

**Lonely Together? Identifying the Determinants of Collective Corrective Action Against
Uncivil Comments**

Marc Ziegele, Teresa K. Naab & Pablo Jost

The published version of this manuscript is available here:

Ziegele, M., Naab, T. K. & Jost, P. (2019). Lonely together? Identifying the determinants of collective corrective action against uncivil comments. *New Media & Society*, advance online publication. <https://doi.org/10.1177/1461444819870130>

Correspondence concerning this article should be addressed to

Teresa K. Naab, Department of Media, Knowledge and Communication, University of

Augsburg, Universitaetsstrasse 10, 86159 Augsburg, Germany. E-Mail: teresa.naab@phil.uni-augsburg.de, phone: +498215985933

Abstract

Journalists, scholars, and citizens have raised concerns regarding the high share of incivility in comment sections of news outlets. The current study surveyed members of the social movement #ichbinhier, which aims at collectively countering uncivil comments to cultivate a civil discussion atmosphere in comment sections. We root the activities of #ichbinhier as corrective action and identify the determinants of the members' engagement by integrating research on bystander behavior and collective action. The findings of our survey show that factors pertaining to individual skills, perceived responsibility, and expected benefits relate to the members' likelihood to engage against uncivil online comments. Regarding factors derived from collective action research, group efficacy and knowledge of the rules and structures of the movement account for higher levels of engagement. These results shed light on the factors that motivate and inhibit #ichbinhier members—and, potentially, Facebook users in general—to engage against uncivil comments.

Keywords: corrective action, collective action, bystander behavior, counter-speech, online comments

Lonely Together? Identifying the Determinants of Collective Corrective Action Against Uncivil Comments

Uncivil user comments in the comment sections of news websites are worrisome to readers and news outlets (Coe et al., 2014; Stroud et al., 2015). Empirical research has documented various negative effects of incivility in comment sections on readers, including attitude polarization (Anderson et al., 2014), arousal of negative affect (Ziegele et al., 2018b), increased levels of incivility in the subsequent comments (Hsueh et al., 2015), and reduced prosocial behavior towards the social groups targeted by uncivil commenters (Ziegele et al., 2018a). Journalists and professional community managers struggle to moderate the seemingly ever-increasing number of uncivil comments (Bergström and Wadbring, 2014). It is therefore important to investigate under which circumstances users themselves are willing to engage against uncivil comments, thereby potentially complementing the efforts of news outlets and professional moderators to cultivate civil and constructive online discussions.

In fact, on many platforms, users can report uncivil comments to site administrators. Additionally, they can write balancing comments or, for example, ask uncivil commenters to discuss more respectfully. Still, many users who would engage against uncivil comments refrain from doing so because they perceive that their individual voice does not impact the overall discussion (Stroud et al., 2016). The current study therefore investigates a social movement that builds a comprehensive structure around citizens' individual and voluntary activities in comment sections: The Facebook group #ichbinhier (#ibh, German for #iamhere), a spin-off of the Swedish group #jagärhär, was founded in 2016 and had more than 39,000 members in 2018 (Kreißel et al., 2018). The movement, which calls itself an "online action group", is also active in the United Kingdom as well as in Norway and Finland. Its overarching goal is to collectively improve the quality of discussions in comment sections on the Facebook sites of established news media by countering uncivil comments with rational and respectful contributions (Ley, 2018). For this purpose, #ibh members collectively organize and execute so-called "actions".

During these actions, they systematically engage in selected comment streams and counter uncivil comments with civil contributions. Members are also encouraged to upvote the comments of their fellow members.

Despite the collective aim of the movement, only a small share of the #ibh members actually intervene in uncivil comment sections (Ley, 2018). The present study therefore investigates the factors that motivate or inhibit #ibh members to actively counter uncivil comments. As a theoretical contribution, we conceptualize the activities of #ibh members as corrective action and integrate motivational factors from research on bystander behavior and collective action into an intervention framework. The resulting hypotheses are tested with an online survey of 576 #ibh members. The findings shed light on the interplay between personal motivations of the members to engage in collective corrective action—such as perceived responsibility and expected personal benefits—and motivations pertaining to the collective structure provided by the group itself – such as a shared social identity and perceived group efficacy. The study thereby illustrates the challenges associated with user-initiated forms of collective corrective action and provides insights into how additional users could be encouraged to become more involved.

Incivility in Comment Sections

The aim of #ibh is to improve online discussions that have drifted to undemocratic and aggravatingly disrespectful opinion expressions (Ley, 2018). Members of the action group regularly scan the Facebook pages of large media outlets to identify “disrespectful, sweeping, inflammatory and/or derogatory comments below the news articles” (Kreißel et al., 2018: 9, authors’ translation). Such comments, from a conceptual perspective, can be considered manifestations of uncivil communication (e.g., Coe et al., 2014). Although various definitions of incivility exist (e.g., Chen, 2017; Herbst, 2010), Coe et al.’s (2014) definition is particularly tailored to user comments. They define incivility as “features of discussion that convey an unnecessarily disrespectful tone toward the discussion forum, its participants, or its topics” (p.

660). These features include name-calling, aspersion, lying, vulgarity, and pejorative speech (Coe et al., 2014). Other scholars have complemented this list with various forms of undemocratic speech, such as the use of racist or sexist terms (Chen, 2017). This understanding of incivility aligns well with the comments targeted by #ibh. However, it should be noted that #ibh members, and online users in general, may have individual perceptions of civil and uncivil communication (e.g., Chen, 2017; Herbst, 2010). In fact, research has shown that “a range of background characteristics—sex, ideology, the personality trait of agreeableness—may influence a given audience member’s response to incivility” (Kenski et al., 2017: 16). These different perceptions could affect the decision of #ibh members to engage in online discussions. Still, previous research has also found evidence that academic conceptions and user perceptions of incivility overlap to some extent (Kalch and Naab, 2017). Therefore, we proceed on the premise that #ibh members primarily campaign against incivility in user comments.

