

**Pabian, S.** (2022). *Defending my online friend: Investigating bystander reactions on influencer bashing among female Instagram users*. Paper presented at the Annual Conference of the International Communication Association, Paris, France, May 26-30 2022.

### **Abstract**

Social media influencers often become the target of ridiculing or insulting messages with the intention to hurt them, also called influencer bashing. Supportive bystander behavior performed by followers, such as sending public support messages or reporting the hateful comments, has the potential to diminish or prevent negative outcomes for the influencer. The goal of the present study was to investigate antecedents of publicly and privately supportive bystander behaviors, as well as, publically and privately inappropriate bystander reactions, such as joining in with the aggressor. To this aim, an online survey was administered among 234 female Instagram users aged 18 to 25. The results indicated that perceived similarity between follower and influencer and state empathy elicited by an influencer bashing incident explain intentions to perform supportive bystander behavior, whereas trait empathy and severity of the influencer bashing incident seem not related. More precisely, a high level of perceived similarity was associated with a high level of state empathy, which in turn was associated with greater intentions to perform public and private supportive bystander behavior (fully mediated relationships). Implications of these findings and limitations are discussed.

*Keywords:* Influencers; online aggression; bashing; bystander; supporting; antecedents.

**Defending my online friend: Investigating bystander reactions on influencer bashing  
among female Instagram users**

Being a social media influencer is something that many youngsters aspire. Recent research among 13-38 year-olds living in the US indicates that more than 50 percent of these young adolescents and adults want to be a social media influencer (MorningConsult, 2019) and also in the United Kingdom social media influencer is the second most popular profession among young people (Skeldon, 2019). Social media influencers are individuals that have a large reach on social media platforms on which they share their interests and aspects of their daily lives by sharing content that can be perceived as authentic (De Veirman et al., 2017; Schouten et al., 2020). Social media influencers have close relationships with their followers and seem successful in influencing their audience, which makes them highly valuable collaborators for organizations that are looking for ways to reach a certain audience and to convince them of certain ideas or products, also called influencer marketing (Lim et al., 2017). Although collaborating with companies and being admired by a large audience seem for many young individuals a dream, the social media influencer life also has aspects that are less attractive (Abidin, 2016).

In spite of the positive feedback that influencers often receive from their followers, recent research has indicated that many social media influencers become also the targets of derogatory messages (Abidin, 2016; Hassan et al., 2018). This behavior has been called in the context of celebrities ‘celebrity bashing’ (Ouvrein, Backer, et al., 2018; Rudnicki et al., 2020). It consists of ridiculing or insulting the target with the intention to hurt, for instance by posting a derogatory or mean comment about the target’s appearance, work, or private life, or by threatening the target (Ouvrein, Pabian, et al., 2018). In the current study, we will further refer to this behavior in the context of influencers by using the term ‘influencer bashing.’ Influencer bashing has not only caught the attention of researchers (e.g., Hassan et al., 2018),

also (emotional) testimonials of social media influencers have been shared in (international) media outlets. For instance, Famke Louise, a Dutch vlogger and singer, has addressed this issue in a documentary by reading out loud negative reactions she receives daily, such as ‘Please kill yourself’ and ‘You are an absolute loser, I hope you are going to die’ (Tratlehner, 2018).

Influencer bashing has the potential to have a large impact on the social media influencer. Research indicates that also public figures, who are expected to have a ‘thick skin’ (Ouvrein et al., 2017), can suffer from these negative online comments (Ouvrein et al., 2021). Examples of potential negative outcomes are related to poor mental health and well-being and externalizing behaviors, such as depressions, excessive alcohol use, or drug addiction (Dalla Pozza et al., 2011; Ouvrein, Pabian, et al., 2018). Research on online aggression indicates that interventions of bystanders can be successful in terms of diminishing or ending the online aggression and/or reducing the harmful effects (Salmivalli, 2010; Wang, 2021). For instance, bystanders can support the victim by sending a supportive message or can ask the perpetrator to stop behaving aggressively (DeSmet et al., 2019). However, bystanders can also decide to do nothing or to support the perpetrator. These behaviors can amplify the negative outcomes for the victims (O’Connell et al., 1999; Wang, 2021). Until now, bystander behavior has not been studied yet in the context of influencer bashing. The goal of the present study is to investigate antecedents of bystander behavior when witnessing influencer bashing.

The research fields on online aggression towards peers and celebrities can provide important insights on different types of bystander behavior and drivers of performing bystander behavior within the influencer context (Ouvrein, De Backer, et al., 2018; Van Cleemput et al., 2014). In what follows, an overview of the relevant literature on bystander

behaviors and drivers of these behaviors is given. Also differences between influencer bashing, celebrity bashing, and online peer aggression are considered.

### **Types of Bystander Behavior**

Researchers have indicated that there are three broad types of bystander behaviors when witnessing online aggression: doing nothing, joining in, and helping the victim (Van Cleemput et al., 2014). The latter is perceived as positive or supportive bystander behavior (DeSmet et al., 2019). Examples of positive bystander behavior are comforting, giving advice, reporting, and confronting the bully (DeSmet et al., 2018). In the context of influencer bashing, positive bystander behavior can consist, for instance, of reporting the hurtful comment to the social media platform, sending a supportive message to the influencer, or sending a message to the perpetrator to ask to stop bashing. One can also decide to perform negative bystander behavior such as reinforcing the negative incidents or joining in, for instance, by liking the influencer bashing incident or by also posting hurtful comments directed towards the influencer. Studies on online bystander behavior have further distinguished specific bystander behaviors, for instance by focusing on how visible an action is to others, i.e. publicness (Wang, 2021). Social media platforms provide contexts in which bystanders can offer help in either a publicly visible or a privately manner. Research indicates that bystanders often prefer to react privately because this is perceived as less risky and, thus, more safe (Dillon & Bushman, 2015; Patterson et al., 2017), whereas there might be a chance that the bystander becomes the new target of the online aggressor if the bystander reacts in a publicly visible way (DeSmet et al., 2014). In order to convince current and future bystanders to perform positive bystander behavior, behavior change theories and evidence-based intervention and prevention programs focusing on online aggression have indicated that it is crucial to understand determinants of bystander behavior (DeSmet et al., 2016, 2018).

### **Antecedents of Bystander Reactions: Empathy**

A large number of determinants have been proposed to understand supportive online bystander behavior in the peer context, relying on theoretical models such as the bystander effect theory (Latané & Darley, 1970), Social Cognitive Theory (Bandura, 1986), Moral Disengagement theory (Bandura et al., 1996), and the Social Ecological Model (Lambe et al., 2019). A recent meta-analysis (Torgal et al., 2021) on online bystander behavior towards peers indicated that of the several personality factors that have been linked to supportive bystander behavior, empathy consistently emerges as an important predictor (e.g., Lambe et al., 2019; Machackova & Pfetsch, 2016; Van Cleemput et al., 2014). Empathy has been defined as sharing the emotional state of another person through taking the perspective of the other and understanding his or her emotions (Davis & Franzoi, 1991; Eisenberg & Strayer, 1987). Positive associations have been found with supportive bystander behavior: Individuals who reported higher levels of empathy were found to be more likely to perform supportive bystander behavior when they witnessed online aggression (e.g., Lambe et al., 2019). In contrast, low levels of empathy have been found to be associated with negative bystander behavior (e.g., Van Cleemput et al., 2014).

Some studies have included and compared different types of empathy, such as cognitive empathy and affective empathy (e.g., Machackova & Pfetsch, 2016). Cognitive empathy is related to understanding the internal state of another individual, whereas affective empathy is related to reacting emotionally on the observed experiences of another individual (Davis & Franzoi, 1991). For instance, in their study among German adolescents, Macháčková and Pfetsch (2016) found that affective empathy was a significant positive correlate of supportive bystander behavior when witnessing online aggression, whereas cognitive empathy was not associated.

