

# **The Effect of Organizational Social Safety Policy on Conflict Perception and Reaction**

Lena Kudelska

S4727517

Department of Psychology, University of Groningen

PSB3E-BT15: Bachelor Thesis

Group 43

Supervisor: Dr. Maja Graso

Second evaluator: Roxana Bucur

In collaboration with: Lisette Abels, Keijzer Laura, Muller-Kickelberg Maike, Oosterhuis Cindy,

Vastova Katerina

July 3, 2024

## Abstract

The present study aims to explore the impact of social safety campaigns on harm perception and the influence of neuroticism on that relationship within the organizational context. I hypothesized that exposure to social safety messages would prime participants to perceive more harm in ambiguous social situations. Additionally, I hypothesized that neurotic individuals would be increasingly vulnerable to social safety messages and consequently perceive more harm in ambiguous situations. To test these hypotheses, our bachelor thesis group recruited participants from personal networks and Prolific and conducted an online vignette-based study. The results indicate that social safety campaigns do not impact harm perception and that neuroticism does not moderate that relationship. The findings of the study suggest the need for longitudinal research and the development of a more comprehensive range of measures to explore further the potential impact of social safety campaigns on the processes underlying harm perception.

## **The Effect of Organizational Social Safety Policy on Conflict Perception and Reaction**

April 17, 2024 will mark the first anniversary of the launch of the Social Safety campaign developed by the University of Groningen (University of Groningen, 2023a). The Social Safety campaign is part of the university's broader initiative, called the Zero Tolerance policy, which explicitly prohibits bullying, discrimination, stalking, and harassment in academia (University of Groningen, 2023a). The campaign employs social media platforms to distribute safety messages and thus initiate a conversation about the nature of undesirable social behaviors. It acknowledges that people interpret social behavior differently; while some might be unaffected by a slight remark, others might find it painful and distressing (University of Groningen, 2023a). The campaign also raises awareness of less severe forms of inappropriate social behavior, such as forwarding an email with someone's negative feedback without permission. Therefore, the campaign encourages its students and staff members to be more mindful of their social environment and assists them in reporting workplace misconduct. Since its inception, the campaign has attracted considerable attention, potentially leading to an increase in the number of reports of unsafe behavior (University of Groningen, 2023a). Yet, little research exists so far on the real impact of such initiatives on the perception of harm in the organizational context. Do social safety campaigns shape the way individuals think about and react to workplace conflict, and if so, how?

Institutions play a particularly significant role in ensuring the social safety of their workforce by establishing norms and legal guidelines for proper workplace conduct (Slavich et al., 2023). They are responsible for openly discussing the topic of psychological safety and encouraging individuals to come forward with potential complaints, hence facilitating social consensus on what constitutes harmful behavior (Hamre et al., 2023). When employees become

aware of the organization's commitment to safety, they might be more willing to detect potential hazards and report on them (Bond et al., 2010). One possible drawback of organizational safety initiatives is that they might inadvertently lead individuals to perceive harm in situations when it is not present (Bleske-Rechek et al., 2023).

Therefore, as organizations are increasingly vocal about their attempts to eradicate practices that threaten psychological safety, the issue of making individuals overly sensitive to harm becomes more salient (McGrath et al., 2019). The social phenomenon of “concept creep” refers to the growing tendency to broaden the psychological definition of what constitutes harm (Dakin et al., 2022). Therefore, plentiful behaviors that were previously considered acceptable are now labeled as damaging to one’s mental or physical health (Furedi, 2016). While the expansion of harm-related concepts increases moral awareness about potentially unethical behaviours, it also leads to greater “self-perceived victimhood” (McGrath et al., 2019, p. 79).

In addition to organizational and cultural factors, individual dispositions might influence the perception of harm and recognition of undesirable behavior in the workplace. In particular, neuroticism is a personality trait related to negative emotionality and more frequent experiences of anxiety (Luminet et al., 2000). People scoring high in neuroticism are more likely to encounter interpersonal difficulties and suffer from mental or physical health problems (Borghuis et al., 2020). Neurotic individuals also pay more attention to negatively valenced stimuli and often interpret ambiguous information as threatening (Ormel et al., 2013).

Individual beliefs and judgments surrounding interpersonal harm most likely develop in the presence of both situational and personality dispositions (Bleske-Rechek et al., 2023). However, the question of how workplace campaigns impact the perception of harm in ambiguous

social interactions, where people's intentions are often not apparent, still needs to be answered. Do they prime individuals to perceive more harm than there, in fact, is? Moreover, as people differ in their sensitivity to hostile and threatening stimuli, the impact of social safety campaigns on harm perception may vary depending on the level of an individual's neuroticism. The research question explored in this bachelor thesis is: Does exposure to social safety messages relate to the perception of harm in ambiguous social scenarios, and how does neuroticism moderate that relationship?

### **Theoretical Foundations**

To address the question regarding the impact of social safety campaigns on harm perception, I will draw from the literature on moral awareness and the phenomenon of concept creep. In the subsequent section, I will explore the rationale for considering neuroticism as a potential moderating variable in the relationship between exposure to social safety messages and the perception of harm.

#### ***Moral awareness and identification of harm***

According to Butterfield et al. (2000), organizational safety management strategies help individuals detect potentially harmful social behaviors by increasing their moral awareness of psychological safety. With improved moral awareness, employees can correctly categorize incoming information as morally relevant or irrelevant (Butterfield et al., 2000). Therefore, when observing or participating in social interactions, individuals with enhanced moral awareness can successfully judge whether and how their behavior affects the well-being of others (Reynolds & Miller, 2015). Individuals are more likely to judge an action as morally relevant if previously exposed to social cues that classify it as harmful (Butterfield et al., 2000). Organizations supply

their employees with social cues that make certain issues more salient, for instance, those related to interpersonal harm, and suggest how individuals should interpret their moral relevance (Jones, 1991). Social scripts are especially valuable when employees encounter an ambiguous situation (Jones, 1991). Individuals who struggle with interpreting their colleague's behavior often base their judgment on the information that is most easily accessible in their minds (Butterfield et al., 2000). If the organization has recently deemed a behavior problematic, an individual is more likely to use that information in their judgment since it is easily retrievable (Butterfield et al., 2000).

### ***Social consensus and harm perception***

Organizations may facilitate moral awareness of psychological safety by establishing a social consensus on what is inappropriate and harmful social behavior. According to Barnett and Valentine (2004, p.339), social consensus is the "perceived degree of social agreement that an action is morally acceptable or morally unacceptable." The degree of social consensus on an issue directly relates to its perceived moral intensity (Barnett & Valentine, 2004). Morally intense issues are more often discussed in moral terms and, therefore, categorized as moral issues (Butterfield et al., 2000). Due to their vividness, morally intense issues attract more attention, which causes individuals to ascribe increased levels of personal responsibility for the given issue (Butterfield et al., 2000). A social consensus on an issue, such as psychological safety, minimizes ambiguity, making it easier for individuals to recognize a potentially threatening situation (Bateman et al., 2012). While informing individuals about the potential threats in their workplace and thus increasing their moral awareness on the ethical issue has numerous advantages, it also potentially leads to an inflated sense of vulnerability to harm (McGrath et al., 2019).

