



Hear me out

How to create an open mind
towards moral criticism

Doctoral Thesis

Inga Katharina Rösler

Hear me out: How to create an open mind towards moral criticism

Luister naar mij: hoe creëer je een open mind voor morele kritiek?

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 1 april 2022 des ochtends te 10.15 uur

door

Inga Katharina Rösler

geboren op 2 november 1989
te Hamburg, Duitsland

Promotor:

Prof. dr. N. Ellemers

Copromotor:

Dr. F. Van Nunspeet

ISBN: 978-94-6423-719-1

Cover Design: M. Alexander Rösler

Printing: www.proefschriftmaken.nl

Copyright © 2022 by Inga Rösler

Table of contents

Chapter 1	General Introduction	5
Chapter 2	Don't Tell Me About My Moral Failures but Motivate Me to Improve: Increasing Effectiveness of Outgroup Criticism by Criticizing One's Competence	30
Chapter 3	Falling on Deaf Ears: The Effects of Sender Identity and Feedback Dimension on How People Process and Respond to Negative Feedback – an ERP Study	60
Chapter 4	Show Me That You Mean Well: The Beneficial Effects of Communicating the Intention to Help when Delivering Moral Critique as an Outgroup Member	98
Chapter 5	General Discussion	137
	Supplementary Materials	167
	References	210
	Dutch summary	230
	Acknowledgments	236
	About the author	240
	KLI dissertation series	241

The background is a teal watercolor wash with a white torn paper effect at the bottom. The text is centered in the upper half of the image.

CHAPTER

1

General Introduction

To be able to live together peacefully in societies and communities, humans have developed laws, social norms, and moral values. These three types of rules form a sort of invisible ‘code’ that specifies what is ‘right’ and what is ‘wrong’ behavior in their society or community. To check whether people adhere to these laws, law enforcement organizations such as the police work on discovering criminal actions. To check whether people adhere to social norms and moral values, on the other hand, no such instances exist. Instead, we often aim to regulate behavior and check whether people adhere to social norms by giving feedback to each other (e.g., Ellemers et al., 2013; Ellemers & van den Bos, 2012; Ellemers & Van der Toorn, 2015). For instance, on a societal level, a government might reward civilians with tax relief for showing cooperative behavior by donating to a charity, which can be seen as a type of positive feedback. Or the government might punish civilians with large fines for selfish behavior such as engaging in tax evasion, which can be seen as a type of negative feedback. On a personal level, we also give feedback to each other. We ‘reward’ a friend on social media with ‘hitting the like’ button of a post showing that they donated to a charity. Or we ‘punish’ a stranger on the train with an angry look for not wearing a facemask in times of a pandemic. We use positive feedback to encourage people to repeat a certain type of behavior (e.g., donating) and we use negative feedback to encourage people to change their behavior (e.g., start wearing a mask on the train). Positive feedback is often received favorably and can motivate people to engage in similar behaviors again. Negative feedback, on the other hand, does not always have the intended effect of changing people’s behavior. Rather, it oftentimes makes people defensive. Consequently, they dismiss negative evaluations of themselves and their behavior, for instance, by labelling the critique as unconstructive or unfair. To demonstrate, consider the following citation from one of my participants (from the research described in Chapter 2) where they describe why they were criticized at their workplace for not doing their share of work.

The critic said: *“That I was not pulling my weight at my work and that morally what I was doing was wrong. I felt it was harsh and unfair.”* – Participant
Chapter 2 Study 1.1