#ichbinhier and Corrective Action

The #ibh members intervene in comment sections that exhibit high degrees of incivility. In such discussions, they add comments that comply to the standards of a civil exchange (Ley, 2018). They demonstrate respect for opposing views, introduce additional perspectives on an issue, present facts, and take position for discriminated and stigmatized groups. Thereby, the members aim at counterbalancing the otherwise negative effects of uncivil discussions (e.g., Hsueh et al., 2015). On an abstract level, these actions can be understood as corrective action. The term has been coined by research on presumed media influence. It describes actions that are reactive to (media) content perceived as harmful and that seek to influence the public sphere and counterbalance the negative influences of the content (Lim, 2017; Rojas, 2010). Recent empirical research on corrective action online includes sharing alternative opinions or information (Bernhard and Dohle, 2015; Golan and Lim, 2016) and engaging in online discussions to correct biases (Lim and Golan, 2011; Rojas, 2010). Some of these activities have also been conceptualized as political participation (e.g., Gil de Zúñiga et al., 2012). Indeed,

corrective action may be understood as a particular form of political participation that is “reactive, based on perceptions of media and media effects, and seek[s] to influence the public sphere” (Rojas, 2010: 347). The activities of #ibh aim at balancing uncivil discussions in the public comment sections of news outlets that could otherwise adversely affect the readers of the discussions and have undesirable effects on public opinion and, eventually, on political decision making (Ley, 2018). We therefore conceptualize the activities of #ibh as corrective action.

Determinants of #ichbinhier’s Corrective Action: A Bystander Perspective

Corrective Action as Bystander Intervention

Bystander research originally investigated why people (or bystanders, respectively) do or do not provide help when they observe emergencies, such as violent attacks and accidents (Fischer et al., 2011). Latané and Darley (1970) described the psychological process of bystander intervention as a five-step model: Bystanders need to 1) notice a critical situation, 2) construe it as an emergency, 3) feel responsible, 4) believe they have the skills to help, and 5) weigh the costs and benefits of a potential intervention (Fischer et al., 2011; Latané and Nida, 1981). During each step, a tendency to intervene may be reconsidered.

Recent research applied the approach to computer-mediated situations and examined, for example, why bystanders intervene in incidents of cyber-bullying (Obermaier et al., 2016) or calls for help in emails and online fora (Barron and Yechiam, 2002; Blair et al., 2005). Uncivil online discussions, from the perspective of bystander research, can be conceptualized as incidents that require intervention. However, these discussions differ from offline situations in several respects. First, in uncivil online discussions, neither victims nor potential helpers are in physical danger. Yet, recent work on bystander behavior supports that the approach holds for highly dangerous as well as less harmful situations, like answering the door or reporting a broken tape recorder (Chekroun and Brauer, 2002). Additionally, various studies have found harmful psychological effects of the use of uncivil language on readers and victims (Chen, 2017; Obermaier et al., 2016), which suggests that public online incivility can be considered a severe

incident. Second, the audience of uncivil user comments is potentially larger than the number of bystanders in offline emergencies. Many online audiences are also at least partly anonymous and only virtually present. Still, previous studies suggested that these conditions do not fundamentally alter the psychological mechanisms why users engage in certain behaviors (Berg, 2016; Rösner and Krämer, 2016). Third, many offline emergencies involve individual victims while many uncivil comments are directed against ideas, social groups, and values in general (Chen, 2017; Coe et al., 2014). Nonetheless, research has shown that the bystander model also applies to cases when people have to decide to intervene against deviant behavior directed against public goods or values; for example, when someone litters a public park or draws graffiti on an elevator (Chekroun and Brauer, 2002). In sum, it seems legitimate to apply the bystander approach to uncivil online discussions and to the analysis of why #ibh members engage in corrective action in comment sections.

Step 1: Recognizing a Critical Situation

Individuals cannot intervene unless they are aware of the potentially critical situation (Latané and Darley, 1970). For the current study, this means that users first have to encounter uncivil discussions and interpret them as uncivil. In survey research, a user's likelihood of encountering critical situations during online use can hardly be measured directly (this is different in experimental research that uses particular referential stimulus material). However, the frequency of using Facebook as a source of political information and the frequency of reading political news and comments on Facebook could serve as proxies for the level of attention #ibh members pay towards potentially uncivil discussions due to various reasons: First, the more political news articles and comments users read on Facebook, the higher is their chance to encounter uncivil comments. Second, political Facebook usage can increase users' sensitivity toward incivility. Research has found that the more frequently users read online news, the more likely they can identify elements of incivility (Kenski et al., 2017). Therefore, #ibh members

who frequently use Facebook as a source of political information could be more likely to encounter uncivil comments and identify them as uncivil:

H1: The frequency of using Facebook as a source for political news and comments will be positively related to the willingness of #ibh members to engage in corrective action.

Step 2: Construing a Situation as an Emergency

Observers are more likely to intervene in situations they unambiguously construe as harmful and calling for intervention (Fischer et al., 2006; Fischer et al., 2011). Ambiguous situations, in contrast, allow bystanders to perceive the event as less harmful, which lowers the costs of not helping (Solomon et al., 1978). The significance of perceiving a situation as harmful is also documented for online incidents. Research confronting participants with specific online content has shown that the more harmful the participants rate the effects of this content, the higher is their tendency to engage in corrective action (Golan and Lim, 2016; Lim and Golan, 2011). Transferred to interventions in online discussions (and applied to a survey design without predetermined stimulus comments), #ibh members should be more willing to counterbalance uncivil comments with corrective action the more harmful these members perceive the effects of current political online discussions and user comments in general.

H2: Perceived harmfulness of current political online discussions and user comments will be positively related to the willingness of #ibh members to engage in corrective action.

Step 3: Feeling Responsible

After perceiving an incident as harmful, bystanders need to feel responsible to engage rather than transfer responsibility to other people (Fischer et al., 2011; Thornberg, 2007). Public debates about uncivil online discussions, however, often emphasize that platform providers and state authorities should regulate deviant comments (Flew et al., 2019). Users may thus attribute the responsibility to regulate comment sections to authorities like media outlets or legal prosecutors. In fact, a study demonstrated that online chat users were less likely to intervene in an act of child grooming when they were informed before that the chat was under computer

surveillance compared to when they were informed that chat users themselves would need to supervise participants' adherence to the community standards (Palasinski, 2012). In contrast, perceiving personal responsibility has been found to increase users' engagement in offline and online corrective action (Baumeister et al., 1988; Naab et al., 2018).

H3: Feelings of personal responsibility will be positively related to the willingness of #ibh members to engage in corrective action.

Step 4: Feeling Able to Intervene

When bystanders feel able and competent to provide help, they are more likely to intervene in emergencies (Burn, 2009; Cramer et al., 1988). The specific factors that increase bystanders' perceptions of feeling able to intervene, however, depend on the nature of the emergency that needs intervention. Therefore, we draw on research on social media activism and online corrective action (e.g., Bruning et al., 2013; Feldman et al., 2017; Golan and Lim, 2016) to discuss internal political efficacy and writing self-efficacy as relevant skills for efficient intervention in uncivil political online discussions.