### *Negative implications of harm inflation*

Since the emergence of the idea of “concept creep,” it has been theorized that the current social preoccupation with harm is a result of the increasing moral sensitivity of people to threatening stimuli in their social environment (Haslam et al., 2020). As moral sensitivity increases, individuals become perceptive of a broader range of violations resulting in harm or maltreatment (McGrath et al., 2019). Another explanation considers the influence of "expansion entrepreneurs" who deliberately broaden harm-related concepts as a "means of enhancing the perceived seriousness of a social problem or threat" (Haslam et al., 2020, p. 268). Subsequently, they enlarge the group of potential victims and offenders involved in that social issue, making the threat more salient to the public (Haslam et al., 2020). In the organizational context, companies increase the salience of psychological safety by developing initiatives that draw attention to a broader range and more subtle forms of interpersonal harm (Bleske-Rechek et al., 2023).

Haslam (2016) argue that the semantic inflation of the concept of harm occurs along a vertical and horizontal dimension. Vertical creep refers to the process by which the threshold for detecting harmful behavior is lowered (Haslam et al., 2020), which means that people notice milder or less severe forms of harm. On the other hand, horizontal creep enables individuals to coin new forms of harm (Haslam et al., 2021). For instance, the expansion of the concept of abuse made it possible to include the experiences of emotional neglect and non-physical aggression, such as verbal assault, under its definition (Haslam, 2016). Moreover, there are individual differences in the extent to which people expand their conceptualizations of harm; people who perceive themselves as more personally vulnerable or entitled tend to hold broader conceptualizations of harm (Haslam et al., 2021).

**Hypothesis 1.** As a result of an increased moral awareness and social consensus, I hypothesize that individuals exposed to safety campaigns (versus those not exposed to a social safety campaign) will perceive more harm in ambiguous social scenarios.

### *The role of neuroticism in harm perception*

Neuroticism is an important personality disposition that potentially underlies the processes related to harm and conflict perception (Bleske-Rechek et al., 2023). Neuroticism has been linked to a broadened conceptualization of harm (Bleske-Rechek et al., 2023), negative emotionality, and attentional bias toward negative stimuli (Bunghez et al., 2023). In the workplace environment, neurotic individuals tend to perceive supervisory behavior as more abusive than people who exhibit lower levels of neuroticism and higher emotional stability (Henle & Groß, 2013). Specifically, neurotic individuals demonstrate increased sensitivity to interpersonal rejection, which makes them more vulnerable to workplace ostracism (Leung et al., 2011). They are also more likely to perceive themselves as vulnerable (personal vulnerability) and are more sensitive to violations of justice (Haslam et al., 2020).

The reason why neuroticism might be associated with greater sensitivity to harm lies in the cognitive adaptive theory, which argues that the attentional and interpretive biases related to neuroticism function as "modes of adaptation to the major, universal challenges of human life" (Matthews, 2008, p. 69). In the social context, greater sensitivity to threats allows neurotic individuals to anticipate and quickly recognize potentially harmful situations (Matthews, 2008). Although they may lack adequate coping strategies, individuals scoring high in neuroticism are more effective in detecting threats in ambiguous social situations (Matthews, 2008).



Since people with neurotic tendencies are more prone to negativity bias, they also exhibit a lower threshold for detecting harm and often identify as victims of mistreatment (Wang et al., 2020). Victim sensitivity, associated with neuroticism, may lead individuals to make inaccurate social judgments regarding other people's trustworthiness or cooperativeness (Gollwitzer et al., 2012). For instance, individuals with higher scores on victim sensitivity pay more attention to information that suggests that the other person is uncooperative and are less interested in the cues that point to someone's cooperativeness (Gollwitzer et al., 2012). As neurotic individuals pay more attention to potentially threatening information and are more sensitive to violations of injustice, they are more likely to interpret ambiguous information as harmful.

**Hypothesis 2.** Therefore, I hypothesize that there will be a positive relationship between the level of neuroticism and perception of harm.

Since neurotic individuals exhibit attentional bias toward negative stimuli, they are more susceptible to messages that warn them about the harmful behavior of others (Bleske-Rechek et al., 2023). The potential moderating effect of neuroticism on harm perception comes from studies on affective priming (Zhang et al., 2010). The phenomenon of affective priming refers to a process by which exposure to a stimulus (“the prime”) aids the processing of a subsequent stimulus when the prime is “evaluatively consistent” with the stimulus it proceeds (Musch & Klauer, 2003, p.8). Therefore, both stimuli must be either positively or negatively valenced, as in the case of the words “loss” and “accident” (Klauer & Musch, 2003). It has been argued that highly neurotic individuals access semantic memory networks related to negative ideas more easily than individuals who score lower on neuroticism (Robinson et al., 2007). Accessed memory nodes, in turn, activate other nodes relevant to negative thoughts via the process of spreading activation (Robinson et al., 2007). Within the organizational context, neurotic

individuals might pay more attention to negatively valenced messages about workplace misconduct and more easily retrieve them when judging the appropriateness of their coworkers' behavior (Bleske-Rechek et al., 2023).

**Hypothesis 3.** Therefore, I hypothesize that neuroticism will moderate the relationship between exposure to social safety campaigns and the perception of harm. Specifically, individuals who are neurotic will be more influenced by the social safety messages than non-neurotic individuals and will interpret ambiguous social situations as more threatening.

## Methods

### Recruitment and Sample Characteristics

The following study was approved by the ethical committee of the Faculty of Behavioral and Social Sciences at the University of Groningen. The participants were recruited through personal networks such as Facebook and LinkedIn, where the questionnaire was posted with a uniform message, and through Prolific, using funding allocated to the thesis group by the faculty.

In total, we recruited 175 participants. To protect individuals' anonymity, the group's bachelor thesis supervisor combined the two samples of participants recruited through personal networks (105 participants) and Prolific (70 participants). There were 14 participants who did not pass the attention check, and therefore, their data was removed from the analysis. Out of 161 participants, 78 were women, 81 were men, and 2 preferred not to identify themselves. The sample consisted of 50 students, 65 employees, 37 participants who identify as both students and employees, as well as 9 participants who do not identify as either students or employees. The average age of participants was 30.01 years ( $SD = 12.50$ ).