Besides the distinction between cognitive and affective empathy, researchers have indicated the difference between empathy as a character trait (trait empathy) and empathy as a situational state (state empathy) (Van der Graaff et al., 2016). The earlier mentioned studies included trait empathy, which refers to the general ability to show empathy. The second type, state empathy, refers to an empathic response in a concrete situation. Also state empathy has been found to be associated with bystander behavior: individuals scoring higher on state empathy, elicited by witnessing the online aggressive incident, are more inclined to perform supportive bystander behavior, both publicly and privately (Wang, 2021), and are less likely to perform negative bystander behavior (Barlińska et al., 2013). One's state empathy depends on the situation and therefore previous research has linked this form of empathy to other characteristics of the online aggressive incidents, including similarity between oneself and the victim and the severity of the incident (Wang, 2021).

### **Perceived Similarity and Bystander Behavior**

Perceived similarity in the context of bystander behavior can be described as the perceived social distance between the victim and the bystander that is established by social comparison: The more similar bystanders perceive the victim with themselves, for instance based on beliefs, attitudes, personality and/or other characteristics, the shorter the perceived social distance (Liviatan et al., 2008; Stets & Burke, 2000; Wang, 2021). Research on online peer aggression and bystander behavior has indicated a positive association between similarity and supportive bystander behavior (DeSmet et al., 2016; Wang, 2021). DeSmet et al. (2016) explained this association by referring to the greater sense of responsibility for the welfare of the victim that bystanders feel when they feel a close connection between themselves and the victim. These studies have been conducted within the peer context and research indicates that the social distance between a celebrity and a bystander is much larger compared to the social distance that one perceives between oneself and a peer (Ouvrein, De

Backer, et al., 2018; Ouvrein et al., 2020). The social distance with an influencer might be much smaller than the perceived social distance with a celebrity. In contrast to celebrities, influencers try to connect with their audience and they try to present themselves as ‘normal’, identifiable, and approachable individuals (Chapple & Cownie, 2017; Djafarova & Rushworth, 2017; Schouten et al., 2020). In the study of Schouten et al. (2020) among young female adults, respondents perceived themselves as being more similar to influencers than to celebrities. Influencers are often perceived as role models or as a big brother or sister, whereas celebrities are often perceived as being unapproachable (Djafarova & Matson, 2021; Djafarova & Rushworth, 2017; Lou & Kim, 2019). This can be explained by the communication style that influencers use, which include strategies to engage with their audience. Also within groups of followers there might be differences in perceived similarity with the influencer and this might be associated with one’s reaction on influencer bashing as a bystander.

### **Severity of the Incidents and Bystander Behavior**

Bystander reactions have also been related to the severity of the incident that the bystander witnesses (Macaulay et al., 2019). Research on online aggression towards peers indicates a positive relationship between severity and supportive bystander behavior: Bystanders are motivated to support the victim when the incident is being perceived as more severe (Bastiaensens et al., 2014). Potential explanations might be that people feel more responsible to support the victim when witnessing a more severe incident compared to a mild incident (Obermaier et al., 2016; Wang, 2021). In addition, a bystander of a more severe incident who decides to not intervene may experience greater shame for not intervening in a positive manner compared to a bystander of a mild incident who decides to not perform positive bystander behavior (Macaulay et al., 2019). Although the likelihood to support the victim is higher when witnessing a more severe incident, bystanders perceive a more severe

incident as less safe to intervene compared to mild forms (Bastiaensens et al., 2014; Bennett & Banyard, 2016; Wang, 2021). Wang (2021) notes that this can influence their choice to react privately instead of publicly as they want to limit the risk for negative outcomes for themselves.

Mild forms of influencer bashing can include criticizing an influencer's appearance, work (e.g., social media posts), or private life, saying mean things, insulting, making fun of an influencer, or sharing a negative news article about the influencer (based on mild forms of celebrity bashing, proposed by Ouvrein, De Backer, et al., 2018). Severe forms of influencer bashing can consist of practices such as modifying online content of an influencer in a negative way and threatening the influencer (based on severe forms of celebrity bashing, proposed by Ouvrein, De Backer, et al., 2018).

### **Relationships Between Antecedents of Bystander Reactions and The Present Study**

Besides the direct associations that have been found between empathy and bystander behavior, similarity and bystander behavior, and severity of the incident and bystander behavior in the context of online peer aggression, researchers have also suggested relationships between some of these antecedents of bystander behavior. Multiple studies have indicated that similarity and state empathy are related to each other in the context of prosocial behavior, including helping: Perceived similarity seems to be associated with higher levels of state empathy, which, in turn, seems to be related to higher levels of prosocial behavior (e.g., Cialdini et al., 1997) This association between similarity and state empathy can be explained by the in-group empathy hypothesis (Brown et al., 2006) or the intergroup empathy bias (Fourie et al., 2017). According to these frameworks, in-group members that are in danger or that are being attacked elicit higher levels of state empathy among bystanders compared to outgroup members (Brown et al., 2006; Fourie et al., 2017). Individuals who are perceived as highly similar are considered to be in-group members. The limited or lack of state empathy



towards outgroup members is also called outgroup empathy failures. The intergroup empathy bias even indicates that outgroup empathy failures do not depend on a person's trait empathy: Even the most deeply empathic individuals can mute their state empathy towards a perceived enemy or outgroup member (Fourie et al., 2017).

Similarity and state empathy have also been linked to the severity of the incidents that are witnessed. For instance, in the study of Wang among adult Facebook users (2021), an interaction effect between severity of the online aggressive incident and perceived similarity had an impact on empathy. More precisely, if online aggressive incidents were performed towards dissimilar victims, severe incidents elicited more state empathy than mild incidents. If online aggressive incidents were performed towards similar victims, both mild and severe forms of online aggression elicited rather high levels of state empathy.

The present study will investigate antecedents of intentions to perform bystander behavior when witnessing influencer bashing. Three potential antecedents were selected based on the literature: empathy, perceived similarity, and severity of the influencer bashing incidents. Although the literature on online peer aggression and celebrity bashing provides inputs for the potential importance of these antecedents in explaining influencer bashing bystander behavior, no expectations were formulated given that the present study focuses on a different form of online aggression. The research question for the present study was: How are empathy, perceived similarity, and severity associated with intentions to perform bystander behavior among young female Instagram users? The presented study was limited to female Instagram users as the majority of influencers tailor to a female audience (Gannon & Prothero, 2018; Schouten et al., 2020) and women between 18 to 25 are most frequently active on Instagram compared to males and older adults (Gazit et al., 2019; McComb et al., 2021; Van der Veer et al., 2021). Moreover, there are indications in the literature regarding

differences between female and male social media users in perceived similarity towards female and male influencers (Gräve, 2017; Hoeghele et al., 2016).

## Method

### Respondents & Procedure

In order to answer the research question, an online survey was administered among young female Instagram users aged 18 to 25 in April 2021. The sample was a convenience sample, consisting of Bachelor and Master students of [University blinded for review]. A total of 240 respondents completed the survey. Six respondents were removed from the dataset during the data cleaning process, because of missing values or outliers. The analytic sample thus consisted of  $N = 234$  young females with a mean age of 21.36 ( $SD = 2.11$ ). The majority had the Dutch nationality (97%,  $N = 227$ ), other nationalities were Belgian (.9%,  $N = 2$ ), Syrian (.9%,  $N = 2$ ), Indonesian (.4%,  $N = 1$ ), Polish (.4%,  $N = 1$ ), and Spanish (.4%,  $N = 1$ ). About half of the sample's highest degree of education (50.8%,  $N = 119$ ) was 'higher education' (college or university). All respondents were Instagram users and almost all of them used Instagram daily (95.7%,  $N = 224$ ). The majority of the respondents followed 1 to 25 (59.4%,  $N = 139$ ) or 25 to 50 (23.9%,  $N = 56$ ) influencers. Nine out of ten respondents indicated that they (almost) never post a reaction under Instagram posts of influencers (91.0%,  $N = 213$ ).