Rather than acknowledging that the person who criticized them might have a point, they dismissed the critique by labelling the feedback as harsh and unfair. Social psychological research has shown that such defensive reactions from receivers of criticism are not uncommon when the criticism refers to the (im)morality of a person (Gausel & Leach, 2011; Täuber et al., 2018; Täuber & van Zomeren, 2013). What often happens in situations where a person is criticized for not acting moral (e.g., not telling the truth) is that they try to deflect blame by denying having engaged in any questionable behavior or by redirecting the blame back to us. An example for such redirection of blame is the above-mentioned example where my participant labelled the feedback as unfair (which it might as well have been, we can't know that for sure). To explain, calling feedback unfair implies that the person criticizing made a mistake or wanted to intentionally hurt by criticizing, thus calling feedback unfair redirects blame back to the critic. Another example of redirecting blame and not taking responsibility for the own immoral actions comes from modern politics. On January 6, 2021, Donald Trump, the former president of the United States, called for his supporters to discount votes that confirmed Joe Biden as the winner of the 46th American presidential election. His speech then resulted in his supporters storming the Capitol in Washington, where the Congress was debating the confirmation of these votes. Even though he received criticism on this dangerous behavior from all over the world, he did not publicly admit being responsible for the violent and deadly storming of the government building. Rather, he responded defensively by dismissing and ignoring all criticism.

Not dealing with such negative feedback on our moral self can guard our own self-views and protect us from having to deal with negative emotions that arise when being criticized such as shame or regret (Gausel, 2013; Gausel & Leach, 2011; Van der Lee et al., 2016). Perceiving ourselves as a 'good person' is important to many of us and our morally questionable behavior stands in contradiction to this. Moreover, signaling immoral and

rule-breaking behavior might lead to social exclusion from the communities we belong to (Van der Lee et al., 2017). Social exclusion is painful and there is research suggesting that it might even be felt as physical pain (Eisenberger & Lieberman, 2004). To avoid feeling excluded or rejected can therefore be a strong motivator to defend the self when facing moral criticism. However, this represents a problem. By defending ourselves and redirecting blame, the criticism is dismissed and loses its power to encourage moral growth and positive change.

1.1 | Comparing criticism on one's morality to criticism on one's competence

The research in this dissertation is inspired by these timely issues of getting defensive when being criticized for one's morality. Defensive responses often occur when one's morality is criticized because being seen as a moral person is important to people, often even more important than being perceived as, for instance, competent (Brambilla & Leach, 2014; Ellemers et al., 2008; Goodwin et al., 2014; Leach et al., 2007; Täuber et al., 2015, 2018; Tetlock, 2002; Van der Lee et al., 2016). Since defensive reactions often entail that people dismiss criticism (Gausel et al., 2012; Gausel & Leach, 2011; Giner-Sorolla, 2012), the first main aim of the research in this dissertation was to find ways to encourage people to be more open towards receiving negative feedback on their morality. To this end, I investigated how people respond and deal with such feedback (Chapters 2 and 3) and tested two strategies to increase the impact of such feedback (Chapters 2,3, and 4). Both strategies were developed to provide easily to be applied tools for feedback senders who aim to encourage others to change their moral behavior. The first strategy entailed shifting the focal point of a feedback message from addressing one's morality to one's competence to reduce defensiveness and encourage motivation to improve in the person being criticized (Chapter 3). The second strategy referred to having feedback senders communicate their good intentions when giving feedback to make people more open to receiving negative feedback on their morality (Chapter 4).

For example, for the research in Chapter 2, I compared situations where participants were criticized for their morality at their workplace with situations where they were criticized for their competence (Chapter 2). One of the main predictions for this research was that criticism on one's competence, rather than on one's morality, should decrease defensiveness and increase people's motivation to improve based on the feedback they received. This prediction built on previous research showing that when people judge others, positive (vs. negative) judgments weigh in more when the competence of another person is judged and negative (vs. positive) judgments are more important when the morality of a person is judged (Martijn et al., 1992; Skowronski & Carlston, 1987). That is, because they suggest that criticism of one's morality has a higher potential to hurt someone's image than criticism of one's competence. Therefore, people might be less defensive when being criticized for their competence compared to their morality. Moreover, other research has shown that people also feel more capable of coping with their past shortcomings when these relate to their competence as compared to their morality (Van der Lee et al., 2016), possibly also making them less defensive to such criticism.