## **Procedure and Materials**

We conducted an online vignette-based experiment to explore the impact of social safety campaigns on harm perception and the moderating effect of neuroticism on that relationship. First, the invitation link to the survey was sent to participants. Participants were told that the study intended to examine how people perceive different kinds of workplace interactions and how individual differences or personality characteristics influence those perceptions. They were also assured about the confidentiality of their responses. In addition, at the beginning of the study, we obtained participants' informed consent and socio-demographic information, such as their gender, political ideology, work status, and experience.

### ***Manipulation of Social Safety Campaign***

In the study, the exposure to a social safety campaign was manipulated to assess whether it affects participants' perceptions of harm. Participants were assigned at random either to an experimental or control condition. In the experimental condition, participants were exposed to a social safety campaign, while those in the control condition were not.

In the experimental condition, participants were primed with 4 different posters of a fictional social safety campaign (see Figure 1, Appendix A). The posters were designed in the typical square format of Instagram posts, as the social media platform is a realistic channel that institutions use to spread social safety campaigns. Both the layout and the content of the fictitious campaign are inspired by the "Just Ask" poster campaign launched by the University of Groningen in April 2023 (University of Groningen, 2023b). The slogans stand out in white lettering against a red background, which gives them a warning appearance at first glance. The aim of the campaign is to make people aware of the potential harm that may arise from social

interactions ("Words can hurt"; "Didn't mean it?") and to define organizational norms of behavior ("Stand firm, speak out"; "Don't ignore the signals").

The two posters pointing out the potential harm of social interactions contain speech bubbles with examples of interactions that can be hurtful even without malicious intent. This makes it clear to the recipient which ambiguous forms of harmful behavior the campaign is targeting. The key message here is that harm can result from verbal interactions and that the assessment of this harm is in the eye of the beholder and does not depend on bad intentions. The other two posters entail standards of behavior and direct calls to action. They point out the individual's responsibility to recognise and address inappropriate behavior.

After familiarizing themselves with the content of the social safety campaign, participants were instructed to consider the goals of the campaign and reflect on how effectively it communicates its message and engages its audience. They were also asked to briefly summarize the main message of the campaign in the designated place.

### *Vignettes*

In the study, the vignettes were used to explore whether a previous manipulation of the exposure to a social safety campaign would affect participants' judgement of harm in ambiguous social situations. We chose to develop two vignettes to increase the external validity of the study and examine participants' reactions to two unrelated workplace scenarios. In the first vignette ("Group outing vignette"), participants saw an interaction among four employees who work together at a large consulting firm (see Figure 2, App. A). In the second vignette, participants were asked to read a text conversation between a manager and a team member ("Clothing vignette") (see Figure 3, App. A). The exchanges in both vignettes took place on a social media platform (WhatsApp) and were displayed in the form of a screenshot of the group chat.

The vignettes were designed as Whatsapp messages to increase the similarity to real-life digital interactions. WhatAapp is a very popular message exchange platform, and it's very likely that the participants are familiar with it. This potentially contributes to making the vignettes more relatable and realistic. It also makes the vignettes accessible to a broader population. Both vignettes are based on workplace issues that should be familiar to most participants. Themes such as (in)appropriate workplace compliments and inclusion or exclusion are intended to touch upon themes that participants are very likely to have been exposed to already. This ensures familiarity and increases the chance of getting genuine responses.

Several components have been incorporated into the vignettes to optimize ambiguity in both. Firstly, the vignettes were designed in a way that leaves room for interpretation. For instance, in the first vignette, the simple reply 'right' was added as a last statement to create an open-ended conclusion. This way, the participants are invited to make their own conclusions about how harmful the behavior in the vignette is. Secondly, through the use of Whatsapp messages, participants are not able to read body language and facial expressions. This way, participants are required to rate intent and tone from text alone. Furthermore, the vignettes have been designed in a way that the harm is not overly explicit. For example, in the first vignette, the exclusion is communicated casually, which could be perceived as either innocent or as deliberate exclusion. Lastly, gender-neutral names were included in both vignettes to lessen the effect of gender bias on the participants' responses and to simplify the study's design.

After familiarizing themselves with each vignette, participants were instructed to answer a set of questions measuring the perceived severity of the harm inflicted on the teammate.

In the last part of the survey, participants were asked to answer questions measuring personality traits and personal characteristics, including neuroticism. Upon completing the

survey, participants were debriefed and were given a chance to reflect on their participation in the study. They were also asked to indicate if they confirm their consent, or if they prefer to rescind it.

## **Measures**

### ***Severity of Harm***

Respondents were asked how much harm they thought the person in each vignette experienced by rating it on a 6-point Likert scale ranging from 1, 'no harm at all' to 6, 'a great deal of harm'. Depending on the vignette, the statements measuring harm severity were phrased in the following way: "Based on the exchange you read, rate the extent of harm you believe Blake has experienced due to team's comments/due to the exchange with the manager (Jordan)." This was based on the measure developed by Dakin et al. (2022).

### ***Neuroticism***

To assess the level of neuroticism, participants were asked to rate their level of agreement with eight different statements (ex. "I see myself as a person who is depressed, blue") on a 6-point Likert scale ranging from 1 "no agree at all" to 6 "agree at all." The statements were derived from the NEO Personality Inventory (Costa & McCrae, 1992) ( $\alpha=0.87$ ).

## **Results**

### **Preliminary Analysis**

To explore the impact of social safety campaigns on harm perception, I performed an independent-samples t-test using the SPSS software. Next, I conducted a regression analysis to check whether there was a relationship between the levels of neuroticism and harm perception.

Additionally, to test for the moderating effect of neuroticism on the relationship between exposure to social safety campaigns and harm perception, I conducted a moderation analysis using the PROCESS Macro for SPSS software (Hayes, 2013). As part of the preliminary analysis, the assumptions pertaining to the independent-samples t-test, simple linear regression, and multiple regression were checked using the SPSS software.

***Assumption check for the independent-samples t-test***

To test the assumptions of normality and homogeneity of variance pertaining to the independent-samples t-test, I performed the Kolmogorov-Smirnov and Levene's tests separately for each vignette. The results of the tests are shown below.

***Group outing vignette.*** For the group outing vignette, the Kolmogorov-Smirnov test indicated that the data were not normally distributed in the control,  $D(84) = 0.25, p < 0.001$  and experimental (social safety campaign) condition,  $D(77) = 0.19, p < 0.000$ . Next, in regard to the first vignette, Levene's test indicated that the assumption of homogeneity of variances was met,  $F(159) = 0.03, p = 0.867$ .

***Clothing vignette.*** For the clothing vignette, the Kolmogorov-Smirnov test indicated that the data were not normally distributed in the control,  $D(84) = 0.18, p < 0.001$ , and experimental (social safety campaign) condition,  $D(77) = 0.15, p < 0.001$ . Next, in regard to the second vignette, Levene's test indicated that the assumption of homogeneity of variances was met,  $F(159) = 0.37, p = 0.547$ .

