Koudenburg, N., Kannegieter, A., Postmes, T., & Kashima, Y. (2021). The subtle spreading of sexist norms. *Group Processes & Intergroup Relations*, 24(8), 1467–1485. <https://doi.org/10.1177/1368430220961838>
- Koudenburg, N., Postmes, T., & Gordijn, E. H. (2013). Resounding silences: Subtle norm regulation in everyday interactions. *Social Psychology Quarterly*, 76(3), 224–241. <https://doi.org/10.1177/0190272513496794>

- Kozyreva, A., Wineburg, S., Lewandowsky, S., & Hertwig, R. (2022). Critical ignoring as a core competence for digital citizens. *Current Directions in Psychological Science*, 32(1), 81–88.
<https://doi.org/10.1177/09637214221121570>
- Kreuder, A., Frick, U., Klütsch, J., Haehn, L., & Schlittmeier, S. J. (2024). The effect of aggressive group norms on young adults' conformity behavior in WhatsApp chats: A vignette-based experiment. *Scientific Reports*, 14(1), Article 17231. <https://doi.org/10.1038/s41598-024-67915-9>
- Kromer, I., Atzmüller, C., & Zartler, U. (2021). Wegschauen oder helfen? Zivilcouragiertes Handeln unter Jugendlichen im Internet [Look the other way or help? Acts of civil courage among young people on the internet]. In F. Berger (Ed.), *Schriftenreihe der ÖFEB-Sektion Sozialpädagogik: Vol. 6. Jugend – Lebenswelt – Bildung: Perspektiven für die Jugendforschung in Österreich [ÖFEB Section on Social Pedagogy Publication Series: Vol. 6. Youth – Social Environment – Education: Perspectives for youth research in Austria]*. Verlag Barbara Budrich.
<https://doi.org/10.2307/j.ctv23r3gbf.12>
- Landesanstalt für Medien NRW. (2023). *Hate Speech Forsa-Studie 2023: Zentrale Untersuchungsergebnisse [Forsa study on hate speech 2023: Key findings]*.
- Latané, B. (1981). The psychology of social impact. *American Psychologist*, 36(4), 343–356.
<https://doi.org/10.1037/0003-066X.36.4.343>
- Latané, B., & Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology*, 10(3), 215–221. <https://doi.org/10.1037/h0026570>
- Latané, B., Liu, J. H., Nowak, A., Bonevento, M., & Zheng, L. (1995). Distance matters: Physical space and social impact. *Personality and Social Psychology Bulletin*, 21(8), 795–805. <https://doi.org/10.1177/0146167295218002>
- Latané, B., & Nida, S. (1981). Ten years of research on group size and helping. *Psychological Bulletin*, 89(2), 308–324. <https://doi.org/10.1037/0033-2909.89.2.308>
- Latané, B., & Wolf, S. (1981). The social impact of majorities and minorities. *Psychological Review*, 88(5), 438–453.
<https://doi.org/10.1037/0033-295X.88.5.438>
- Leiner, D. (2024). *SoSci Survey* (Version 3.5.02) [Computer software]. <https://www.soscisurvey.de>
- Leiner, D. (2025). *SoSci Panel*. Ludwig-Maximilians-Universität München. <https://www.soscipanel.de/index.php>
- Leiser, A. (2022). Psychological perspectives on participatory culture: Core motives for the use of political internet memes. *Journal of Social and Political Psychology*, 10(1), 236–252. <https://doi.org/10.5964/jspp.6377>
- Leonhard, L., Rueß, C., Obermaier, M., & Reinemann, C. (2018). Perceiving threat and feeling responsible. How severity of hate speech, number of bystanders, and prior reactions of others affect bystanders' intention to counterargue against hate speech on Facebook. *Studies in Communication & Media*, 7(4), 555–579.
<https://doi.org/10.5771/2192-4007-2018-4-555>
- Lutz, C., & Hoffmann, C. P. (2017). The dark side of online participation: exploring non-, passive and negative participation. *Information Communication & Society*, 20(6), 876–897.
<https://doi.org/10.1080/1369118x.2017.1293129>
- Medienpädagogischer Forschungsverbund Südwest. (2022). *JIMplus 2022: Fake news und hate speech*.
<https://www.lfk.de/fileadmin/PDFs/Publikationen/Studien/JIMplus-2022/JIMplus-2022.pdf>
- Meerson, R., Koban, K., & Matthes, J. (2025). Too much of what? Two-wave panel evidence for selective (de-)sensitization through frequent exposure to different kinds of digital hate. *Journal of Computer-Mediated Communication*, 30(2), Article zmaf002. <https://doi.org/10.1093/jcmc/zmaf002>
- Mullen, B. (1985). Strength and immediacy of sources: A meta-analytic evaluation of the forgotten elements of social impact theory. *Journal of Personality and Social Psychology*, 48(6), 1458–1466. <https://doi.org/10.1037/0022-3514.48.6.1458>
- Neubaum, G. (2016). *Monitoring and expressing opinions on social networking sites: Empirical investigations based on the spiral of silence theory* [Doctoral dissertation, Universität Duisburg-Essen]. https://duepublico2.uni-due.de/servlets/MCRFileNodeServlet/duepublico_derivate_00042707/Neubaum_Diss.pdf