Political efficacy is comprised of internal and external political efficacy. We focus on internal efficacy, which describes a person's beliefs to have the competencies to understand and participate effectively in politics (Kenski and Stroud, 2006). It is closely related to other concepts that predict political participation and activism, such as political knowledge (Jung et al., 2011; Kenski and Stroud, 2006). Reviewing various studies, Feldman et al. (2017) note that internal efficacy is a consistent predictor of political participation, including (offline) activism and collective action. Velasquez and LaRose (2015) also found a link between internal political efficacy and the likelihood of participating in social media activism. As #ibh campaigns for a better discussion culture in *political* online discussions, it can be assumed that a high level of internal political efficacy goes along with a high willingness to intervene in such discussions:

H4: Perceived internal political efficacy will be positively related to the willingness of #ibh members to engage in corrective action.

To engage in specific forms of participation, users may additionally need skills that are closely related to these activities. For comment writing, one important skill involves the ability of efficiently putting one's thoughts into words (Ziegele, 2016). Such confidence can be seen as a specific form of self-efficacy and consequently has been investigated as a part of the concept of writing self-efficacy (Bruning et al., 2013). It can be assumed that high writing self-efficacy decreases the time needed to write comments. High time requirements, in turn, are a significant inhibitor of comment writing (Springer et al., 2015). Therefore, we hypothesize:

H5: Perceived writing self-efficacy will be positively related to the willingness of #ibh members to engage in corrective action.

Step 5: Costs and Benefits of Intervening

Bystanders estimate the costs and benefits of getting involved before intervening (Dovidio et al., 1997). When the perceived costs of helping outweigh the perceived benefits, intervention is less likely (Burn, 2009). Corrective action in uncivil online discussions comes with little physical costs. However, users run the risk of getting into conflict with other commenters or even falling victim to online harassment (Springer et al., 2015; Stroud et al., 2016). Additionally, users may fear to be negatively judged by other observers when engaging in an uncivil discussion. Such evaluation apprehension makes bystanders reluctant to intervene in general (Schwartz and Gottlieb, 1980).

At the same time, engaging in corrective action can have several benefits. For example, users may receive social rewards like recognition and praise for good comments from other users or fellow #ibh members (van Bommel et al., 2012). Disagreeing with uncivil comments can also yield affective gratifications, such as feeling better afterwards (Springer et al., 2015). In sum, we therefore hypothesize:

H6: Perceived risks of intervening in uncivil discussions will be negatively related to the willingness of #ibh members to engage in corrective action.

H7: Perceived individual benefits of intervening in uncivil discussions will be positively related to the willingness of #ibh members to engage in corrective action.

Determinants of #ichbinhier's Corrective Action: A Collective Action Perspective

Collective action “refers to actions undertaken by (...) groups for a collective purpose, such as the advancement of a particular ideology or idea or the political struggle with another group” (Brunsting and Postmes, 2002: 527). Groups that perform collective action vary in their aims and formal organization, but many of their underlying psychological mechanisms and motivations are similar (van Zomeren et al., 2008) and include the perception of a social grievance or injustice, a shared social identity, and a sense of collective efficacy to change society for the better (Kelly and Kelly, 1994; Postmes and Brunsting, 2002).

The #ibh movement shares important characteristics of such groups. First, by defining civil and constructive online discussions as a desirable condition, #ibh offers its members a distinct ideology that resembles normative concepts of online deliberation (e.g., Friess and Eilders, 2015). The #ibh administrators also frame the engagement of the members as a struggle between opposing groups, or, in other words, between democratic and civil citizens on the one hand and less democratic and uncivil commenters on the other one (Ley, 2018).

Second, #ibh and social movements share the principles of formal organization and organized actions (e.g., Bimber et al., 2005). The #ibh members gather in a closed Facebook group where they collectively plan and organize their actions. They also talk about the rules and procedures of the movement and discuss new tools that could potentially support their engagement.

Third, the psychological mechanisms and motivations underlying the #ibh movement align with those of collective action. By arguing that uncivil voices have unjustifiably gained the upper hand in online discussions (Ley, 2018), #ibh declares a current social grievance and a relative injustice. At the same time, the group offers its members a collective identity (i.e., being

a campaigner for a better discussion culture; Ley, 2018) and a sense of group efficacy (i.e., an organized group of users can change the quality of online discussions).

Reviewing these arguments, it seems promising to integrate factors from collective action research to explain #ibh members' engagement in uncivil comment sections. Important factors include a shared social identity, perceived group efficacy, perceived collective benefits, and knowledge of the formal rules and structures of the social movement (Bimber et al., 2005; Brunsting and Postmes, 2002; Kelly and Kelly, 1994). Considering these factors acknowledges the collective character of the #ibh movement and allows investigating its relative importance to the members' personal motives.

Collective Identity

Research has shown that sharing an ingroup identity increases the willingness of group members to participate in collective action (Kelly and Kelly, 1994; Postmes and Brunsting, 2002). This should also apply to #ibh members who strongly identify with the proposition of the group to *collectively* engage in corrective action against uncivil comments, who feel similar to other group members, and who share the idea of being a campaigner for a better discussion culture.

H8: The strength of a shared social identity with the #ibh movement will be positively related to the willingness of #ibh members to engage in corrective action.

Group Efficacy

Group efficacy is defined as the “collective belief that group-related problems can be solved by collective effort” (Van Zomeren et al., 2004: 651). Previous research has demonstrated that members of social movements are more likely to engage in collective action when they perceive their group as capable of transforming social grievances through unified efforts (Postmes and Brunsting, 2002; van Zomeren et al., 2008). Since organized actions that aim at improving the quality of online discussions are at the core of the activities of #ibh, members

should be more willing to engage in uncivil comment sections if they perceive that the collective efforts of the group can make a difference.

H9: Perceived group efficacy will be positively related to the willingness of #ibh members to engage in corrective action.

Collective Benefits

Collective benefits or motives can be seen as a complement to individual benefits of engaging in collective action (Brunsting and Postmes, 2002; Klandermans, 1984; Simon et al., 1998). They describe how likely group members think that the overarching collective goal of the social movement (in the case of #ibh: a civil and deliberative discussion culture) will be achieved and how much they value this goal. Previous research has shown that the expectation of such collective benefits increases people's participation in various social movements (Postmes and Brunsting, 2002; Simon et al., 1998).