***Assumption check for the regression analysis***

To test the assumptions of normality, linearity, and homoscedasticity of variance pertaining to the simple linear regression and multiple linear regression, I analyzed the normal Predicted Probability plot, scatterplot of standardized residuals and the histogram of standardized residuals. Additionally, to check for multicollinearity and independence of errors, I calculated the Variance Inflation Factor and performed the Durbin-Watson test. Each assumption check was conducted separately for each vignette, as indicated below.

***Group outing vignette.*** The normal Predicted Probability plot (see Figure 1, App. B) indicated that the data was not normally distributed since the data points are located next to the diagonal normality line in a curve-shaped manner. Therefore, the assumption of normality was not met.

However, the scatterplot of standardized residuals (see Figure 2, App. B) revealed that the error terms are equally spread along the value of zero on the x-axis and the value of zero on the y-axis, which means that the data does not violate the assumptions of linearity and homoscedasticity of variance. Subsequent examination of the histogram of standardized residuals confirmed that the assumptions of linearity and homoscedasticity were met (see Figure 3, App. B).

Lastly, the values of the Variance Inflation Factor (VIF) are less than 10 (Interaction variable, Tolerance = 0.95, VIF = 1.05; Neuroticism, Tolerance = 0.95, VIF = 1.05), which suggests that the assumption of multicollinearity was met. Also, the assumption of independent errors was also met (Durbin-Watson value = 2.18)

***Clothing vignette.*** The examination of the normal Predicted Probability plot (see Figure 4, App. B) indicated that the assumption of normality was not violated, with the data points



located close to and along the diagonal normality line. The assumptions of homoscedasticity and linearity were also met, as indicated by the histogram of standardized residuals (see Figure 5, App. B) and the scatterplot of standardized residuals (see Figure 6, App. B).

Moreover, the values of the Variance Inflation Factor (VIF) are less than 10 (Interaction variable, Tolerance = 0.95, VIF = 1.05; Neuroticism, Tolerance = 0.95, VIF = 1.05), which suggests that the assumption of multicollinearity was met. The assumption of independent errors was also met (Durbin-Watson value = 1.79)

## **Hypothesis testing**

### ***Examining the effect of social safety campaign on harm perception***

Following the preliminary analysis, an independent-samples t-test was performed to test whether exposure to a social safety campaign had a significant effect on the participants' perception of harm, measured by the severity of harm. An independent sample t-test was performed separately for each vignette. Since Levene's test for equality of variances indicated that the assumption of homogeneity of variances was met for both vignettes, the following analysis reports on the statistics of the independent-sample t-test with equal variances assumed. For each vignette, the results of an independent-sample t-test are shown below.

***Group outing vignette.*** For the group outing vignette, there were no significant differences ( $t(159) = 0.23, p = 0.820$ ) in harm severity for the control condition in which participants were not exposed to a social safety campaign ( $M = 5.12, SD = 1.44$ ) and the experimental condition in which participants were exposed to a social safety campaign ( $M = 5.17, SD = 1.31$ ). The magnitude of the differences in the means (mean difference = 0.05, 95%

$CI: -0.38$  to  $0.48$ ) was nonsignificant. Hence, the first hypothesis was not supported. This means that exposing people to a social safety campaign did not have a significant effect on participants' perception of harm.

***Clothing vignette.*** For the clothing vignette, there were no significant differences ( $t(159) = 1.47, p = 0.144$ ) in harm severity for the control condition in which participants were not exposed to a social safety campaign ( $M = 2.95, SD = 1.71$ ) and the experimental condition in which participants were exposed to a social safety campaign ( $M = 3.35, SD = 1.73$ ). The magnitude of the differences in the means (mean difference =  $0.40$ ,  $95\% CI: -0.14$  to  $0.93$ ) was nonsignificant. Hence, the first hypothesis was not supported.

### ***Examining the effect of neuroticism on harm perception***

Next, a simple linear regression analysis was performed to explore the extent to which neuroticism could predict the severity of harm. The results of the simple linear regression are shown below, separately for each vignette.

***Group outing vignette.*** For the group outing vignette, it was found that neuroticism did not predict a significant amount of variance in the value of the severity of harm,  $R^2 = 0.00, F(1, 158) = 0.53, p = 0.469$ . The levels of neuroticism did not predict the levels of severity of harm,  $\beta = 0.10, p = 0.469$ . Therefore, our second hypothesis was not supported, as neuroticism was not significantly and positively associated with the severity of harm.

***Clothing vignette.*** For the clothing vignette, it was found that neuroticism did not predict a significant amount of variance in the value of the severity of harm,  $R^2 = 0.00, F(1, 158) = 0.02, p = 0.903$ . The levels of neuroticism did not predict the levels of severity of harm,  $\beta = -$

0.02,  $p = .903$ . Therefore, our second hypothesis was not supported, as neuroticism was not significantly and positively associated with the severity of harm.

***Examining the moderating effect of neuroticism***

Finally, a moderation analysis was conducted, using the PROCESS Macro for SPSS software (Hayes, 2013), with the exposure to social safety campaign as the predictor, harm severity as the dependent and neuroticism as the moderator. It was hypothesized that neuroticism would moderate the relationship between the exposure to social safety campaign and harm perception measured by severity of harm.

***Group outing vignette.*** For the group outing vignette, the results indicated that neither exposure to social safety campaigns ( $B = 0.06, p = 0.779$ ) nor neuroticism ( $B = 0.21, p = 0.284$ ) had significant effects on harm severity. There was also no significant interaction between exposure to social safety campaign and neuroticism ( $B = -0.21, p = 0.436$ ). The results indicate that 1% of the variability in harm perception was predicted by all of the variables, that is, exposure to social safety campaign and neuroticism,  $R^2 = 0.01, F(3, 156) = 0.40, p = 0.750$ . Therefore, our third hypothesis was not supported. This means that neurotic individuals do not perceive more harm in ambiguous social interactions following exposure to social safety campaigns.

***Clothing vignette.*** For the group outing vignette, the results indicated that neither exposure to social safety campaigns ( $B = 0.40, p = 0.141$ ) nor neuroticism ( $B = 0.30, p = 0.215$ ) had significant effects on harm severity. There was also no significant interaction between exposure to social safety campaign and neuroticism ( $B = -0.58, p = 0.084$ ). The findings indicate that 3% of the variability in harm perception was predicted by all of the variables, that is,

exposure to social safety campaign and neuroticism,  $R^2 = 0.03$ ,  $F(3, 156) = 1.74$ ,  $p = 0.161$ .

Therefore, our third hypothesis was not supported.

After performing the preliminary analysis and hypothesis testing, it can be concluded that there is no significant association between exposure to social safety campaign and harm perception. Furthermore, neuroticism was not found to moderate the relationship between the exposure to social safety campaign and harm perception.