The study followed APA Ethical Guidelines for research with human subjects and the advices of the ethical advisory board of [Faculty and University blinded for review]. Informed consent was obtained from all respondents. After providing informed consent, respondents first filled in a block of questions to measure socio-demographic questions, as well as questions to measure trait empathy. In a next block, they were introduced to a Dutch Instagram influencer by means of pictures and a short text and respondents were asked to rate their perceived similarity with the influencer. Next, respondents were exposed, ad random, to

two examples of either mild or severe influencer bashing incidents performed towards the influencer. More precisely,  $N = 114$  respondents were shown two examples of mild incidents, whereas  $N = 120$  respondents saw two examples of severe incidents. Respondents indicated for each of the two examples their affective state empathy, their (general) intention to take action, and their intentions to perform specific types of bystander behavior. After these questions, respondents reached the end of the survey and received a debriefing which included a warning that some of the incidents were fictional and created, based on previous research, for this study. The data and materials of the present study are available on the Open Science Framework [url blinded for review].

### **Materials & Measures**

In order to investigate antecedents of and intentions to perform bystander behavior, respondents were exposed to two examples of influencer bashing, either mild or severe incidents. These (fictional) incidents were shown under two (real) Instagram posts of the Dutch female lifestyle influencer Juultje Tieleman (@juultjetieleman). The influencer had 932.000 followers at the moment of administering the survey and is in the same age group as the respondents of the present study. A pre-test among  $N = 23$  young female Instagram users ( $M_{\text{age}} = 22.0$ ,  $SD = 1.84$ ) indicated that this age group is, on average, slightly familiar with the influencer and does not have an extreme negative or positive attitude towards the influencer. More precisely, respondents of the pre-test indicated, on average,  $M = 3.35$  ( $SD = 2.31$ ), to be somewhat familiar with the influencer (single-item measurement with answer options ranging from 1 to 7; Wei et al., 2008). Attitudes towards these influencer are, on average, neutral to slightly positive,  $M = 4.61$ ,  $SD = 1.09$  (measured with four items on a scale ranging from 1 to 7; Reliability of the scale:  $\alpha = .66$ ; adapted from Dimofte et al., 2003). Similar scores for familiarity and attitude were found in the analytical sample (familiarity:  $M = 4.11$ ,  $SD = 2.08$ ; attitude:  $M = 4.45$ ,  $SD = 1.12$ ). Besides familiarity with and attitude towards the influencer,

the pre-test also tested the perceived seriousness of the influencer bashing incidents. The influencer bashing incidents were created based on the study of Ouvrein et al. (2018). Respondents of the pretest rated nine influencer bashing incidents from 1 (*mild*) to 10 (*severe*) and based on these scores two mild and two severe incidents were selected.

The two mild influencer bashing incidents that were selected based on the pre-test scores were the following negative comments towards the influencer ‘You are so fake’ and ‘You are such a bad influencer’, which are, according to Ouvrein et al. (2018) examples of ‘saying mean things to a celebrity’ and ‘criticizing the performance of a celebrity’. The two severe negative comments that were selected were ‘I’ll be waiting for you at your house soon’ and ‘If you don’t respond, I will share pictures of you that you don’t want to see all over the internet’, these are, according to Ouvrein et al. (2018), examples of threats towards the celebrity.

Before analyzing the data, a manipulation check was performed to test whether the used influencer bashing incidents were interpreted as intended (mild or severe) by the analytical sample. Respondents were asked at the end of the survey to indicate on a slider bar, ranging from 1 (*mild*) to 10 (*severe*) whether they considered the Instagram post they had seen as a mild or severe incident. Respondents who were exposed to mild bashing incidents ( $N = 114$ ) rated the incidents they had seen as somewhat mild ( $M_{Mild\ incident\ 1} = 4.76$ ,  $SD_{Mild\ incident\ 1} = 2.13$ ;  $M_{Mild\ incident\ 2} = 4.77$ ,  $SD_{Mild\ incident\ 2} = 2.23$ ), whereas respondents who saw severe influencer bashing incidents ( $N = 120$ ) rated the two incidents they had seen as rather severe ( $M_{Severe\ incident\ 1} = 6.10$ ,  $SD_{Severe\ incident\ 1} = 2.08$ ;  $M_{Severe\ incident\ 2} = 6.86$ ,  $SD_{Severe\ incident\ 2} = 2.14$ ), and, therefore, the manipulation was considered to be successful.

The following scales were used to measure perceived similarity, trait empathy, state empathy, and intention to perform bystander behavior.

**Perceived similarity.** In order to measure perceived similarity with the influencer, four items were assessed with answer options ranging from 1 (*totally disagree*) to 7 (*totally agree*), based on Hoffner and Buchanan (2005). An example of an item is: ‘This influencer thinks in the same way as I do’. For further analyses a mean score for each respondents was calculated. The internal consistency (Cronbach’s alpha) was .87.

**Affective trait empathy.** The Basic Empathy Scale (Jolliffe & Farrington, 2006) was used to measure affective trait empathy. Eleven items were administered, with answer options ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). An example of an item is ‘After being with a friend who is sad about something, I usually feel sad.’ A mean score was calculated for each respondent ( $\alpha = .74$ ).

**Affective state empathy.** A scale consisting of five items was used to measure affective state empathy after exposure to each of the two influencer bashing incidents (either mild or severe influencer bashing incidents) (based on Davis, 1983). The five items were: After seeing this, I feel: ‘worried’, ‘uncomfortable’, ‘sad’, ‘disturbed’, and ‘upset’. Answer options ranged from 1 (*not at all*) to 5 (*very much*). A mean score was calculated based on the answers for the two influencer bashing incidents respondents were exposed to ( $\alpha = .92$ ).

**Intention to perform bystander behavior.** First, the general intention to take action was measured with a single item: ‘What is the probability you would react after seeing this comment?’, on a seven-point Likert scale ranging from 1 (*very unlikely*) to 7 (*very likely*) (based on Bastiaensens et al., 2014). This question was administered twice, after exposure to each example of either mild influencer bashing or either severe influencer bashing ( $\alpha = .68$ ). Next, intentions to perform specific forms of bystander behavior were measured. Based on existing measurements used to measure bystanders’ response towards online peer aggression and celebrities bashing (Bastiaensens et al., 2014; DeSmet et al., 2019; Ouvrein, Backer, et

al., 2018) both supportive bystander behavior and negative bystander behavior were included and a differentiation was made between publicly visible forms of bystander behavior and private (non-publicly visible) forms of bystander behavior, which resulted in four types of bystander behavior. All types were administered twice, after exposure to the two examples of either mild or severe influencer bashing, with a seven-point Likert scale ranging from 1 (*very unlikely*) to 7 (*very likely*). Public supportive bystander reactions were measured with two items, consisting of 'Leaving a comment to support the victim' and 'Leaving a comment to point out the inappropriate behavior of the perpetrator' ( $\alpha = .88$ ). Private supportive bystander reactions were measured with four items: 'Reporting the hurtful message of the perpetrator', 'Reporting the profile of the perpetrator', 'Sending a private message to the influencer to support her', and 'Sending a private message to the perpetrator to point out the inappropriateness of his/her behavior' ( $\alpha = .88$ ). Public negative bystander reactions were measured with one item, 'Leaving a publically visible comment to support the perpetrator' ( $\alpha = .81$ ). Finally, private negative bystander reactions were measured with one item, 'Sending a private message to the perpetrator to support him/her' ( $\alpha = .86$ ).

### **Data-analysis**

To analyze the data, first descriptive and bivariate statistics were calculated. In order to test the associations between the potential antecedents and intention to perform (specific forms of) bystander behavior, bootstrapped moderation (model 1), mediation (model 4), double moderation (model 3), and moderated mediation (model 7) analyses were performed by using Hayes' Process Macro for SPSS (Hayes, 2017) in SPSS version 28. Moderation models (see Figure 1) were calculated to test the associations between perceived similarity, trait empathy, and bystander behavior intentions, more precisely, intention to (1) take action (general intention to perform bystander behavior), (2) perform public supportive bystander behavior, (3) perform private supportive bystander behavior, (4) perform public negative

bystander behavior, and (5) perform private negative bystander behavior. In order to test similar associations for state empathy, mediation analyses were performed as state empathy might be influenced by perceived similarity (see Figure 2). Next, associations between the severity of the influencer bashing incident, trait empathy, and bystander behavior intentions were calculated by means of double moderation models (see Figure 3). In a final step, moderated mediation models were calculated to see whether the severity of the influencer bashing incident has an effect on the indirect relationship between perceived similarity and bystander behavior intention via state empathy (see Figure 4). Bias-corrected 95% confidence intervals based on 5,000 bootstrap samples were used to evaluate the effects. Significant effects are supported by the absence of zero within the confidence intervals.