H10: Perceived collective benefits of intervening in uncivil discussions will be positively related to the willingness of #ibh members to engage in corrective action.

Knowledge of #ibh Rules and Structures

Social movements can overcome various obstacles related to effective collective action by setting up rules to coordinate the efforts of their members or by designating decision makers with differentiated roles (Bimber et al., 2005). However, an abundance of rules and procedures can also inhibit the engagement of individual members (e.g., Guldberg and Mackness, 2009). Becoming an active campaigner for #ibh requires members to understand the different social roles in the group (e.g., administrators, moderators, and regular users), the process of selecting comment threads for intervention, and the rules of respectful comment writing. Members might feel less confident to engage for #ibh's collective aim if they feel that they do not (yet) fully understand these rules and structures (Ley, 2018). We therefore hypothesize:

H11: Perceived knowledge about the rules and formal structure of #ibh will be positively related to the willingness of #ibh members to engage in corrective action.

Method

In July 2018, the administrators of the #ibh group distributed our online questionnaire among their members. They posted three messages that included the link to our survey in the closed Facebook group. At the time of the survey, #ibh had 44,783 members. Overall, 4,640 users visited the introduction page and 632 individual respondents completed the survey. We excluded 56 respondents because they did not identify as #ibh members. The final sample consisted of $N = 576$ respondents. Seventy-eight percent were female and the average age of the respondents was 46.29 years ($SD = 11.61$). Fifteen percent had completed secondary school or less. Thirty-three percent had received a university-entrance diploma and 52 percent of the participants were university graduates.

While this is the first study to investigate an internationally growing social movement that campaigns for a better online discussion culture, we acknowledge that our sample covers a specific population in a specific context. Still, the respondents show sufficient variance in intervention behavior: Eighty percent said that they had already engaged in discursive corrective action against uncivil comments for #ibh, although most of them (74 percent) did not do so on a regular basis. Nevertheless, the high share of #ibh members who had already written public comments shows that these members are generally more active than average users (across most Western countries, only between ten and thirty percent of online users write comments, see Newman, Fletcher, Kalogeropoulos, Levy, & Nielsen, 2018).

Measures

All constructs, scales (including reliabilities), and items are documented in the Online Appendix.

Perceptions of uncivil/harmful comments. We approach incivility from a users' perspective. Yet, in terms of language, it is problematic that there is no equivalent German term for the English word "incivility". Certainly, the phenomena that are subsumed under this concept (e.g., name-calling) are known phenomena in Germany, too. Members of #ibh thus aim at

decreasing what academia calls incivility without actually using that academic term. In our survey, we therefore used the term “harmful discussions” or “harmful comments” as substitute terms. To ensure that the members’ perceptions of harmful comments relate to the academic concept of incivility, we asked them to rate the harmfulness of five types of comments that previous research had classified as uncivil. These comments included a) name-calling/impoliteness, b) lies, c) threats, d) antagonistic stereotypes, or e) pejorative speech.

Willingness to engage in corrective action. After providing their ratings of harmful comments, respondents were asked how likely they would intervene in such discussions by writing balancing comments and by directly replying to one or several harmful comments ($1 =$ very unlikely to $5 =$ very likely). Additionally, as an indirect form of corrective action, we asked respondent whether they would respond with supportive comments to high-quality comments in otherwise harmful discussions. The resulting *discursive corrective action scale* showed a high internal consistency ($M = 3.43$, $SD = 1.03$, $\alpha = .84$).

Paying attention towards potentially uncivil discussions. Three questions were asked to measure how often respondents read Facebook news and user comments as a source of political information (e.g., “I read political news of established news media outlets on Facebook”, $M = 4.55$, $SD = 1.02$, $\alpha = .78$). Answers were recorded on six-point scales ($0 =$ never to $6 =$ very often).

Construing the situation as an emergency. Respondents indicated whether they perceive the effects of current online discussions and user comments on political topics as a chance or a threat to society. We assumed that the more the respondents perceive comments and discussions as a threat, the more unambiguously they also construe uncivil comments as an emergency. Answers were recorded with a slider on a scale ranging from $1 =$ significant chance to $100 =$ significant threat ($M = 67.05$, $SD = 20.04$).

Feeling responsible. Respondents were asked who they think is responsible for taking action against uncivil comments. The options included “I myself” as well as news outlets, law

enforcement, Facebook, and other users. Respondents indicated their agreement with each actor on a five-point scale ($1 =$ strongly disagree, $5 =$ strongly agree). To build a *personal responsibility scale*, we subtracted the mean value of the respondents' agreement with the perceived responsibility of "others" from their agreement with the statement that they themselves are responsible for taking action ($M = -0.60$, $SD = 0.90$).

Feeling able to intervene. Following Beierlein et al. (2012), respondents' internal political efficacy was measured using three items (e.g., "I am good at understanding and assessing important political issues", $M = 4.09$, $SD = 0.72$, $\alpha = .75$). Three items measured their perceived *writing self-efficacy* (e.g., "In general, it is easy for me to write down what I want to express", $M = 4.01$, $SD = 0.84$, $\alpha = .86$). Answers were recorded on five-point scales ($1 =$ strongly disagree to $5 =$ strongly agree).

Costs and benefits of intervening. We relied on an internal document from #ibh that included findings from qualitative interviews with activists who were asked about their perceived risks and benefits of intervening in uncivil discussions. Based on these findings, five items reflected the *perceived personal benefits* of intervening (e.g., "If I intervene with own comments in uncivil discussions, it will make me feel good", $M = 3.08$, $SD = 0.63$, $\alpha = .73$). The *risks of intervening* were measured with four items (e.g., "If I intervene (...), other users will insult me", $M = 3.52$, $SD = 0.70$, $\alpha = .83$). All answers were recorded on five-point scales ($1 =$ very unlikely to $5 =$ very likely).

Collective social identity. Borrowing a scale from collective action research (Brunsting and Postmes, 2002), respondents' level of perceived *collective social identity* was measured with five items on five-point scales ($1 =$ strongly disagree to $5 =$ strongly agree; e.g., "I have much in common with other members of #ichbinhier", $M = 3.82$, $SD = 0.60$, $\alpha = .71$).

Group efficacy. The perceived effectiveness of the movement's coordinated actions for reaching its overarching goal was measured with three items (e.g., "When people gather in groups to counter uncivil comments, they can sustainably improve the discussions in the

comment sections of news outlets on Facebook”; $1 =$ very unlikely to $5 =$ very likely, $M = 4.16$, $SD = 0.68$, $\alpha = .77$). These items were adapted from previous research on collective action (Brunsting and Postmes 2002; van Zomeren et al. 2004).