### **Discussion**

Developing adequate strategies to counter bullying and harassment in the workplace is a prerequisite to fostering an appropriate organizational safety environment (Liang, 2021). Such interventions have the potential to shape employees' understanding of harm-related concepts and instill appropriate skills for conflict recognition and management (Jones, 1991). In particular, social safety campaigns may facilitate social consensus on what constitutes inappropriate behavior in the workplace and increase moral awareness of psychological safety (Butterfield, 2000). Therefore, the following study hypothesized that exposure to social safety messages would lead participants to perceive more harm in ambiguous social situations.

Besides testing the role of situational factors in harm perception, I hypothesized that individual differences in the level of neuroticism would moderate the relationship between exposure to social safety campaigns and perceived severity of harm since neurotic individuals exhibit increased threat recognition and negative emotionality. Therefore, it was hypothesized that participants exposed to social safety messages and displaying higher levels of neuroticism would perceive more harm in ambiguous social situations than participants who scored lower on the neuroticism scale and were also familiar with social safety messages.

In contrast to the predictions, the study's findings did not support the hypotheses mentioned above. Firstly, there was no significant difference in the responses relating to the severity of perceived harm of participants exposed to social safety messages and those in the control condition. Therefore, exposing people to social safety messages did not affect their assessment of how harmful behaviors were in ambiguous social situations. Secondly, neuroticism did not have a significant effect on the levels of perceived severity of harm. Thirdly, statistical analysis did not reveal a significant moderating effect of neuroticism on the relationship between exposure to social safety warnings and harm perception. Individual differences in neuroticism did not influence the perception of harm following exposure to a social safety campaign.

### **Theoretical Implications**

The findings of the present study stand in contrast to research, which found that alerting individuals of the potentially harmful aspects of their daily social interactions led them to perceive more harm in ambiguous situations, regardless of whether the statements they received were intentionally hostile or not (Bleske-Rechek et al., 2023). Informing individuals about the subtle and nuanced nature of some forms of aggression has been hypothesized to increase awareness of undesirable social behavior, thus making it more likely that individuals detect and report on it (Bleske-Rechek et al., 2023).

The present study also challenges the notion that repeated exposure to warning signs of psychological harm gradually lowers the threshold for detecting such threats, making individuals sensitive to a broader range of interpersonal violations (McGrath et al., 2019). The phenomenon of "concept creep" offers a valuable framework for understanding the increased societal

vulnerability to harm and preoccupation with harm-related issues (Haslam et al., 2020). It also suggests that social institutions may purposefully expand notions of psychological harm (Haslam et al., 2020). Considering the present findings, however, it is unclear whether social safety campaigns indeed affect individuals' perceptions of harm by broadening the concept of psychological harm and shedding light on more subtle forms of workplace aggression. It is likely that in the organizational context, characterized by a hierarchical structure, the fear of authority holds some individuals from communicating their disapproval of undesirable workplace practices or behaviors (Dai et al., 2022). In the clothing vignette, participants might have displayed an avoidant coping style and withheld their judgment due to the power imbalance inherent in the relationship between the manager and an employee (Dai et al., 2022).

The present study also makes a significant contribution to the existing literature by exploring the role of individual differences in how people conceptualize harm and interpret ambiguous social situations in the organizational setting. Thus far, it has been found that neurotic individuals tend to assess ambiguous stimuli as more threatening and perceive themselves as more vulnerable to interpersonal rejection or violations of justice (Leung et al., 2011). In the work environment, neurotic employees were more likely to rate their supervisors as abusive than their colleagues, who scored low in neuroticism (Henle & Groß, 2013). Therefore, the lack of a significant positive relationship between neuroticism and interpreting ambiguous social interactions as potentially harmful greatly contrasts previous research.

One explanation for such a discrepancy comes from studies that found that neurotic individuals possess poorer perspective-taking skills and social self-efficacy (Guo et al., 2018). Neurotic individuals often doubt if they correctly interpret other people's feelings, especially if that person needs help (Guo et al., 2018). Therefore, in the context of the study's vignettes,

neurotic participants might have had issues taking the perspective of the allegedly harmed individual and were less confident in their ability to resolve potential issues between the individuals in the given social scenarios (Guo et al., 2018).

Another significant implication of the current study pertains to the observed nonsignificant moderating effect of neuroticism on the relationship between exposure to social safety campaigns and perception of harm. The literature on the effects of priming on neurotic individuals suggests that they are more vulnerable to negatively valenced stimuli (Robinson et al., 2007), which was hypothesized to relate to a greater perception of harm in ambiguous social situations following the exposure to social safety warning signs. However, some studies found that neurotic individuals are less vulnerable to priming effects if it means aligning and coordinating their behavior with that of their interlocutor (interpersonal priming) (Gill et al., 2004). They were found to display lower levels of interaction with their environment and more "inward focus" (paying more attention to themselves rather than their partners in the conversation) (Gill et al., 2004). In the context of the present study, neurotic participants might not have paid attention to the potentially harmful aspects of conversational exchanges in the given vignettes due to the interpersonal nature of the social safety campaign.

### **Practical implications**

The findings of the present study, besides their theoretical implications, have direct relevance for the public and policymakers. They provide insights into the environmental and individual factors underlying harm perception and shed light on the critical issue of psychological safety in the workplace. Instead of being implemented unaided, social safety campaigns could complement other initiatives that aim to heighten moral sensitivity to

interpersonal harm in the organizational context. Company executives may potentially shape employees' understanding of harm by organizing workshops or regular check-ins with staff to discuss unhealthy working habits and their experiences of mistreatment. Such an active approach could also increase employee engagement with the company's core values (Bakker & Schaufeli, 2008).

### **Limitations and Directions for Future Research**

In addition to the theoretical and practical implications, the study has several limitations that must be acknowledged to establish the applicability of findings to real-life workplace situations. Firstly, due to the time constraints stemming from the nature of a dissertation project, the present study could only expose its participants to the social safety campaign once. Therefore, it did not explore the potential long-term effects of such initiatives on the processes underlying harm identification. Companies frequently implement social campaigns for a more extended period (for instance, several months or years), which helps their target audience become more acquainted with the campaign due to repeated exposure to its messages. It would be also beneficial to pretest the designed vignettes to increase their internal validity and ensure that they are sufficiently ambiguous for the research topic.

Secondly, the present study utilized several self-report measures, which raises the issue of a social desirability bias. Therefore, when completing the questionnaires pertaining to the study variables, participants might have distorted their answers to present themselves in a more favorable light. The issue of social desirability is especially salient in the context of reporting attitudes regarding sexism, workplace diversity, and violence (Boring & Delgauw, 2024). Future research would benefit from incorporating a more comprehensive range and type of



measures, including interviews, to assess the study variables. Additionally, since the scale measuring the severity of harm observed in the vignettes consisted of only one item, future studies should employ a more detailed measure, which could reveal nuances in harm assessment.