To demonstrate what this could look like in real life, please consider the following criticism situation. It is the same situation as the one you just read, but I've slightly adapted it in such a way that the competence of the participants, rather than their morality, is the focal point of the critique.

"That I was not pulling my weight at my work and that because of that I was perceived as lazy. I felt it was harsh, but I tried to put in more effort." –

Fictional response for demonstration purposes

As a reminder, the previous example was: The critic said: *"That I was not pulling my weight at my work and that morally what I was doing was wrong. I felt it was harsh and unfair."* – Participant Chapter 2 Study 1.1

This may seem like only a small adjustment – referring to the lack of competence rather than a lack of morality. But such reframing may lead to a vastly different response of the person receiving such a message. To explain, criticism of one's morality may be perceived as referring to something that

is difficult to change, such as being an immoral person in general. Criticism of one's competence, on the other hand, may offer the person criticized an easier way to address the shortcoming by changing their behavior (e.g., Skowronski & Carlston, 1987; Täuber & van Zomeren, 2013; Van der Lee et al., 2016). That being said, being criticized for being incompetent, rather than immoral, will most likely still evoke some negative emotions in people. But research shows that it also helps people to better cope with the negative evaluation of themselves and that it might motivate them to improve in the future (Täuber et al., 2018; Täuber & van Zomeren, 2013; Van der Lee et al., 2016)

What this can look like, is demonstrated with a second citation from a participant of the research in Chapter 2. The participant explains how the criticism of their competence helped them to improve their professionalism at the workplace:

“My coworker saw me with my hands in the pockets while we were performing an installation of an air conditioner at some client’s house, and he told me to remove them from there, because it was perceived as unprofessional, and also as if I was doing nothing there. I must admit he’s been right, and I appreciated his criticism.” – Participant Chapter 2 Study 1.1

Rather than blaming the feedback sender for being unfair, here the participant accepts the criticism and is motivated to change their behavior in the future. Of course, these two examples are about very different situations. But they nicely demonstrate the underlying psychological processes that play a role in comparing criticism on one's morality to criticism on one's competence. They demonstrate that the effectiveness of negative feedback, whether the receiver of the feedback is getting defensive or whether they are motivated to change their behavior, might depend on factors of the situation and how the feedback message is framed. To clarify this point further, consider the following, again adapted, response, now referring to the morality of the person.

“My coworker saw me with my hands in the pockets while we were performing an installation of an air conditioner at some client’s house, and he told me to remove them from there, because it was perceived as having no work ethic, and also as if I was letting my coworker do all the work.” – Fictional response for demonstration purposes

Here, the participant is criticized for the unethical and uncooperative behavior and the immorality (rather than incompetence) of the participant is emphasized. One can imagine that the participant here would have not admitted that the critic was right, but rather would have responded defensively to protect their image of being a good coworker.

1.2 | Comparing criticism delivered by ingroup members to criticism delivered by outgroup members

Another important characteristic of the feedback situation that might influence the degree of receiver’s defensiveness is related to the characteristics of the actual person who is giving the feedback (Esposito et al., 2013; Hornsey & Imani, 2004; Hornsey et al., 2002). To elucidate which characteristics might influence how feedback is received, it is important to consider the structure of our social world. It is divided into many different groups, for instance, by countries, religions, or cities. Most of us identify with some of these groups (i.e., ingroups) and feeling similar and close to members of a group can give us self-esteem and feelings of belonging (i.e., Ellemers, 2012; Tajfel & Turner, 1979). Groups that we don’t identify with are referred to as *outgroups*. There is a large body of research in both social psychology and neuroscience showing that such group-memberships influence how we perceive and react to a person when interacting with them (Amodio & Cikara, 2021; Cikara & Van Bavel, 2014; Ellemers, 2012; Ellemers et al., 2013; Tajfel & Turner, 1979). In the context of feedback, research has shown that criticism coming from ingroups is perceived more favorably than feedback coming from outgroups (intergroup sensitivity effect, Hornsey et al., 2002). The intergroup sensitivity effect refers to group-level criticism, meaning that your ingroup is criticized by an outgroup. However, it isn’t clear whether something similar happens when you as a person, rather than your ingroup, are criticized by someone belonging to a different group (i.e., individual-level criticism). Because we