Collective benefits. Based on a review of the overarching goals and aims of the #ibh movement (Ley, 2018), we set up four items to measure the *perceived collective benefits* of intervening in uncivil online discussions (e.g., “I will contribute to a more balanced public opinion in our society”, $M = 4.01$, $SD = 0.63$, $\alpha = .71$). Answers were recorded on five-point scales ($1 =$ very unlikely to $5 =$ very likely).

Knowledge of the rules and formal structure of #ibh. This construct was measured using three self-developed items (e.g., “I fully understand the rules and procedures of #ibh”, $M = 4.30$, $SD = 0.70$, $\alpha = .84$). Answers were recorded on five-point scales ($1 =$ strongly disagree to $5 =$ strongly agree).

Results

Perceptions of Incivility / Harmfulness

Members of #ibh rated the comments that previous research had classified as uncivil as highly harmful. This was the case for impolite or insulting comments ($M = 4.74$, $SD = 0.52$), comments that contain lies ($M = 4.84$, $SD = 0.36$), comments that threaten others ($M = 4.91$, $SD = 0.36$), comments that include antagonistic stereotypes ($M = 4.92$, $SD = 0.33$), and comments that reject democracy ($M = 4.87$, $SD = 0.42$). These findings suggest that there is significant overlap between the users’ concept of harmfulness and the academic concept of incivility.

Hypotheses Testing

We computed a stepwise linear regression to test our hypotheses (Table 1). Listwise elimination of cases with missing values reduced the number of cases to $n = 545$. All variance inflation factors (VIF) were below the value of 1.3, indicating no significant multicollinearity issues. Participants’ willingness to engage in discursive corrective action in uncivil discussions was entered as the dependent variable. Research on citizen engagement has shown that this

engagement also differs depending on sociodemographic variables, such as age, sex, and education (Gil de Zúñiga et al., 2012). Therefore, we included these control variables in our models. Additionally, in each model (Models 1-6), specific predictors along the five steps of the bystander intervention framework and the collective action framework were added.

The results show that the more frequently #ibh members consumed political information and user comments on Facebook, the higher was their willingness to engage in corrective action (Model 6: $\beta = .09$, $p = .007$). These findings support H1. In contrast, the extent to which members perceived political online discussions and comments as a threat to society was unrelated to their willingness to engage in corrective action (Model 6: $\beta = -.05$, $p = .12$). H2 is rejected. The #ibh members' feelings of being personally responsible to regulate uncivil discussions—as opposed to attributing responsibility to authorities—increased their willingness to engage in corrective action (Model 5: $\beta = .11$, $p = .002$). These findings support H3. Regarding perceived skills, internal political efficacy initially was a marginally significant predictor of corrective action (Model 3: $\beta = .08$, $p = .082$) but became insignificant in the final model (Model 6: $\beta = .05$, $p = .142$). H4 is not supported. Consistent with our predictions, writing self-efficacy was positively related to the willingness of #ibh members to engage in corrective action (Model 6: $\beta = .25$, $p < .001$, H5 supported). The members were also more willing to post corrective comments in uncivil discussions the more likely they perceived their interventions to yield personal benefits (Model 6: $\beta = .20$, $p < .001$, H7 supported). However, negative consequences were positively related to their willingness to engage as well (Model 6: $\beta = .13$, $p < .001$, H6 rejected). We will discuss this unexpected finding in the next section.

Regarding collective factors, members' identification with #ibh, that is, their shared social identity, was unrelated to a higher willingness to post corrective comments (Model 5: $\beta = .01$, $p = .93$). H8 is rejected. Perceived group efficacy (Model 6: $\beta = .10$, $p = .005$) increased the willingness of the members to engage in corrective action (H9 supported), while their expectations of collective benefits were unrelated to their participation intentions (Model 6: $\beta =$

.07, $p = .09$, H10 rejected). Finally, perceived knowledge of #ibh members about the rules and structures of the movement was positively related to their willingness to engage in discursive corrective action (Model 6: $\beta = .11$, $p = .004$, H11 supported).

Discussion

The current study surveyed members of the social movement #ichbinhier, which aims at cultivating a civil discussion atmosphere in comment sections by collectively engaging in discursive corrective actions. We used bystander research and integrated research on collective action to develop a framework of the individual and collective determinants of online engagement against uncivil user comments. Overall, this framework was helpful to understand what motivates or inhibits #ibh members to engage in corrective action.

Members who frequently consumed political information on Facebook were more inclined to engage in corrective action. This finding aligns with research on political participation reporting that hard news consumption increases various forms of engagement (e.g., McLeod et al., 1999). We primarily interpreted our finding in terms of an increased awareness of potentially uncivil discussions. However, research has also shown that hard news consumption increases various competencies needed for efficient intervention, such as political knowledge and discussion skills (Kenski and Stroud, 2006; McLeod et al., 1999). Thus, frequent political Facebook usage seems to increase both users' awareness and skills that facilitate their engagement in corrective action.

Additionally, #ibh members who felt a high degree of personal responsibility to regulate online discussions were more willing to engage against uncivil comments than members who attributed this responsibility to external authorities. This finding lends support to the diffusion of responsibility thesis (e.g., Barron & Yechiam, 2002; Latané & Nida, 1981) and suggests that politics and news outlets could benefit from emphasizing the importance of user-driven engagement. Currently, many debates focus on how new state laws could make platform providers responsible for deleting inappropriate content (e.g., Flew et al., 2019). It is undisputed

that such regulatory procedures are necessary in some cases. However, these measures should be complemented by “bottom-up” corrective action of users. Our findings suggest that this could be stimulated by demonstrating to users the importance of their interventions and by appealing to their feeling of personal responsibility.

Besides users’ feelings of personal responsibility, their perceived ability to intervene was a relevant factor increasing the engagement in corrective action against uncivil comments. This is in line with previous bystander research (e.g., Burn, 2009). Particularly members who felt a high efficacy in comment writing were more likely to intervene in uncivil discussions, while a general sense of political efficacy was less important. This finding underpins the importance of using efficacy measures that are tailored to the specific context of intervention (Velasquez and LaRose, 2015). From a practical point of view, the findings suggest that measures to increase users’ level of engagement should aim at developing skills that are directly related to comment writing. For example, highlighting successful interventions could be an important motivation for citizens who hesitate whether their comments are “good enough”. In fact, a previous study reported that showing users how to intervene increased their likelihood of doing so (Naab et al., 2018).