Thirdly, the present study utilized a sample representative of the WEIRD population: individuals who live in Western, Educated, Industrialized, Rich, and Democratic societies (Henrich et al., 2010). Therefore, when designing future studies, the issue of cultural diversity and inclusivity needs to be addressed so that findings are generalizable to and representative of a broader range of individuals. Cultural differences are especially salient in the organizational context since collectivistic versus individualistic countries may differ in how companies conceptualize and react to workplace mistreatment (Dai et al., 2022). Individuals working in countries characterized by a high power distance (a greater acceptance of "unequal distribution of power, state or wealth") are more likely to depend on their leaders when making decisions and accept abuse by remaining silent about the harmful behaviors in their workplace (Dai et al., 2022, p.2). Regarding the sampling method, the present study relied on convenience sampling instead of random selection, which restricts the findings from being generalizable to the whole target population.

Lastly, it is worth noting that the preliminary statistical analysis revealed that the assumption of normality for the independent-samples t-test was not met. However, such violations should not impact the accuracy of the results since the study sample was sufficiently large (Ernst & Albers, 2017). Additionally, the assumption check for the moderation analysis and linear regression revealed that the assumption of normality was violated for the "Group outing" vignette.

Future research should focus on conducting longitudinal studies in which participants are repeatedly exposed to social safety messages. Such an approach would enable the exploration of the potential long-term effects of such initiatives on the processes underlying harm perception. Additionally, researchers could devise a combination of informational (similar to those implemented in the present study) and behavioral strategies to manipulate the individuals' assessment of harm in ambiguous social interactions.

The next step in the research on the issue of harm perception and its link to social safety campaigns could involve exploring alternative moderators, such as political correctness and field dependence/independence. Political correctness was recently found to be associated with a greater perception of racial discrimination in ambiguous social scenarios and increased moral sensitivity to harm (Lueke, 2023). This suggests that political correctness may not only serve as a mechanism for creating a safe, non-discriminatory environment, but it may also make people overly sensitive to harm by highlighting the division between the potential perpetrators and victims and promoting the “cult of victimization” (Scalcau, 2020, p. 54). Moreover, besides exploring personality dimensions, such as neuroticism, it would be beneficial to investigate the effects of cognitive styles on conflict and harm perception in the organizational context. For instance, the concept of field dependence posits that some individuals rely more heavily on external cues present in their social environment ("social referents") when making judgments about ambiguous interpersonal behavior (Witkin & Goodenough, 1977). Field-independent individuals, on the other hand, are more autonomous in social interactions and more reliant on internal cues when exposed to an ambiguous situation (Witkin & Goodenough, 1977).