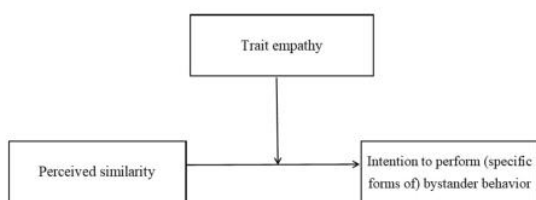


Figure 1. Tested moderation models

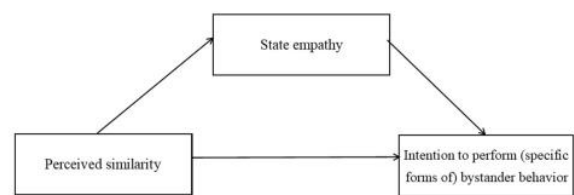


Figure 2. Tested mediation models

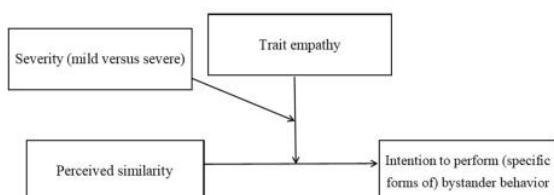


Figure 3. Tested double moderation models

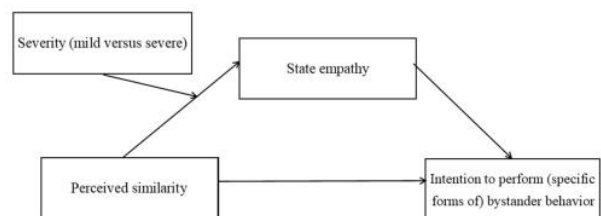


Figure 4. Tested moderated mediation models

## Results

### Descriptive and Bivariate Statistics

The descriptive statistics of the main variables (Table 1) indicated that, on average, respondents had a very low intention to take action ( $M = 1.49$ ,  $SD = .84$ , range 1-7). The mean scores of the specific types of bystander behavior indicated that if respondents would take action, they would rather perform private supportive bystander behavior ( $M = 2.18$ ,  $SD =$

1.15, range 1-7), followed by public supportive bystander behavior ( $M = 1.73$ ,  $SD = .96$ , range 1-7), public negative bystander behavior ( $M = 1.24$ ,  $SD = .67$ , range 1-7), and private negative bystander behavior ( $M = 1.24$ ,  $SD = .65$ , range 1-7). Bootstrapped paired samples  $t$ -tests (see Appendix A) indicated that all types of bystander behavior differed from each other in terms of intentions to perform, except for the two types of negative bystander behavior (public and private). Furthermore, the descriptives indicated that respondents perceived rather low similarity between themselves and the influencer ( $M = 2.77$ ,  $SD = 1.08$ , range 1-7) and they scored below the midpoint on state empathy ( $M = 2.26$ ,  $SD = .86$ , range 1-5). Finally, the average score on trait empathy indicated that respondents are somewhat empathic ( $M = 3.64$ ,  $SD = .46$ , range 1-5).

**Table 1**

*Spearman correlation coefficients and descriptive statistics of the main variables of the present study*

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Severity of the influencer bashing incidents	-								
2. Intention to take action	.13	-							
3. Public supportive bystander behavior	-.04	.62***	-						
4. Private supportive bystander behavior	.36***	.59***	.57***	-					
5. Public negative bystander behavior	.02	.40***	.43***	.25***	-				
6. Private negative bystander behavior	.07	.42***	.39***	.28***	.90***	-			
7. Similarity	.13	.06	.14*	.20**	-.05	-.05	-		
8. State empathy	.24***	.43***	.43***	.56***	.16*	.18**	.25***	-	
9. Trait empathy	-.11	.05	.13	.11	-.05	-.10	.11	.19**	-
Mean	/	1.49	1.73	2.18	1.24	1.24	2.77	2.26	3.64
SD	/	.84	.96	1.15	.67	.65	1.08	.86	.46
Range	/	1-7	1-7	1-7	1-7	1-7	1-7	1-5	1-5

Note. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

Spearman correlation coefficients (Table 1) indicated moderate to strong correlations between the different forms of bystander behavior. Perceived similarity correlated only significantly with two specific types of bystander behavior, public supportive bystander behavior ( $r = .14$ ,  $p < .05$ ) and private supportive bystander behavior ( $r = .20$ ,  $p < .01$ ). State



empathy had a significant positive correlation with the overall intention to take action ( $r = .24, p < .01$ ) and the intention to perform all specific types of bystander behavior. Trait empathy was not related to intention to take action, nor to intentions to perform specific types of bystander behavior.

Bootstrapped independent  $t$ -tests were performed to discover significant differences between the two types (severity) of influencer bashing, mild and severe, and their average score on overall intention to take action and their average scores on intentions to perform specific bystander behaviors. Significant differences between being exposed to mild versus severe influencer bashing incidents were only found for private supportive bystander behavior,  $t(232) = -5.29, p < .001$ . Respondents who were exposed to severe forms of influencer bashing had a higher intention to perform private supportive bystander behavior ( $M = 2.55, SD = 1.11$ ), compared to respondents who were exposed to mild forms of influencer bashing ( $M = 1.80, SD = 1.06$ ). In addition, a bootstrapped independent  $t$ -test was also performed to see whether there were differences in respondents' state empathy. Those who were exposed to severe forms of influencer bashing scored significantly higher on state empathy ( $M = 2.46, SD = .81$ ) compared to those who were exposed to mild forms of influencer bashing ( $M = 2.05, SD = .87$ ),  $t(232) = -3.67, p < .001$ . Finally, two bootstrapped independent  $t$ -tests showed there were no differences between those who were exposed to mild influencer bashing incidents versus those who were exposed to severe influencer bashing incidents with regard to perceived similarity and trait empathy. This means that respondents were successfully randomly distributed to either mild or severe examples of influencer bashing for these two independent variables.

### **Understanding Bystander Behavior Intentions: Perceived Similarity & Empathy**

First, bootstrapped moderation models for intention to take action and for each of the specific types of bystander behaviors as dependent variables were calculated with similarity as independent variable and *trait empathy* as potential moderator for the whole sample (independent of the severity of the influencer bashing incidents they had seen) (see Figure 1). No significant associations were found in these models, indicating that similarity and trait empathy seems not directly associated with the intention to perform (any type of) bystander behavior, nor moderates trait empathy the relationships between similarity and intentions to perform (specific types of) bystander behavior.

Next, bootstrapped mediation models were calculated to investigate the relationship between similarity and intention to take action and intention to perform the specific types of bystander behaviors, with *state empathy* as potential mediator of these relationships (see Figure 2). An indirect association ( $\beta = .11$ ,  $B = .09$ ,  $SE = .03$ ,  $CI\ 95\% [.04; .14]$ ) was found between perceived similarity and general intention to take action via state empathy: Perceived similarity was positively associated with state empathy ( $\beta = .25$ ,  $B = .20$ ,  $SE = .05$ ,  $CI\ 95\% [.10; .30]$ ,  $p < .001$ ), which was, in turn, positively associated with general intention to take action ( $\beta = .45$ ,  $B = .44$ ,  $SE = .06$ ,  $CI\ 95\% [.32; .55]$ ,  $p < .001$ ). In other words, a higher perceived similarity between oneself and the influencer is associated with a higher level of state empathy, which in turn is associated with a higher intention to take action. This relationship seems fully mediated by state empathy as no direct relationship was found between perceived similarity and general intention to take action. The mediation analysis indicated that the model explained 19.15% of the variance in intention to take action. Similar mediation models were also calculated for the specific types of bystander behaviors and these resulted in similar findings. An indirect association ( $\beta = .10$ ,  $B = .09$ ,  $SE = .03$ ,  $CI\ 95\% [.04; .15]$ ) was found between perceived similarity and intention to perform public supportive