Members of #ibh were also more willing to engage in corrective action when they expected to obtain personal benefits, such as reputation and appreciation. However, expectations of collective benefits that pertain to the overarching goal of the #ibh movement, such as contributing to a more balanced climate of opinion, did not increase the members’ willingness to intervene. Possibly, these benefits are too abstract to serve as motivators. In fact, there is no directly tangible outcome to judge the success of the #ibh actions to change the overall climate of online discussions. Another possible reason pertains to a “free ride dilemma” (Simon et al., 1998: 647): If the movement reaches its collective goal, each member of the society will benefit no matter if they had contributed to achieving it. This potentially lowers the value of the collective motive as a driving force of their engagement. Consequently, previous research has

suggested that potential free riders need “additional selective incentives to participate, which can involve both social and material (...) benefits” (Simon et al., 1998: 647). Our findings support this view by showing that the social and affective gratifications #ibh members gain from their actions are important predictors of their further engagement. Future research on online corrective action should therefore also focus on users who are willing to assign positive reactions to other comments. This kind of participation is less demanding and effortful (Karpf, 2010), yet it can motivate others to engage in corrective action against uncivil comments.

The current study also yielded an unexpected result: Expectations of personal risks increased users’ likelihood to engage in corrective action against uncivil comments. We assume that this indicates a problem of causality. Instead of personal risks being *expectations*, they could likely be *consequences* of engaging in corrective action. The more often members write corrective comments, the more often they may have experienced the risks of writing such comments. Then, personal risks do not *drive* these members’ engagement. Instead, they see insults and other risks as inevitable and *still* engage. Future longitudinal or experimental research is needed to investigate whether this issue of causality also applies to the benefits that #ibh members expect from their engagement.

Regarding the further factors derived from collective action research, perceived group efficacy predicted #ibh members’ engagement in collective corrective action. Group efficacy can be considered a complement to individual skills and abilities on the collective level. The more able the group is perceived in reaching its goal and the more the members share the belief that they are more powerful together, the more motivated they are to engage. Knowledge of the rules and structures of the #ibh group was also positively related to the members’ willingness to participate. While these rules and structures may be necessary to coordinate a movement’s actions (Bimber et al., 2005), this finding also implies that members who do not entirely see through these rules and structures will be inhibited from participating. To prevent the creation of

new “knowledge barriers”, movements such as #ibh therefore should keep their formal organization as simple as possible.

A shared social identity did not predict the willingness of #ibh members to engage in corrective action. Possibly, many members do not share an identification with the movement that goes beyond its basic idea to collectively improve the quality of online discussions. In fact, this interpretation resonates with research suggesting that some movements derive their collective identity from their shared aim of improving the recognition of their group or social category in society, while other movements derive their collective identity from their shared interest in defending “fundamental values that govern their social life” (van Zomeren and Spears, 2009: 663). In other words, the identity formation of #ibh members may not be based on a shared social category (which our measure of collective identity may have captured) but rather on the values that the movement collectively campaigns for.

Overall, these results suggest that the engagement of #ibh members in uncivil comment sections is primarily driven by their perceptions of personal responsibility, personal abilities, and expected personal benefits. On the level of the collective, only perceived group efficacy and knowledge of the rules and structures of #ibh served as additional motivating factors. These results differ from previous studies on collective action which found that a shared social identity increased the members’ level of participation (Brunsting and Postmes, 2002; van Zomeren et al., 2008). The results thus raise the question whether #ibh members act “lonely together”; in other words, although they can be considered a part of a collective action group, they ultimately contribute to the group’s aims based on an assessment of their personal motives. We did not measure whether other collective factors stimulated the members’ willingness to engage in corrective action, such as a perceived relative deprivation (Kelly and Kelly, 1994; Postmes and Brunsting, 2002). Still, if movements such as #ibh aim at increasing participation rates, it seems necessary that they appeal to their members’ personal responsibility as well as to their personal skills and benefits they can expect.

Limitations

The study has several limitations that need to be mentioned. First, we asked the respondents about their general assessment of uncivil comments and about their hypothetical engagement with these comments. While this procedure increased external validity compared to an experimental design with a limited number of predefined stimuli, it also somewhat blurs the results. We cannot fully reconstruct which kinds of uncivil user comments and discussions participants had in mind and whether specific forms of incivility in online discussions would stimulate corrective action. In line with this, we did not find a significant influence of the perceived harmfulness of online comments and discussions in general on participants' willingness to engage with corrective action (step 2 of the model). At a first glance, this finding contradicts the assumptions from bystander research. However, it could also be that our measure of risk perception was too abstract.

Additionally, we did not measure users' actual engagement in corrective action but rather their self-reported behavioral intention to do so. Still, more than two-thirds of the members surveyed reported that they had already written comments for #ibh. Additionally, research on the theory of planned behavior has shown that the intention to perform a specific behavior and the actual execution of this behavior are correlated (Armitage & Conner, 2001). In fact, in our sample, the correlation between respondents' willingness to engage in corrective action and their self-reported commenting frequency was $r = .65$. Nevertheless, these findings should be complemented by experimental designs in which users' actual engagement in corrective action is measured in specific, manipulated comment threads. This allows for measuring participants' perception of the emergency level in a specific comment thread and for analyzing their actual engagement in corrective action.

Although we repeatedly distributed the link to our survey in the #ibh group, only a limited share of members completed the questionnaire. This renders questions regarding the generalizability of our results. The #ibh group does not systematically survey its members,

therefore representative data on the population are not available. Yet, it seems unlikely that more than 50 percent of the members participate in corrective action against uncivil comments.

Therefore, our sample could be biased towards the more involved members. As we are interested in correlations rather than frequencies, it is still possible that the patterns of corrective action reported in this study also apply to #ibh members or to users in general.

Conclusion

Given the pressing social challenges arising from uncivil comments, further knowledge is needed on how to counter incivility in comment sections. Current debates focus on the need for regulation by state authorities and platform providers. In contrast, the social movement #ichbinhier appeals to users themselves and to their power to collectively engage for a better discussion atmosphere. This study is one of the first to analyze the collective engagement of such organized groups in online comment sections. Beyond the implications of the findings for the #ibh movement and for other forms of corrective action, the study integrated findings and concepts from different research directions and thereby contributes an interdisciplinary perspective on users' engagement in comment sections. Adapting the intervention framework from bystander research proved useful to explain users' engagement with uncivil discussions. It shows that the bystander model not only holds for "traditional" emergencies but also for less obvious "emergencies" such as uncivil attacks against social values and a deliberative discussion atmosphere in online comment sections. Research on collective action additionally allowed for a more nuanced view on bystander behavior in the context of mutual group support and situations that individuals can barely handle alone. This theoretical framework can be transferred to other forms of online and offline engagement with opposing standpoints and issues perceived as socially or individually harmful. Meanwhile, the findings of the current survey may stimulate further attention to user-initiated forms of corrective action that may complement institutionalized regulation and journalistic moderation of online discussions.