- Neubaum, G., & Krämer, N. C. (2017). Monitoring the opinion of the crowd: Psychological mechanisms underlying public opinion perceptions on social media. *Media Psychology*, 20(3), 502–531. <https://doi.org/10.1080/15213269.2016.1211539>
- Neubaum, G., & Krämer, N. C. (2018). What do we fear? Expected sanctions for expressing minority opinions in offline and online communication. *Communication Research*, 45(2), 139–164. <https://doi.org/10.1177/0093650215623837>
- Nisic, N., & Auspurg, K. (2009). Faktorieller Survey und klassische Bevölkerungsumfrage im Vergleich - Validität, Grenzen und Möglichkeiten beider Ansätze [A comparison of factorial surveys and traditional population surveys—Validity, limitations, and potential of both approaches]. In P. Kriwy & C. Gross (Eds.), *Klein aber fein! Quantitative empirische Sozialforschung mit kleinen Fallzahlen [Small but mighty! Quantitative empirical social research with small sample sizes]* (pp. 211–246). VS Verl. für Sozialwiss; GWV Fachverlage. https://doi.org/10.1007/978-3-531-91380-3_9
- Obermaier, M., Fawzi, N., & Koch, T. (2016). Bystanding or standing by? How the number of bystanders affects the intention to intervene in cyberbullying. *New Media & Society*, 18(8), 1491–1507. <https://doi.org/10.1177/1461444814563519>
- Pabian, S. (2025). Bystanders' perceptions on online hate speech: Investigating the effects of perpetrators' justifications and the bystander's role on bystanders' attitude and prosocial intervention intentions. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 19(4), Article 6. <https://doi.org/10.5817/CP2025-4-6>
- Peng, R. X., & Chen, J. (2025). Meme-ingful connections: Unleashing the power of memes, GIFs, and emojis in relationship-oriented online communication. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 19(4), Article 11. <https://doi.org/10.5817/CP2025-4-11>
- Ping, K., Kumar, A., Ding, X., & Rho, E. H. (2025). Behind the counter: Exploring the motivations and barriers of online counterspeech writing. *ACM Transactions on Computer-Human Interaction*, 32, Article 55. <https://doi.org/10.1145/3745769>
- Preece, J., Nonnecke, B., & Andrews, D. (2004). The top five reasons for lurking: Improving community experiences for everyone. *Computers in Human Behavior*, 20(2), 201–223. <https://doi.org/10.1016/j.chb.2003.10.015>
- Prentice, D. A., & Miller, D. T. (1996). Pluralistic ignorance and the perpetuation of social norms by unwitting actors. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 161–209). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60238-5](https://doi.org/10.1016/S0065-2601(08)60238-5)
- Quintal, É., & Collen, D. (2023). *Weaponized laughter: Memes and hate in the Canadian digital landscape: A report by the online hate research and education project produced by the Toronto Holocaust Museum*. Toronto Holocaust Museum. <https://cdn.sanity.io/files/rdq6owfff/production/6b78f8630669069025ea145da2221ef2c1fac032.pdf>
- Rendsvig, R. K. (2014). Pluralistic ignorance in the bystander effect: Informational dynamics of unresponsive witnesses in situations calling for intervention. *Synthese*, 191(11), 2471–2498. <https://doi.org/10.1007/s11229-014-0435-0>
- Rivis, A., & Sheeran, P. (2003). Descriptive norms as an additional predictor in the theory of planned behaviour: A meta-analysis. *Current Psychology*, 22(3), 218–233. <https://doi.org/10.1007/s12144-003-1018-2>
- Roberts, S. O., Bareket-Shavit, C., Dollins, F. A., Goldie, P. D., & Mortenson, E. (2020). Racial inequality in psychological research: Trends of the past and recommendations for the future. *Perspectives on Psychological Science*, 15(6), 1295–1309. <https://doi.org/10.1177/1745691620927709>
- Robertson, C. E., Del Rosario, K. S., & van Bavel, J. J. (2024). Inside the funhouse mirror factory: How social media distorts perceptions of norms. *Current Opinion in Psychology*, 60, Article 101918. <https://doi.org/10.1016/j.copsyc.2024.101918>
- Rossi, L., Segerberg, A., Arminio, L., & Magnani, M. (2025). Do you see what I see? Emotional reaction to visual content in the online debate about climate change. *Environmental Communication*, 19(3), 449–467. <https://doi.org/10.1080/17524032.2024.2420787>