bystander behavior via state empathy: Perceived similarity was positively associated with state empathy ( $\beta = .25$ ,  $B = .20$ ,  $SE = .05$ ,  $CI\ 95\% [.10; .30]$ ,  $p < .001$ ), which was, in turn, positively associated with intention to perform public supportive bystander behavior ( $\beta = .41$ ,  $B = .45$ ,  $SE = .07$ ,  $CI\ 95\% [.32; .59]$ ,  $p < .001$ ). No direct relationship was found between perceived similarity and intention to perform public supportive bystander behavior. The model explained 18.25% of the variance in intention to take action. An indirect association ( $\beta = .14$ ,  $B = .15$ ,  $SE = .04$ ,  $CI\ 95\% [.07; .22]$ ) was also found between perceived similarity and intention to perform private supportive bystander behavior via state empathy: Perceived similarity was positively associated with state empathy ( $\beta = .25$ ,  $B = .20$ ,  $SE = .05$ ,  $CI\ 95\% [.10; .30]$ ,  $p < .001$ ), which was, in turn, positively associated with intention to perform public supportive bystander behavior ( $\beta = .55$ ,  $B = .74$ ,  $SE = .08$ ,  $CI\ 95\% [.59; .88]$ ,  $p < .001$ ). No direct relationship was found between perceived similarity and intention to perform private supportive bystander behavior. The model explained 31.94% of the variance in intention to take action. For the intention to perform negative bystander behavior, no (indirect) relationship was found between perceived similarity and intention to perform public or private negative bystander behavior and also no relationship was found between state empathy and intention to perform public or private negative bystander behavior.

### **The Effect of Severity of Influencer Bashing Incidents in Understanding Bystander Behavior Intentions**

In order to investigate the effect of the severity of the influencer bashing incidents on bystander behavior intentions while controlling for potential moderation effects of perceived similarity and *trait empathy*, bootstrapped double moderation models were calculated (see Figure 3). No significant direct effects of severity, similarity, and trait empathy were found, nor significant two-way or three-way interactions.

Finally, bootstrapped moderated mediation models were calculated to investigate the influence of severity on the indirect relationship between perceived similarity and intention to perform bystander behavior via *state empathy* (see Figure 4). The bootstrapped 95% confidence intervals of the indices of moderated mediation (Hayes, 2015) indicated that moderated mediation was not supported by the data for the general intention to take action, as well as for the intentions to perform specific types of bystander behavior (public supportive bystander behavior, private supportive bystander behavior, public negative bystander behavior, and private negative bystander behavior). Taken together, it seems that the severity of the influencer bashing incident (mild or severe) has no influence on the relationships between similarity, state/trait empathy, and intention to perform (specific forms of) bystander behavior.

### **Discussion**

The main aim of the present study was to explore antecedents of bystander behaviors when witnessing influencer bashing. The present study focused on antecedents of bystander behavior that are proposed by the literature on online peer aggression and celebrity bashing.

First, with regard to empathy, our results indicated that state empathy seems associated with higher intentions to perform supportive bystander behavior when witnessing influencer bashing, but not trait empathy. In previous literature, both trait and state empathy consistently emerged to be positively associated with supportive bystander behavior (e.g., Barlińska et al., 2013; Lambe et al., 2019; Wang, 2021). In contrast, only a limited amount of studies emphasize the importance of state empathy instead of trait empathy for explaining supportive bystander behavior (Mackay, 2019) or prosocial behavior (Ma et al., 2021). It should be noted that the standard deviation for trait empathy was low, which indicates only very small to no differences between the respondents of the present study. Therefore, it seems that potential differences in intentions to perform bystander behavior when witnessing

influencer bashing between those who score high and those who score low on trait empathy could not have been investigated within the present sample. It seems that more research is needed to better understand associations between bystander behavior and both state and trait empathy. In addition, also more research is needed to further investigate the relationship between state empathy and negative bystander behavior. The correlation coefficients indicated a positive correlation between state empathy and negative bystander behavior, which indicates that those who score higher on state empathy not only have higher intentions to support the victim, but also higher intentions to support the perpetrator. This is in contrast with previous research on online peer aggression (Barlińska et al., 2013). A potential explanation for this relationship might be that the included scale for measuring state empathy did not explicitly measure state empathy towards the influencer. It could be that respondents indicated their state empathy towards the aggressor instead of the influencer. They might agree with the aggressor and/or understand why the aggressor is writing hurtful comments.

Second, regarding the perceived similarity, the results of the correlation analyses showed that similarity seems associated with supportive bystander behavior, both publicly and privately, but not with negative bystander behavior. The results of the bootstrapped mediation models indicated that these relationships can be fully explained by one's state empathy: A higher perceived similarity is related to a higher state empathy, which in turn is associated with a higher intention to perform supportive bystander behavior. These findings were similar to previous studies (Cialdini et al., 1997; DeSmet et al., 2016; Wang, 2021) and emphasize the importance of investigating both state empathy and perceived similarity simultaneously in order to better understand bystander behavior. They also provide support for theoretical frameworks such as the intergroup empathy bias (Fourie et al., 2017).

Third, regarding the severity of the influencer bashing, the results of the present study provided limited evidence for the relationship between severity and intentions to perform

(specific forms of) bystander behavior. Those who were exposed to severe influencer bashing incidents had, compared to those who were exposed to mild influencer bashing incidents, only higher intentions to perform one specific form of bystander behavior, namely private supportive bystander behavior. This finding is in line with previous research that indicated that bystanders are more inclined to help the victim if the incident is serious (e.g., Bastiaensens et al., 2014) and rather prefer providing non-publicly visible support than publicly visible support (e.g., Bennett & Banyard, 2016; Wang, 2021). No associations were found between severity and intentions to perform negative bystander behavior. Furthermore, the results indicated that there were differences in associations between severity and state empathy: State empathy was higher among respondents that were exposed to severe forms of influencer bashing compared to respondents that were exposed to mild forms of influencer bashing. The models that tested associations between severity, similarity, trait empathy, and bystander behavior showed no differences between mild and severe forms. The earlier mentioned indirect association between perceived similarity and (public and private) supportive bystander behavior via state empathy are true for both respondents who were exposed to mild influencer bashing incidents and for those who had seen severe forms of influencer bashing. The overall insignificant effects of severity seems to indicate that respondents feel equally responsible to take action when witnessing mild or severe influencer bashing incidents and the severity of the incidents seems not a determinant of their bystander actions.

Taken together, the findings of the present study point to the direction that witnesses of influencer bashing are more inclined to help the influencer if they perceive similarities between themselves and the influencer and if the bashing incidents provoke an empathic response. The perceived similarity has been found crucial for experiencing a close bond or parasocial relationship with the influencer (Rubin & McHugh, 1987; Shan et al., 2020; Yuan

& Lou, 2020). Researchers have indicated that followers who perceive a strong bond between themselves and the influencer are more inclined to follow up advices of the influencer, both within the marketing context (e.g., higher purchase intentions towards the products the influencer promotes; Lou & Kim, 2019), as well as in the context of social and health-related issues, such as climate change (e.g., Boykoff & Goodman, 2009), smoking (e.g., Yoo, 2016), healthy foods (e.g., Folkvord et al., 2020), and sexual health (e.g., Janssen et al., 2021). The results of the present study seem to contribute to these research fields as they suggest that a stronger connection (caused by a higher perceived similarity) can also motivate followers to support (their favorite) influencers when they are victimized online. This means that for influencers, putting hard work in building close relationships with their community of followers not only benefits them for being able to affect followers' behaviors in domains such as purchase behaviors and social and health-related issues, but also for receiving support from the community in case the influencer is being ridiculed or insulted. The support that influencers receive from community members can reduce the aversive outcomes for the influencer, such as in terms of mental-health and well-being problems. Influencers who have been victimized might want to share their negative experiences afterwards with the whole community. Their personal narratives, and especially their appropriate emotional, cognitive, or behavior strategies used for dealing with the situation, might be relevant for young followers who have been the target of harmful online actions themselves (Ouvrein et al., 2019). Reading the stories of fellow-sufferers can help victims of online aggression cope with their own problems.