## **Appendix A**

**Figure 1**

*Posters of the social safety campaign*



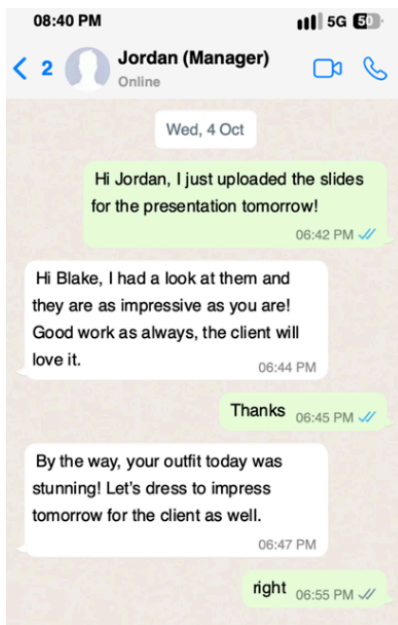
**Figure 2**

*Group outing vignette*



Figure 3

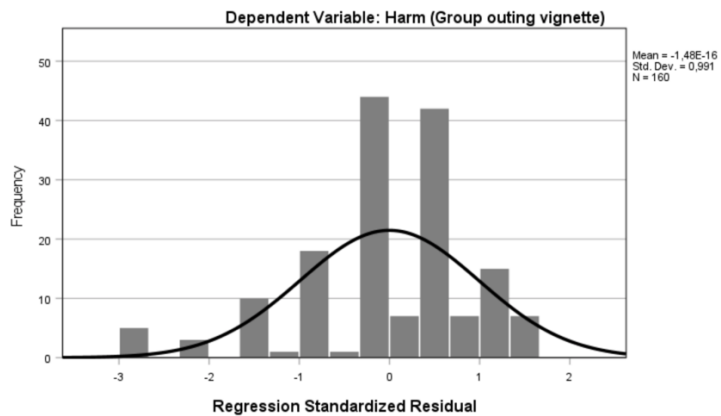
*Clothing vignette*



Appendix B

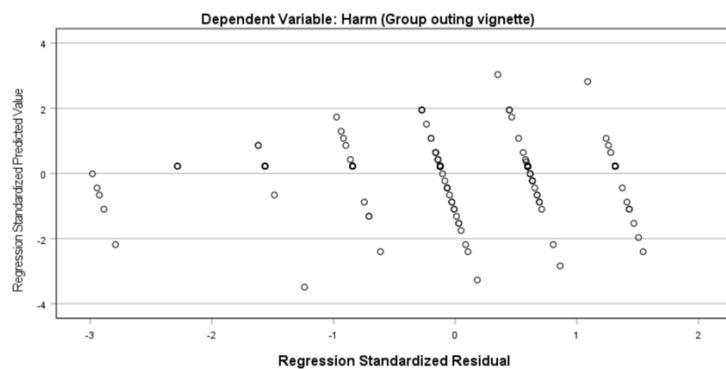
Figure 1

*Histogram of standardized residuals (Group outing vignette)*



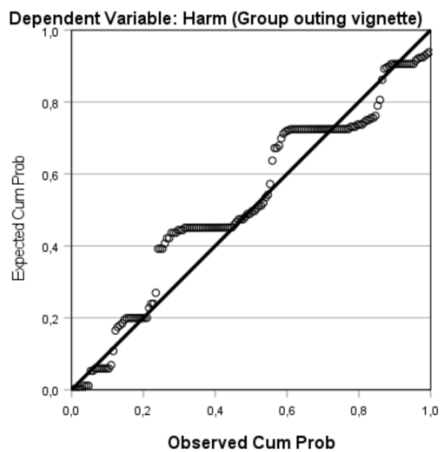
**Figure 2**

*Scatterplot of standardized residuals (Group outing vignette)*



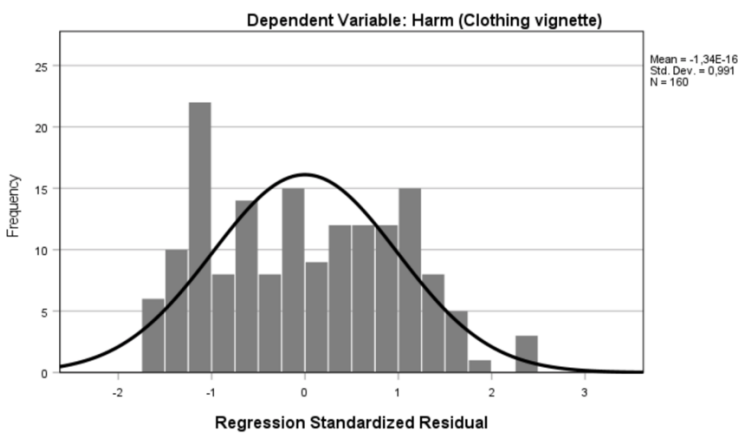
**Figure 3**

*Normal Predicted Probability plot (Group outing vignette)*



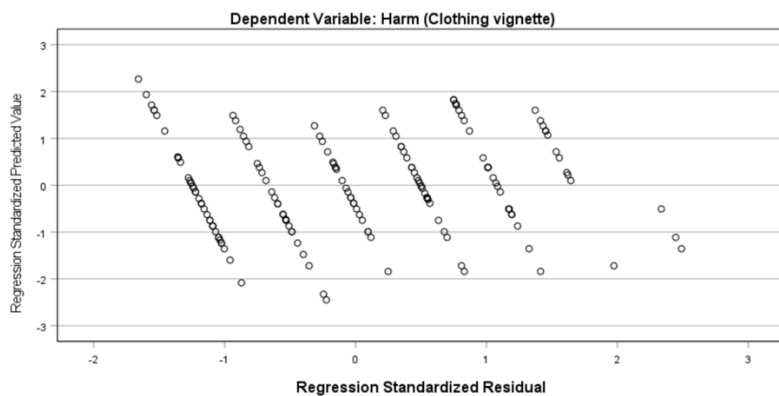
**Figure 4**

*Histogram of standardized residuals (Clothing vignette)*



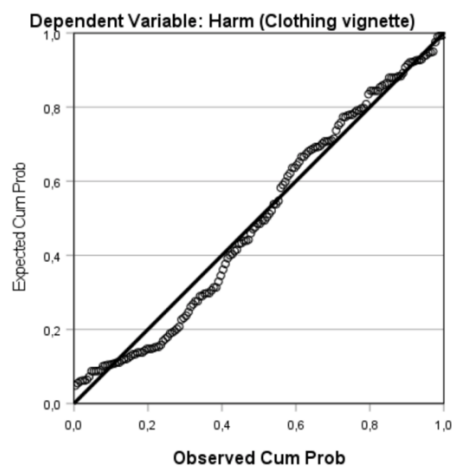
**Figure 5**

*Scatterplot of standardized residuals (Clothing vignette)*



**Figure 6**

*Normal Predicted Probability plot (Clothing vignette)*



## References

Bakker, A. B., & Schaufeli, W. B. (2008). Positive organizational behavior: engaged employees in flourishing organizations. *Journal of Organizational Behavior*, 29(2), 147–154. <https://doi.org/10.1002/job.515>

- Barnett, T., & Valentine, S. (2004). Issue contingencies and marketers' recognition of ethical issues, ethical judgments and behavioral intentions. *Journal of Business Research*, 57(4), 338–346. [https://doi.org/10.1016/s0148-2963\(02\)00365-x](https://doi.org/10.1016/s0148-2963(02)00365-x)
- Bateman, C. R., Valentine, S., & Rittenburg, T. L. (2012). Ethical decision making in a Peer-to-Peer file sharing situation: the role of moral absolutes and social consensus. *Journal of Business Ethics*, 115(2), 229–240. <https://doi.org/10.1007/s10551-012-1388-1>
- Bleske-Rechek, A., Deaner, R. O., Paulich, K. N., Axelrod, M., Badenhorst, S., Nguyen, K., Seyoum, E., & Lay, P. S. (2023). In the eye of the beholder: Situational and dispositional predictors of perceiving harm in others' words. *Personality and Individual Differences*, 200, 111902. <https://doi.org/10.1016/j.paid.2022.111902>
- Bond, S., Tuckey, M. R., & Dollard, M. F. (2010). Psychosocial safety climate, workplace bullying, and symptoms of posttraumatic stress. *Organization Development Journal*, 28(1), 37. <https://psycnet.apa.org/record/2010-06252-003>
- Borghuis, J., Bleidorn, W., Sijtsma, K., Branje, S., Meeus, W., & Denissen, J. J. A. (2020). Longitudinal associations between trait neuroticism and negative daily experiences in adolescence. *Journal of Personality and Social Psychology*, 118(2), 348–363. <https://doi.org/10.1037/pspp0000233>
- Boring, A., & Delfgaauw, J. (2024). Social desirability bias in attitudes towards sexism and DEI policies in the workplace. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.4773184>
- Bunghez, C., Rusu, A., De Houwer, J., Perugini, M., Boddez, Y., & Sava, F. A. (2023b). The Moderating Role of Neuroticism on Evaluative Conditioning: Evidence From Ambiguous



- Learning Situations. *Social Psychological & Personality Science*. <https://doi.org/10.1177/19485506231191861>
- Butterfield, K. D., Treviño, L. K., & Weaver, G. R. (2000). Moral Awareness in business Organizations: Influences of Issue-Related and Social Context Factors. *Human Relations*, 53(7), 981–1018. <https://doi.org/10.1177/0018726700537004>
- Costa, P. T., & McCrae, R. R. (1992). The Five-Factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders*, 6(4), 343–359. <https://doi.org/10.1521/pedi.1992.6.4.343>
- Dai, Y., Li, H., Xie, W., & Deng, T. (2022). Power distance belief and workplace communication: the mediating role of fear of authority. *International Journal of Environmental Research and Public Health/International Journal of Environmental Research and Public Health*, 19(5), 2932. <https://doi.org/10.3390/ijerph19052932>
- Dakin, B. C., McGrath, M. J., Rhee, J. J., & Haslam, N. (2022). Broadened concepts of harm appear less serious. *Social Psychological and Personality Science*, 14(1), 72–83. <https://doi.org/10.1177/19485506221076692>
- Ernst, A. F., & Albers, C. J. (2017). Regression assumptions in clinical psychology research practice—a systematic review of common misconceptions. *PeerJ*, 5, e3323. <https://doi.org/10.7717/peerj.3323>
- Furedi, F. (2016). The cultural underpinning of concept creep. *Psychological Inquiry*, 27(1), 34–39. <https://doi.org/10.1080/1047840x.2016.1111120>
- Gill, A. J., Harrison, A. J., & Oberlander, J. (2004). Interpersonality: Individual differences and interpersonal priming. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 26(26). <https://escholarship.org/content/qt5fz9q55f/qt5fz9q55f.pdf?t=op2kxj>