- Roth, J., & Mazziotta, A. (2015). Adaptation and validation of a German Multidimensional and Multicomponent Measure of Social Identification. *Social Psychology*, 46(5), 277–290. <https://doi.org/10.1027/1864-9335/a000243>
- Schmid, U. K. (2023). Humorous hate speech on social media: A mixed-methods investigation of users' perceptions and processing of hateful memes. *New Media & Society*, 27(3), 1588–1606. <https://doi.org/10.1177/14614448231198169>
- Schmid, U. K., Kümpel, A. S., & Rieger, D. (2022). How social media users perceive different forms of online hate speech: A qualitative multi-method study. *New Media & Society*, 26(5), 2614–2632. <https://doi.org/10.1177/14614448221091185>
- Schmid, U. K., Schulze, H., & Drexel, A. (2025). Memes, humor, and the far right's strategic mainstreaming. *Information, Communication & Society*, 28(4), 537–556. <https://doi.org/10.1080/1369118X.2024.2329610>
- Schmitt, J. B., Harles, D., & Rieger, D. (2020). Themen, motive und mainstreaming in rechtsextremen online-Memes [Themes, motifs, and mainstreaming in far-right online memes]. *Medien & Kommunikationswissenschaft*, 68(1–2), 73–93. <https://doi.org/10.5771/1615-634X-2020-1-2-73>
- Sherif, M. (1936). *The psychology of social norms*. Harper.
- Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on Generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, 12(6), 1123–1128. <https://doi.org/10.1177/1745691617708630>
- Sorrels, J. P., & Kelley, J. (1984). Conformity by omission. *Personality and Social Psychology Bulletin*, 10(2), 302–305. <https://doi.org/10.1177/0146167284102017>
- Southern, R. (2022). Memes; a new emerging logic: Evidence from the 2019 British general election. In D. Taras & R. Davis (Eds.), *Electoral campaigns, media, and the new world of digital politics* (pp. 200–220). University of Michigan Press. <https://www.jstor.org/stable/10.3998/mpub.12013603.13>
- Sponholz, L. (2021). Hass mit Likes: Hate Speech als Kommunikationsform in den Social Media [Hate with likes: Hate speech as a communication form in social media]. In S. Wachs, B. Koch-Priewe, & A. Zick (Eds.), *Hate Speech - Multidisziplinäre Analysen und Handlungsoptionen [Hate Speech: Multidisciplinary Analyses and Action Options]* (pp. 15–37). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-31793-5_2
- Steiner, P., Atzmüller, C., & Su, D. (2016). Designing valid and reliable vignette experiments for survey research: A case study on the fair gender income gap. *Journal of Methods and Measurement in the Social Sciences*, 7(2), 52–94. <https://doi.org/10.2458/v7i2.20321>
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of inter-group conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of inter-group relations* (pp. 33–47). Brooks/Cole.
- Tontodimamma, A., Nissi, E., Sarra, A., & Fontanella, L. (2021). Thirty years of research into hate speech: Topics of interest and their evolution. *Scientometrics*, 126(1), 157–179. <https://doi.org/10.1007/s11192-020-03737-6>
- Tuters, M. (2019). Larping & liberal tears. Irony, belief and idiocy in the deep vernacular web. In M. Fielitz & N. Thurston (Eds.), *Post-digital cultures of the far right: Online actions and offline consequences in Europe and the US* (pp. 37–48). Transcript. <https://doi.org/10.1515/9783839446706-003>
- Wachs, S., Wettstein, A., Bilz, L., & Gámez-Guadix, M. (2022). Adolescents' motivations to perpetrate hate speech and links with social norms. *Comunicar*, 30(71), 9–20. <https://doi.org/10.3916/C71-2022-01>
- Wachs, S., & Wright, M. F. (2018). Associations between bystanders and perpetrators of online hate: The moderating role of toxic online disinhibition. *International Journal of Environmental Research and Public Health*, 15(9), Article 2030. <https://doi.org/10.3390/ijerph15092030>
- Wagener, A. (2021). The postdigital emergence of memes and GIFs: Meaning, discourse, and hypernarrative creativity. *Postdigital Science and Education*, 3(3), 831–850. <https://doi.org/10.1007/s42438-020-00160-1>
- Wiggins, B. E., & Bowers, G. B. (2015). Memes as genre: A structurational analysis of the memescape. *New Media & Society*, 17(11), 1886–1906. <https://doi.org/10.1177/1461444814535194>
- Wijenayake, S., & Goncalves, J. (2025). A review of online social conformity: Outcomes and determinants. *International Journal of Human-Computer Interaction*, 41(15), 9274–9303. <https://doi.org/10.1080/10447318.2024.2424385>