The present study has limitations that can be used to suggest directions for future research. First, the present study measured bystander behavior intentions towards an influencer that was selected, based on a pre-test, by the authors. Although these incidents were shown under a real Instagram post, the preselection of one influencer by the authors

had a serious impact on the external validity of the study. The results indicated that the respondents, on average, had rather low intentions to take action and saw rather small similarities between themselves and the influencer used in the present study. Future research might want to investigate bystander behavior intentions among actual followers of influencers. In addition, future research might also want to focus on actual bystander behavior instead of intention, or might want to focus on other proximal determinants of actual behavior, such as willingness to intervene in bystander behavior (see Prototype Willingness Model; Gerrard et al., 2008). Furthermore, the present study only focused on active bystander behavior, but research indicates that many bystanders rather remain passive (e.g., Levy & Sela-Shayovitz, 2020; Van Cleemput et al., 2014). This type of bystander behavior was not explicitly measured in the present study, but as the average intentions to perform supportive and negative bystander behavior were rather low, it can be expected that many bystanders of online aggressive incidents towards influencers rather have the intention to do nothing.

Second, the present study made a selection of possible determinants of bystander behavior. Examples of determinants that future research might want to include are past involvement in online aggression and the frequency of being exposed to influencer bashing. Research on bystander behavior when witnessing online peer aggression has indicated that former victims of online aggression, as well as former victims of offline aggression, are more likely to help the victim when witnessing an online aggressive incident, whereas former perpetrators of online aggression and former perpetrators of offline aggression are more likely to join in (e.g., Van Cleemput et al., 2014). Besides past involvement in online aggression, future research might also want to include previous exposure to influencer bashing as a bystander. Longitudinal research suggests that frequent exposure to online aggression can lead to lower levels of empathic responsiveness (Pabian et al., 2016). This can be explained by the desensitization mechanism: Repeated exposure to a certain stimulus can



lead to reduced physiological, emotional, cognitive, and/or behavioral responsiveness to it, including a flattening of emotional reactions and a cognitive switch towards a more favorable attitude towards online aggression (Pabian et al., 2016; Rule & Ferguson, 1986; Strasburger & Wilson, 2014). A final suggestion that we want to formulate for future research is related to the bystander behaviors of other followers. Research on online peer aggression indicates that individual levels of supportive bystander behavior are related to levels of supportive bystander behavior in the peer group and positive peer norms towards supportive bystander behavior (e.g., Lambe et al., 2019). Applied to the context of followers, supportive bystander behaviors of other followers and community norms favoring performing supportive bystander behavior might also determine individuals' intentions to perform supportive bystander behavior.

**Contribution to CAT:** *The present paper will contribute to CAT scholarship as it aims to shed more light on a potentially harmful form of technologically-mediated social interaction. The division and its members have already shown their interest in topics such as social media influencers, online conflicts, and public and private online bystander behavior.*

**Note:** *This paper is submitted solely to the "Communication & Technology Division" of ICA. The full paper nor parts of the paper have NOT been published, presented, or accepted for publication or presentation anywhere else.*

### References

- Abidin, C. (2016). "Aren't these just young, rich women doing vain things online?": Influencer selfies as subversive frivolity. *Social Media + Society*, 2(2), 1–17. <https://doi.org/10.1177/2056305116641342>
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory* (pp. xiii, 617). Prentice-Hall.
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology*, 71(2), 364–374. <https://doi.org/10.1037/0022-3514.71.2.364>
- Barlińska, J., Szuster, A., & Winiewski, M. (2013). Cyberbullying among adolescent bystanders: Role of the communication medium, form of violence, and empathy. *Journal of Community & Applied Social Psychology*, 23(1), 37–51. <https://doi.org/10.1002/casp.2137>
- Bastiaensens, S., Vandebosch, H., Poels, K., Van Cleemput, K., DeSmet, A., & De Bourdeaudhuij, I. (2014). Cyberbullying on social network sites. An experimental study into bystanders' behavioural intentions to help the victim or reinforce the bully. *Computers in Human Behavior*, 31, 259–271. <https://doi.org/10.1016/j.chb.2013.10.036>
- Bennett, S., & Banyard, V. L. (2016). Do friends really help friends? The effect of relational factors and perceived severity on bystander perception of sexual violence. *Psychology of Violence*, 6(1), 64–72. <https://doi.org/10.1037/a0037708>
- Boykoff, M. T., & Goodman, M. K. (2009). Conspicuous redemption? Reflections on the promises and perils of the 'Celebrization' of climate change. *Geoforum*, 40(3), 395–406. <https://doi.org/10.1016/j.geoforum.2008.04.006>

- Brown, L. M., Bradley, M. M., & Lang, P. J. (2006). Affective reactions to pictures of ingroup and outgroup members. *Biological Psychology, 71*(3), 303–311.  
<https://doi.org/10.1016/j.biopsycho.2005.06.003>
- Chapple, C., & Cownie, F. (2017). An investigation into viewers' trust in and response towards disclosed paid-for-endorsements by YouTube lifestyle vloggers. *Journal of Promotional Communications, 5*(2), Article 2.  
<https://www.promotionalcommunications.org/index.php/pc/article/view/95>
- Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C., & Neuberg, S. L. (1997). Reinterpreting the empathy–altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology, 73*(3), 481–494. <https://doi.org/10.1037/0022-3514.73.3.481>
- Dalla Pozza, V., Di Pietro, A., Morel, S., & Psaila, P. (2011). *Justice, freedom and security. Cyberbullying among young people*. European Union Publications Office.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113–126. <https://doi.org/10.1037/0022-3514.44.1.113>
- Davis, M. H., & Franzoi, S. L. (1991). Stability and change in adolescent self-consciousness and empathy. *Journal of Research in Personality, 25*(1), 70–87.  
[https://doi.org/10.1016/0092-6566\(91\)90006-c](https://doi.org/10.1016/0092-6566(91)90006-c)
- De Veirman, M., Cauberghe, V., & Hudders, L. (2017). Marketing through Instagram influencers: The impact of number of followers and product divergence on brand attitude. *International Journal of Advertising, 36*(5), 798–828.  
<https://doi.org/10.1080/02650487.2017.1348035>
- DeSmet, A., Bastiaensens, S., Van Cleemput, K., Poels, K., Vandebosch, H., Deboutte, G., Herrewijn, L., Malliet, S., Pabian, S., Van Broeckhoven, F., De Troyer, O., Deglorie,

- G., Van Hoecke, S., Samyn, K., & De Bourdeaudhuij, I. (2018). The efficacy of the Friendly Attac serious digital game to promote prosocial bystander behavior in cyberbullying among young adolescents: A cluster-randomized controlled trial. *Computers in Human Behavior, 78*, 336–347. <https://doi.org/10.1016/j.chb.2017.10.011>
- DeSmet, A., De Bourdeaudhuij, I., Walrave, M., & Vandebosch, H. (2019). Associations between bystander reactions to cyberbullying and victims' emotional experiences and mental health. *Cyberpsychology, Behavior, and Social Networking, 22*(10), 648–656. <https://doi.org/10.1089/cyber.2019.0031>
- DeSmet, A., Van Cleemput, K., Bastiaensens, S., Poels, K., Vandebosch, H., Malliet, S., Verloigne, M., Vanwolleghem, G., Mertens, L., Cardon, G., & De Bourdeaudhuij, I. (2016). Bridging behavior science and gaming theory: Using the Intervention Mapping Protocol to design a serious game against cyberbullying. *Computers in Human Behavior, 56*, 337–351. <https://doi.org/10.1016/j.chb.2015.11.039>
- DeSmet, A., Veldeman, C., Poels, K., Bastiaensens, S., Van Cleemput, K., Vandebosch, H., & De Bourdeaudhuij, I. (2014). Determinants of self-reported bystander behavior in cyberbullying incidents amongst adolescents. *Cyberpsychology, Behavior and Social Networking, 17*(4), 207–215. <https://doi.org/10.1089/cyber.2013.0027>
- Dillon, K. P., & Bushman, B. J. (2015). Unresponsive or un-noticed? Cyberbystander intervention in an experimental cyberbullying context. *Computers in Human Behavior, 45*, 144–150. <https://doi.org/10.1016/j.chb.2014.12.009>
- Dimofte, C. V., Forehand, M. R., & Deshpandé, R. (2003). Ad schema incongruity as elicitor of ethnic self-awareness and differential advertising response. *Journal of Advertising, 32*(4), 7–17. <https://doi.org/10.1080/00913367.2003.10639142>