- Gollwitzer, M., Rothmund, T., Alt, B., & Jekel, M. (2012). Victim sensitivity and the accuracy of social judgments. *Personality and Social Psychology Bulletin*, 38(8), 975–984. <https://doi.org/10.1177/0146167212440887>
- Guo, Q., Sun, P., & Li, L. (2018). Why neurotic individuals are less prosocial? A multiple mediation analysis regarding related mechanisms. *Personality and Individual Differences*, 128, 55–61. <https://doi.org/10.1016/j.paid.2018.02.026>
- Hamre, K. V., Einarsen, S., & Notelaers, G. (2023). Psychosocial safety climate as a moderator in role stressor- bullying relationships: A multilevel approach. *Safety Science*, 164, 106165. <https://doi.org/10.1016/j.ssci.2023.106165>
- Haslam, N. (2016). Concept Creep: Psychology’s expanding concepts of harm and pathology. *Psychological Inquiry*, 27(1), 1–17. <https://doi.org/10.1080/1047840x.2016.1082418>
- Haslam, N., Dakin, B. C., Fabiano, F., McGrath, M. J., Rhee, J. J., Vylomova, E., Weaving, M., & Wheeler, M. (2020). Harm inflation: Making sense of concept creep. *European Review of Social Psychology*, 31(1), 254–286. <https://doi.org/10.1080/10463283.2020.1796080>
- Haslam, N., Tse, J. S. Y., & De Deyne, S. (2021). Concept creep and psychiatrization. *Frontiers in Sociology*, 6. <https://doi.org/10.3389/fsoc.2021.806147>
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
- Henle, C. A., & Groß, M. (2013). What have I done to deserve this? Effects of employee personality and emotion on abusive supervision. *Journal of Business Ethics*, 122(3), 461–474. <https://doi.org/10.1007/s10551-013-1771-6>

- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral And Brain Sciences*, *33*(2–3), 61–83. <https://doi.org/10.1017/s0140525x0999152x>
- Jones, T. M. (1991). Ethical Decision making by Individuals in Organizations: an Issue-Contingent Model. *Academy of Management Review*, *16*(2), 366–395. <https://doi.org/10.5465/amr.1991.4278958>
- Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7–49). Lawrence Erlbaum Associates Publishers.
- Leung, A. S., Wu, L., Chen, Y., & Young, M. N. (2011). The impact of workplace ostracism in service organizations. *International Journal of Hospitality Management*, *30*(4), 836–844. <https://doi.org/10.1016/j.ijhm.2011.01.004>
- Liang, H. (2021). Does workplace bullying produce employee voice and physical health issues? Testing the mediating role of emotional exhaustion. *Frontiers in Psychology*, *12*. <https://doi.org/10.3389/fpsyg.2021.610944>
- Lueke, A. (2023). That’s racist!: Political correctness predicts accusations of racism in ambiguous situations. *Basic and Applied Social Psychology*, *45*(6), 173–191. <https://doi.org/10.1080/01973533.2023.2251636>
- Luminet, O., Zech, E., Rimé, B., & Wagner, H. (2000). Predicting cognitive and social consequences of emotional episodes: the contribution of emotional intensity, the five factor model, and alexithymia. *Journal of Research in Personality*, *34*(4), 471–497. <https://doi.org/10.1006/jrpe.2000.2286>

- Matthews, G. (2008). *Personality and Information Processing: A Cognitive-Adaptive Theory*. In *SAGE Publications Ltd eBooks*(pp. 56–79). <https://doi.org/10.4135/9781849200462.n3>
- McGrath, M. J., Randall-Dziedz, K., Wheeler, M., Murphy, S. C., & Haslam, N. (2019). Concept creepers: Individual differences in harm-related concepts and their correlates. *Personality and Individual Differences, 147*, 79–84. <https://doi.org/10.1016/j.paid.2019.04.015>
- Ormel, J., Bastiaansen, J. A., Riese, H., Bos, E. H., Servaas, M. N., Ellenbogen, M. A., Rosmalen, J., & Alemán, A. (2013). The biological and psychological basis of neuroticism: Current status and future directions. *Neuroscience & Biobehavioral Reviews, 37*(1), 59–72. <https://doi.org/10.1016/j.neubiorev.2012.09.004>
- Reynolds, S. J., & Miller, J. A. (2015). The recognition of moral issues: moral awareness, moral sensitivity and moral attentiveness. *Current Opinion in Psychology, 6*, 114–117. <https://doi.org/10.1016/j.copsyc.2015.07.007>
- Robinson, M. D., Ode, S., Moeller, S. K., & Goetz, P. W. (2007). Neuroticism and affective priming: Evidence for a neuroticism-linked negative schema. *Personality and Individual Differences, 42*(7), 1221–1231. <https://doi.org/10.1016/j.paid.2006.09.027>
- Scalcau, A. (2020). The paradoxes of political correctness. *Theoretical and Empirical Researches in Urban Management, 15*(4), 53–59. <https://www.jstor.org/stable/26955974>
- Slavich, G. M., Roos, L. G., Mengelkoch, S., Webb, C. A., Shattuck, E. C., Moriarity, D. P., & Alley, J. (2023). Social Safety Theory: Conceptual foundation, underlying mechanisms, and future directions. *Health Psychology Review, 17*(1), 5–59. <https://doi.org/10.1080/17437199.2023.2171900>

University of Groningen. (2023a, June 26). *Social safety campaign encourages dialogue*.

Retrieved July 3, 2024, from <https://www.rug.nl/news/2023/06/social-safety-campaign-encourages-dialogue?lang=en>

University of Groningen. (2023b, November 21). *Social Safety*. Retrieved July 3, 2024,

from <https://www.rug.nl/about-ug/policy-and-strategy/social-safety/?lang=en>

Valentine, S., & Gódkin, L. (2019). Moral intensity, ethical decision making, and whistleblowing intention. *Journal of Business Research*, *98*, 277–

288. <https://doi.org/10.1016/j.jbusres.2019.01.009>

Wang, X., Cui, S., Wu, M., Wang, Y., Gao, Q., & Zhou, Y. (2020). Victim sensitivity and its neural correlates among patients with major depressive disorder. *Frontiers in*

*Psychiatry*, *11*. <https://doi.org/10.3389/fpsy.2020.00622>

Witkin, H. A., & Goodenough, D. R. (1977). Field dependence and interpersonal

behavior. *Psychological Bulletin*, *84*(4), 661–689. <https://doi.org/10.1037/0033-2909.84.4.661>

Zhang, Q., Li, X., Gold, B. T., & Jiang, Y. (2010). Neural correlates of cross-domain affective priming. *Brain Research*, *1329*, 142–151. <https://doi.org/10.1016/j.brainres.2010.03.021>