Funding

The first author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Parts of this research were supported by the Digital Society research program funded by the Ministry of Culture and Science of the German State of North Rhine-Westphalia (grant 1706dgn015).

References

- Anderson AA, Brossard D, Scheufele DA, et al. (2014) The “nasty effect:” Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication* 19: 373–387.
- Armitage CJ and Conner M (2001) Efficacy of the Theory of Planned Behavior: A meta-analytic reviews. *British Journal of Social Psychology* 40: 471–499.
- Barron G and Yechiam E (2002) Private e-mail requests and the diffusion of responsibility. *Computers in Human Behavior* 18: 507–520.
- Baumeister RF, Chesner SP, Senders PS, et al. (1988) Who’s in charge here? Group leaders do lend help in emergencies. *Personality & social psychology bulletin* 14: 17–22.
- Berg J (2016) The impact of anonymity and issue controversiality on the quality of online discussion. *Journal of Information Technology & Politics* 13: 37–51.
- Bergström A and Wadbring I (2014) Beneficial yet crappy: Journalists and audiences on obstacles and opportunities in reader comments. *European Journal of Communication* 30: 137–151.
- Bernhard U and Dohle M (2015) Corrective or confirmative actions? Political online participation as a consequence of presumed media influences in election campaigns. *Journal of Information Technology & Politics* 12: 285–302.
- Bimber B, Flanagin AJ and Stohl C (2005) Reconceptualizing collective action in the contemporary media environment. *Communication Theory* 15: 365–388.

- Blair CA, Foster Thompson L and Wuensch KL (2005) Electronic helping behavior: The virtual presence of others makes a difference. *Basic and Applied Social Psychology* 27: 171–178.
- Bruning R, Dempsey M, Kauffman DF, et al. (2013) Examining dimensions of self-efficacy for writing. *Journal of educational psychology* 105: 25–38.
- Brunsting S and Postmes T (2002) Social movement participation in the digital age. *Small Group Research* 33: 525–554.
- Burn SM (2009). A situational model of sexual assault prevention through bystander intervention. *Sex Roles* 60: 779–792.
- Chekroun P and Brauer M (2002) The bystander effect and social control behavior: the effect of the presence of others on people's reactions to norm violations. *European Journal of Social Psychology* 32: 853–867.
- Chen GM (2017) *Online incivility and public debate: Nasty talk*. Cham: Springer International Publishing.
- Coe K, Kenski K and Rains SA (2014) Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *Journal of Communication* 64: 658–679.
- Cramer RE, McMaster MR, Bartell PA, et al. (1988) Subject competence and minimization of the bystander effect. *Journal of Applied Social Psychology* 18: 1133–1148.
- Dovidio JF, Gaertner SL, Validzic A, et al. (1997) Extending the benefits of recategorization: Evaluations, self-disclosure, and helping. *Journal of Experimental Social Psychology* 33: 401–420.
- Feldman L, Hart PS, Leiserowitz A., et al. (2017) Do hostile media perceptions lead to action? The role of hostile media perceptions, political efficacy, and ideology in predicting climate change activism. *Communication Research* 44: 1099-1124.

- Fischer P, Greitemeyer T, Pollozek F, et al. (2006) The unresponsive bystander: are bystanders more responsive in dangerous emergencies? *European Journal of Social Psychology* 36: 267–278.
- Fischer P, Krueger JI, Greitemeyer T, et al. (2011) The bystander-effect: a meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin* 137: 517–537.
- Flew T, Martin F and Suzor N (2019) Internet regulation as media policy: Rethinking the question of digital communication platform governance. *Journal of Digital Media & Policy* 10: 33–50.
- Friess D and Eilders C (2015). A systematic review of online deliberation research. *Policy & Internet* 7: 319–339.
- Gil de Zúñiga H, Jung N and Valenzuela S (2012) Social media use for news and individuals' social capital, civic engagement and political participation. *Journal of Computer-Mediated Communication* 17: 319–336.
- Golan GJ and Lim JS (2016) Third person effect of ISIS's recruitment propaganda: Online political self-efficacy and social media activism. *International Journal of Communication* 10: 4681–4701.
- Guldberg K and Mackness J (2009) Foundations of communities of practice: Enablers and barriers to participation. *Journal of computer assisted learning* 25: 528–538.
- Herbst S (2010) *Rude democracy: Civility and incivility in American politics*. Philadelphia, PA: Temple University Press.
- Hsueh M, Yogeewaran K and Malinen S (2015) “Leave your comment below”: Can biased online comments influence our own prejudicial attitudes and behaviors? *Human Communication Research* 41: 557–576.

- Jung N, Kim Y and De Zúñiga, HG (2011). The mediating role of knowledge and efficacy in the effects of communication on political participation. *Mass Communication and Society* 14: 407–430.
- Kalch A and Naab TK (2017) Replying, disliking, flagging: How users engage with uncivil and impolite comments on news sites. *Studies in Communication and Media* 6: 395–419.
- Karpf D (2010) Online political mobilization from the advocacy group's perspective: Looking beyond clicktivism. *Policy & Internet* 2: 7–41.
- Kelly C and Kelly J (1994) Who gets involved in collective action? Social psychological determinants of individual participation in trade unions. *Human Relations* 47: 63–88.
- Kenski K, Coe K and Rains SA (2017) Perceptions of uncivil discourse online: An examination of types and predictors. *Communication Research: 1–20*
- Kenski K and Stroud NJ (2006) Connections between internet use and political efficacy, knowledge, and participation. *Journal of Broadcasting & Electronic Media* 50: 173–192.
- Klandermans PG (1984) Mobilization and participation in trade union action: An expectancy-value approach. *Journal of Occupational Psychology* 57: 107–120.
- Kreißel P, Ebner J, Urban A et al. (2018). *Hass auf Knopfdruck [Hate at the touch of a button]*. Institute for Strategic Dialogue.
- Latané B and Darley JM (1970) *The unresponsive bystander: Why doesn't he help?* Englewood Cliffs, NJ: Prentice-Hall.
- Latané B and Nida S (1981) Ten years of research on group size and helping. *Psychological Bulletin* 89: 308–324.
- Ley H (2018) *#ichbinhier: Zusammen gegen Fake News und Hass im Netz: [#iamhere: Together against fake news and hate in the net]*. Köln: DuMont.