- Djafarova, E., & Matson, N. (2021). Credibility of digital influencers on YouTube and Instagram. *International Journal of Internet Marketing and Advertising*, *15*(2), 131–148. <https://doi.org/10.1504/IJIMA.2021.114338>
- Djafarova, E., & Rushworth, C. (2017). Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users. *Computers in Human Behavior*, *68*, 1–7. <https://doi.org/10.1016/j.chb.2016.11.009>
- Eisenberg, N., & Strayer, J. (1987). *Empathy and its development* (pp. x, 406). Cambridge University Press.
- Folkvord, F., Roes, E., & Bevelander, K. (2020). Promoting healthy foods in the new digital era on Instagram: An experimental study on the effect of a popular real versus fictitious fit influencer on brand attitude and purchase intentions. *BMC Public Health*, *20*(1), 1677. <https://doi.org/10.1186/s12889-020-09779-y>
- Fourie, M. M., Subramoney, S., & Madikizela, P. G. (2017). A less attractive feature of empathy: Intergroup empathy bias. In M. Kondo (Ed.), *Empathy—An Evidence-based Interdisciplinary Perspective*. InTech. <https://doi.org/10.5772/intechopen.69287>
- Gannon, V., & Prothero, A. (2018). Beauty bloggers and YouTubers as a community of practice. *Journal of Marketing Management*, *34*(7–8), 592–619. <https://doi.org/10.1080/0267257X.2018.1482941>
- Gazit, T., Aharony, N., & Amichai-Hamburger, Y. (2019). Tell me who you are and I will tell you which SNS you use: SNSs participation. *Online Information Review*, *44*(1), 139–161. <https://doi.org/10.1108/OIR-03-2019-0076>
- Gerrard, M., Gibbons, F. X., Houlihan, A. E., Stock, M. L., & Pomery, E. A. (2008). A dual-process approach to health risk decision making: The prototype willingness model. *Developmental Review*, *28*(1), 29–61. <https://doi.org/10.1016/j.dr.2007.10.001>

- Gräve, J.-F. (2017). Exploring the perception of influencers vs. traditional celebrities: Are social media stars a new type of endorser? *Proceedings of the 8th International Conference on Social Media & Society - #SMSociety17*, 1–5.  
<https://doi.org/10.1145/3097286.3097322>
- Hassan, S., Yacob, M. I., Nguyen, T., & Zambri, S. (2018). *Social media influencer and cyberbullying: A lesson learned from preliminary findings*. 9th Knowledge Management International Conference (KMICe).  
<http://www.kmice.cms.net.my/kmice2018>
- Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research*, 50(1), 1–22. <https://doi.org/10.1080/00273171.2014.962683>
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (Third Edition). Guilford Press.  
<https://www.guilford.com/books/Introduction-to-Mediation-Moderation-and-Conditional-Process-Analysis/Andrew-Hayes/9781462549030>
- Hoeghele, D., Schmidt, S. L., & Torgler, B. (2016). The importance of key celebrity characteristics for customer segmentation by age and gender: Does beauty matter in professional football? *Review of Managerial Science*, 10(3), 601–627.  
<https://doi.org/10.1007/s11846-015-0172-x>
- Hoffner, C., & Buchanan, M. (2005). Young adults' wishful Identification with television characters: The role of perceived similarity and character attributes. *Media Psychology*, 7(4), 325–351. [https://doi.org/10.1207/S1532785XMEP0704\\_2](https://doi.org/10.1207/S1532785XMEP0704_2)
- Janssen, L., Croes, E., & Schouten, A. P. (2021). *#safesex – The effectiveness of social media influencers as ambassadors of sexual health behavior*. ARPH (Association for Researchers in Psychology and Health), Netherlands.

- Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, *29*(4), 589–611.  
<https://doi.org/10.1016/j.adolescence.2005.08.010>
- Lambe, L. J., Cioppa, V. D., Hong, I. K., & Craig, W. M. (2019). Standing up to bullying: A social ecological review of peer defending in offline and online contexts. *Aggression and Violent Behavior*, *45*, 51–74. <https://doi.org/10.1016/j.avb.2018.05.007>
- Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* Appleton-Century-Crofts. Educational Division.
- Levy, M., & Sela-Shayovitz, R. (2020). Cyberaggression: The effect of parental monitoring on bystander roles. *International Journal of Child, Youth and Family Studies*, *11*(4.2), 13–36. <https://doi.org/10.18357/ijcyfs114.2202019986>
- Lim, X. J., Mohd Radzol, A. R. bt, Cheah, J.-H. (Jacky), & Wong, M. W. (2017). The Impact of social media influencers on purchase intention and the mediation effect of customer attitude. *Asian Journal of Business Research*, *7*(2).  
<https://doi.org/10.14707/ajbr.170035>
- Liviatan, I., Trope, Y., & Liberman, N. (2008). Interpersonal similarity as a social distance dimension: Implications for perception of others' actions. *Journal of Experimental Social Psychology*, *44*(5), 1256–1269. <https://doi.org/10.1016/j.jesp.2008.04.007>
- Lou, C., & Kim, H. K. (2019). Fancying the new rich and famous? Explicating the roles of influencer content, credibility, and parental mediation in adolescents' parasocial relationship, materialism, and purchase intentions. *Frontiers in Psychology*, *10*, 2567. <https://doi.org/10.3389/fpsyg.2019.02567>
- Ma, X., Luo, D., Deng, T., & Tao, Y. (2021). *The influence of trait empathy on implicit altruism*. 2021 5th International Seminar on Education, Management and Social

Sciences (ISEMSS 2021), Chengdu, China.

<https://doi.org/10.2991/assehr.k.210806.109>

Macaulay, P. J. R., Boulton, M. J., & Betts, L. R. (2019). Comparing early adolescents' positive bystander responses to cyberbullying and traditional bullying: The impact of severity and gender. *Journal of Technology in Behavioral Science*, 4(3), 253–261.

<https://doi.org/10.1007/s41347-018-0082-2>

Machackova, H., & Pfetsch, J. (2016). Bystanders' responses to offline bullying and cyberbullying: The role of empathy and normative beliefs about aggression. *Scandinavian Journal of Psychology*, 57(2), 169–176.

<https://doi.org/10.1111/sjop.12277>

Mackay, Y. O. C. (2019). *Teens who intervene: Identifying factors related to adolescent cyber-bystander intervention in cyberbullying*. Canterbury Christ Church University.

McComb, S. E., Gobin, K. C., & Mills, J. S. (2021). The effects of self-disclaimer Instagram captions on young women's mood and body image: The moderating effect of participants' own photo manipulation practices. *Body Image*, 38, 251–261.

<https://doi.org/10.1016/j.bodyim.2021.04.011>

MorningConsult. (2019). *Influencer report: Engaging gen Z and millennials*.

<https://morningconsult.com/influencer-report-engaging-gen-z-and-millennials/>

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>

O'Connell, P., Pepler, D., & Craig, W. (1999). Peer involvement in bullying: Insights and challenges for intervention. *Journal of Adolescence*, 22(4), 437–452.