- Winkler, Y., Bachl, M., & Scharkow, M. (2022). Individual users' participation on political Facebook pages. *Journal of Quantitative Description: Digital Media*, 2, 1–40. <https://doi.org/10.51685/jqd.2022.019>
- Wu, T.-Y., & Atkin, D. J. (2018). To comment or not to comment: Examining the influences of anonymity and social support on one's willingness to express in online news discussions. *New Media & Society*, 20(12), 4512–4532. <https://doi.org/10.1177/1461444818776629>
- Yoon, I. (2016). Why is it not just a joke? Analysis of internet memes associated with racism and hidden ideology of colorblindness. *Journal of Cultural Research in Art Education*, 33(1), 92–123. <https://doi.org/10.2458/jcrae.4898>
- You, L., & Lee, Y.-H. (2019). The bystander effect in cyberbullying on social network sites: Anonymity, group size, and intervention intentions. *Telematics and Informatics*, 45, Article 101284. <https://doi.org/10.1016/j.tele.2019.101284>
- YouGov. (2024). *Leading methods for personal communication among global consumers in December 2023*. <https://www.statista.com/statistics/1459143/most-used-communication-ways-worldwide/>
- Zhao, S. Z., Luk, T. T., Guo, N., Wang, M. P., Lai, A. Y. K., Wong, B. Y. M., Fong, D. Y. T., Chan, S. S. C., & Lam, T. H. (2021). Association of mobile instant messaging chat group participation with family functioning and well-being: Population-based cross-sectional study. *Journal of Medical Internet Research*, 23(3), Article e18876. <https://doi.org/10.2196/18876>
- Zick, A., Wolf, C., Küpper, B., Davidov, E., Schmidt, P., & Heitmeyer, W. (2008). The syndrome of group-focused enmity: The interrelation of prejudices tested with multiple cross-sectional and panel data. *Journal of Social Issues*, 64(2), 363–383. <https://doi.org/10.1111/j.1540-4560.2008.00566.x>

Appendix

Table A1. *Sociodemographic Characteristics of Participants.*

Characteristics	Full sample	
	<i>n</i>	%
Gender		
Female	652	60
Male	429	39
Divers	9	1
Highest educational level ^a		
No formal school degree	1	0
Lower secondary school certificate (Hauptschule)	11	1
Intermediate secondary school certificate (Mittlere Reife)	66	6
Completed vocational training	75	7
High school diploma (Abitur or Fachabitur)	228	21
University or postgraduate degree	685	63
Experiences of discrimination ^b		
Based on race	25	2
Based on other factors	212	19
No experience	744	68

Note. *N* = 1,090. Participants were on average 49.8 years old (*SD* = 13.8). ^a “other”-option not displayed, ^b “not sure”-option and no answer-option not displayed.

Table A2. *Measurement of the Group's Importance. Item Overview and Translation.*

Original German item	Translation
Für mich wäre diese Chatgruppe von großer Bedeutung.	This chat group would be very important to me.
Es wäre mir wichtig, was diese Chatgruppe über mich denkt.	It would matter to me what this chat group thinks of me.
Es wäre mir total egal, was in dieser Chatgruppe passiert. (R)	I wouldn't care at all about what happens in this chat group. (R)
Ich würde mich der Chatgruppe eng verbunden fühlen.	I would feel a strong sense of connection to the chat group.
Was die Chatgruppe denkt oder macht, würde mich auch beschäftigen.	What the chat group thinks or does would affect me as well.

About Author

Christina Josupeit is a Ph.D. candidate in the Faculty of Educational Sciences at the University of Duisburg-Essen. She is interested in how society is held together during times of multiple crises and polarization. Consequently, her research focuses on digital civic courage and group dynamics. She is also a research assistant and lecturer at the Faculty of Applied Social Sciences at TH Köln – University of Applied Sciences.

<https://orcid.org/0000-0003-4722-7774>

Correspondence to

Christina Josupeit, Faculty of Applied Social Sciences, TH Köln, Gustav-Heinemann-Ufer 54, 50968 Köln, Germany.
christina.josupeit@th-koeln.de

© Author(s). The articles in Cyberpsychology: Journal of Psychosocial Research on Cyberspace are open access articles licensed under the terms of the [Creative Commons BY-SA 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/) which permits unrestricted use, distribution and reproduction in any medium, provided the work is properly cited and that any derivatives are shared under the same license.