- Lim JS (2017) The third-person effect of online advertising of cosmetic surgery: A path model for predicting restrictive versus corrective actions. *Journalism & Mass Communication Quarterly* 94: 972–993.
- Lim JS and Golan GJ (2011) Social media activism in response to the influence of political parody videos on YouTube. *Communication Research* 38: 710–727.
- McLeod JM, Scheufele DA and Moy P (1999) Community, communication, and participation: The role of mass media and interpersonal discussion in local political participation. *Political Communication* 16: 315–336.
- Naab TK, Kalch A and Meitz, TG (2018). Flagging uncivil user comments: Effects of intervention information, type of victim, and response comments on bystander behavior. *New Media & Society* 20: 777–795.
- Newman N, Fletcher R, Kalogeropoulos A, Levy DAL, Nielsen, RK (2018) *Reuters Institute Digital News Report 2018*. Reuters Institute for the Study of Journalism.
- Obermaier M, Fawzi N and Koch T (2016) Bystanding or standing by? How the number of bystanders affects the intention to intervene in cyberbullying. *New Media & Society* 18: 1491–1507.
- Palasinski M (2012) The roles of monitoring and cyberbystanders in reducing sexual abuse. *Computers in Human Behavior* 28: 2014–2022.
- Postmes T and Brunsting S (2002) Collective action in the age of the internet. *Social Science Computer Review* 20(3): 290–301.
- Rojas H (2010) “Corrective” actions in the public sphere: How perceptions of media and media effects shape political behaviors. *Journal of Personality and Social Psychology* 22: 343–363.
- Rösner L and Krämer NC (2016) Verbal venting in the social web: Effects of anonymity and group norms on aggressive language use in online comments. *Social Media+Society* 2: 1–13.

- Schwartz SH and Gottlieb A (1980) Bystander anonymity and reactions to emergencies. *Journal of Personality and Social Psychology* 39: 418–430.
- Simon B, Loewy M, Stürmer S, et al. (1998) Collective identification and social movement participation. *Journal of personality and social psychology* 74: 646–658.
- Solomon LZ, Solomon H and Stone R (1978) Helping as a function of number of bystanders and ambiguity of emergency. *Personality & social psychology bulletin* 4: 318–321.
- Springer N, Engelmann I and Pfaffinger C (2015) User comments: motives and inhibitors to write and read. *Information, Communication & Society* 18: 798–815.
- Stroud NJ, Scacco JM, Muddiman A, et al. (2015) Changing deliberative norms on news organizations' Facebook sites. *Journal of Computer-Mediated Communication* 20: 188–203.
- Stroud NJ, van Duyn E and Peacock C (2016) *News commenters and news comment readers*. Available at: <https://mediaengagement.org>.
- Thornberg R (2007) A classmate in distress: schoolchildren as bystanders and their reasons for how they act. *Social Psychology of Education* 10: 5–28.
- van Bommel M, van Prooijen JW, Elffers H, et al. (2012) Be aware to care: Public self-awareness leads to a reversal of the bystander effect. *Journal of Experimental Social Psychology* 48: 926–930.
- Van Zomeren M, Spears R, Fischer AH and Leach CW (2004) Put your money where your mouth is! Explaining collective action tendencies through group-based anger and group efficacy. *Journal of personality and social psychology* 87: 649–664.
- van Zomeren M, Postmes T and Spears R (2008) Toward an integrative social identity model of collective action: a quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin* 134: 504–535.

- Velasquez A and LaRose R (2015). Social media for social change: Social media political efficacy and activism in student activist groups. *Journal of Broadcasting & Electronic Media* 59: 456–474.
- Ziegele M (2016) *Nutzerkommentare als Anschlusskommunikation. Theorie und qualitative Analyse des Diskussionswerts von Online-Nachrichten [User comments as media-stimulated interpersonal communication: Theory and qualitative analysis of the discussion value of online news]*. Wiesbaden: Springer VS.
- Ziegele M, Koehler C and Weber M (2018a) Socially destructive? Effects of negative and hateful user comments on readers' donation behavior toward refugees and homeless persons. *Journal of Broadcasting & Electronic Media* 62: 636-653.
- Ziegele M, Weber M, Quiring O, et al. (2018b) The dynamics of online news discussions: effects of news articles and reader comments on users' involvement, willingness to participate, and the civility of their contributions. *Information, Communication & Society* 21: 1419–1435.

Tables

Table 1

Hierarchical Multiple Regression Analysis Predicting Likelihood of Engagement With Uncivil Comments

	Model 1: Attention (+ controls)	Model 2: Situation as emergency	Model 3: Taking responsib.	Model 4: Ability to intervene	Model 5: Costs and benefits	Model 6: Collective factors
	β	β	β	β	β	β
Controls						
Sex (1 = female)	-.07	-.07	-.06	-.02	-.04	-.05
Formal education (1 = high)	-.07	-.07	-.05	-.12**	-.10**	-.11**
Age	.26***	.26***	.26***	.22***	.22***	.22***
Step 1: Paying attention						
Political Facebook usage	.22***	.22***	.22***	.19***	.11**	.09**
Step 2: Construing the situation as an emergency						
Perceived threat of comments		-.06	-.05	-.04	-.05	-.05
Step 3: Feeling responsible						
Perceived personal responsibility			.15***	.13**	.11**	.11**
Step 4: Feeling able to intervene						
Internal political efficacy				.08 [†]	.06	.05
Perceived writing self-efficacy				.30***	.27***	.25***
Step 5: Costs and benefits						
Perceived personal benefits					.27***	.20***
Perceived risks					.14***	.13***
Collective factors						
Collective identity #ibh						-.01
Perceived group effectiveness						.10**
Perceived collective benefits						.06
Perceived knowledge about #ibh						.11**
<i>N</i>	545	545	545	545	545	545
<i>R</i> ²	.14	.14	.16	.26	.35	.38
ΔR^2	-	.00	.02	.11	.09	.03
Sig. ΔR^2	***	ns.	***	***	***	***

Notes. $n = 545$, * $p < .05$, ** $p < .01$, *** $p < .001$