<https://doi.org/10.1006/jado.1999.0238>



- Ouvrein, G., Backer, C. J. S. D., & Vandebosch, H. (2018). Joining the clash or refusing to bash? Bystanders reactions to online celebrity bashing. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *12*(4), Article 4.  
<https://doi.org/10.5817/CP2018-4-5>
- Ouvrein, G., De Backer, C. J. S., & Vandebosch, H. (2018). Online celebrity aggression: A combination of low empathy and high moral disengagement? The relationship between empathy and moral disengagement and adolescents' online celebrity aggression. *Computers in Human Behavior*, *89*, 61–69.  
<https://doi.org/10.1016/j.chb.2018.07.029>
- Ouvrein, G., Hallam, L., J. S. De Backer, C., & Vandebosch, H. (2021). Bashed at first sight: The experiences and coping strategies of reality-TV stars confronted with celebrity bashing. *Celebrity Studies*, *12*(3), 389–406.  
<https://doi.org/10.1080/19392397.2019.1637269>
- Ouvrein, G., Pabian, S., Machimbarrena, J. M., De Backer, C. J. S., & Vandebosch, H. (2018). Online Celebrity Bashing: Wrecking Ball or Good for You? Adolescent girls' attitudes toward the media and public bashing of Miley Cyrus and Selena Gomez. *Communication Research Reports*, *35*(3), 261–271.  
<https://doi.org/10.1080/08824096.2018.1467833>
- Ouvrein, G., Vandebosch, H., & De Backer, C. J. S. (2017). Celebrity critiquing: Hot or not? Teen girls' attitudes on and responses to the practice of negative celebrity critiquing. *Celebrity Studies*, *8*(3), 461–476. <https://doi.org/10.1080/19392397.2017.1307126>
- Ouvrein, G., Vandebosch, H., & De Backer, C. J. S. (2019). Celebrities' Experience with cyberbullying: A framing analysis of celebrity stories in online news articles in teen magazines. In H. Vandebosch & L. Green (Eds.), *Narratives in Research and*

- Interventions on Cyberbullying among Young People* (pp. 181–198). Springer International Publishing. [https://doi.org/10.1007/978-3-030-04960-7\\_12](https://doi.org/10.1007/978-3-030-04960-7_12)
- Ouvrein, G., Vandebosch, H., & De Backer, C. J. S. (2020). Online celebrity bashing: Purely relaxation or stressful confrontation? An experimental study on the effects of exposure to online celebrity bashing on the emotional responses and physiological arousal among adolescent bystanders. *Cyberpsychology, Behavior and Social Networking*, 23(9), 588–594. <https://doi.org/10.1089/cyber.2019.0642>
- Pabian, S., Vandebosch, H., Poels, K., Van Cleemput, K., & Bastiaensens, S. (2016). Exposure to cyberbullying as a bystander: An investigation of desensitization effects among early adolescents. *Computers in Human Behavior*, 62, 480–487. <https://doi.org/10.1016/j.chb.2016.04.022>
- Patterson, L. J., Allan, A., & Cross, D. (2017). Adolescent perceptions of bystanders' responses to cyberbullying. *New Media & Society*, 19(3), 366–383. <https://doi.org/10.1177/1461444815606369>
- Rubin, R. B., & McHugh, M. P. (1987). Development of parasocial interaction relationships. *Journal of Broadcasting & Electronic Media*, 31(3), 279–292. <https://doi.org/10.1080/08838158709386664>
- Rudnicki, K., Ouvrein, G., De Backer, C., & Heidi, V. (2020). Intranasal oxytocin administration reduces bystanders' acceptance of online celebrity bashing. *International Journal of Bullying Prevention*, 2(1), 29–40. <https://doi.org/10.1007/s42380-019-00054-2>
- Rule, B. G., & Ferguson, T. J. (1986). The Effects of Media Violence on Attitudes, Emotions, and Cognitions. *Journal of Social Issues*, 42(3), 29–50. <https://doi.org/10.1111/j.1540-4560.1986.tb00241.x>

- Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior, 15*(2), 112–120. <https://doi.org/10.1016/j.avb.2009.08.007>
- Schouten, A. P., Janssen, L., & Verspaget, M. (2020). Celebrity vs. Influencer endorsements in advertising: The role of identification, credibility, and Product-Endorser fit. *International Journal of Advertising, 39*(2), 258–281. <https://doi.org/10.1080/02650487.2019.1634898>
- Shan, Y., Chen, K.-J., & Lin, J.-S. (Elaine). (2020). When social media influencers endorse brands: The effects of self-influencer congruence, parasocial identification, and perceived endorser motive. *International Journal of Advertising, 39*(5), 590–610. <https://doi.org/10.1080/02650487.2019.1678322>
- Skeldon, P. (2019, January 31). Young affiliates: Nearly a fifth of British children aspire to be social media influencers. *Telemedia Online*. <https://www.telemediaonline.co.uk/young-affiliates-nearly-a-fifth-of-british-children-aspire-to-be-social-media-influencers/>
- Stets, J. E., & Burke, P. J. (2000). Identity Theory and Social Identity Theory. *Social Psychology Quarterly, 63*(3), 224. <https://doi.org/10.2307/2695870>
- Strasburger, V. C., & Wilson, B. J. (2014). Television violence: Sixty years of research. In D. A. Gentile & M. Rich (Eds.), *Media Violence and Children. A complete guide for parents and professionals, second edition*. (pp. 135–178). Praeger.
- Torgal, C., Espelage, D. L., Polanin, J. R., Ingram, K. M., Robinson, L. E., El Sheikh, A. J., & Valido, A. (2021). A meta-analysis of school-based cyberbullying prevention programs' impact on cyber-bystander behavior. *School Psychology Review, 0*(0), 1–15. <https://doi.org/10.1080/2372966X.2021.1913037>
- Tratlehner, E. (2018). *Famke Louise* [Documentary]. RTL Nederland.

- Van Cleemput, K., Vandebosch, H., & Pabian, S. (2014). Personal characteristics and contextual factors that determine “helping,” “joining in,” and “doing nothing” when witnessing cyberbullying. *Aggressive Behavior, 40*(5), 383–396.  
<https://doi.org/10.1002/ab.21534>
- Van der Graaff, J., Meeus, W., de Wied, M., van Boxtel, A., van Lier, P. A. C., Koot, H. M., & Branje, S. (2016). Motor, affective and cognitive empathy in adolescence: Interrelations between facial electromyography and self-reported trait and state measures. *Cognition & Emotion, 30*(4), 745–761.  
<https://doi.org/10.1080/02699931.2015.1027665>
- Van der Veer, N., Boekee, S., & Hoekstra, N. (2021). *National social media research 2021*. Newcom Research & Consultancy. <https://www.newcom.nl/socialmediaonderzoek/>
- Wang, S. (2021). Standing up or standing by: Bystander intervention in cyberbullying on social media. *New Media & Society, 23*(6), 1379–1397.  
<https://doi.org/10.1177/1461444820902541>
- Wei, M.-L., Fischer, E., & Main, K. J. (2008). An examination of the effects of activating persuasion knowledge on consumer response to brands engaging in covert marketing. *Journal of Public Policy & Marketing, 27*(1), 34–44.  
<https://doi.org/10.1509/jppm.27.1.34>
- Yoo, W. (2016). The influence of celebrity exemplars on college students’ smoking. *Journal of American College Health, 64*(1), 48–60.  
<https://doi.org/10.1080/07448481.2015.1074238>
- Yuan, S., & Lou, C. (2020). How social media influencers foster relationships with followers: The roles of source credibility and fairness in parasocial relationship and product interest. *Journal of Interactive Advertising, 20*(2), 133–147.  
<https://doi.org/10.1080/15252019.2020.1769514>

**Appendix A**

*Bootstrapped paired samples t-tests for intentions to perform specific forms of bystander behavior*

Pair	Mean difference	SD	95% C.I.	<i>t</i>	df	<i>p</i> -value
Public supportive bystander behavior – Private supportive bystander behavior	-.44	.98	-.57; -.32	-6.92	237	< .001
Public supportive bystander behavior – Public negative bystander behavior	.48	.85	.37; .59	8.73	237	< .001
Public supportive bystander behavior – Private negative bystander behavior	.49	.86	.38; .60	8.76	237	< .001
Private supportive bystander behavior – Public negative bystander behavior	.92	1.15	.78; 1.07	12.36	237	< .001
Private supportive bystander behavior – Private negative bystander behavior	.93	1.14	.78; 1.07	12.61	237	< .001
Public negative bystander behavior – Private negative bystander behavior	.01	.22	-.02; .03	.29	237	.769