interact with ingroup and outgroup members daily, for instance at the (digital) workplace, it is important to research whether criticism from outgroup members on our behavior or character is also less favorably received than when it is delivered by ingroup members. Consider, for instance, two colleagues from different work departments discussing a financial deal that one of the two has recently closed (i.e., the seller). Such deals are frequently closed in the seller's department but would be seen as morally questionable in the other department and by the public. Would the seller accept criticism from their (outgroup) colleague questioning such deals for ethical reasons? And would the response depend on the fact that they work in different departments and therefore perceive each other as outgroup members? And finally, if so, is there a way to prevent such group effects from occurring? Finding answers to these questions is important to prevent fraudulent behavior in groups who have developed morally questionable norms such as closing risky financial deals. In such groups, group members are often afraid to speak up about morally questionable behavior because they are afraid of being ostracized (e.g., Cavazza et al., 2014; Ellemers & van Nunspeet, 2020; Thau et al., 2015; Van der Lee et al., 2017). Consequently, immoral behavior often has to be called out by outsiders. However, as outlined above, outsiders are often not listened to (Esposito et al., 2013; Hornsey & Imani, 2004; Hornsey et al., 2002). The second main aim of the research in this dissertation was therefore to find ways to make people more open to receiving criticism from outgroup members. I investigated how people respond to and deal with criticism coming from ingroup and outgroup members (Chapters 2, 3, and 4) and tested whether focusing criticism on competence (vs. morality, Chapters 2 and 3) and communicating having good intentions when delivering negative feedback (Chapter 4) can decrease such group-effects.

To summarize, the two central questions addressed in this dissertation are: 1) How do people deal with being criticized for their moral behavior and how can we make people more open to such criticism? and 2) Do people react differently to criticism if it is delivered by an outgroup compared to an ingroup member – and if so – how can we intervene to prevent this from happening? Contrary to most research in the field of moral psychology, but in the spirit of recent research on moral evaluations (Van der Lee, Ellemers, & Scheepers,

2016; Van Nunspeet, Derks, Ellemers, & Nieuwenhuis, 2015; Van Nunspeet, Ellemers, Derks, & Nieuwenhuis, 2014), the subject of investigation is therefore the person being evaluated, rather than the person who is doing the evaluations (e.g., moral reasoning, moral judgments, Ellemers et al., 2019). In the following I will shortly elaborate on the importance of such research and why morality plays a central role in the regulation of behavior in a society.

1.3 | The role of morality (feedback) in society

Why is it important to research how people respond to feedback coming from different sources and referring to their (im)morality? One obvious reason is that moral criticism is omnipresent: At the workplace, in the economy, and in our homes. We often observe others being – or are ourselves – criticized and are not aware of how characteristics of the context (e.g., who is giving the feedback) is influencing others' and our own reactions to being criticized. This can lead to faulty assessments of a situation, such as inferring that senders have harmful intentions when giving feedback. For example, a colleague from another work team criticizes someone's presentation and the presenter infers that this colleague did so out of spite or own self-enhancement rather than the actual intention of wanting to help them improve their presentation style. As another example, a letter from a regulatory agency including criticism of possible fraudulent behavior of a bank in the financial sector raises defensive reactions in employees of the bank because their integrity is being questioned. Or, criticism coming from someone with different political views might be disregarded because one feels the views and actions of a political party are 'morally wrong' and unacceptable.

Another, more far reaching, reason for why it is important to investigate how we can make people more open to being criticized for their morality and by outgroup members, is because it can help with preventing future conflict in our societies. As mentioned in the beginning of this introduction, social norms and moral values act as a guide so people know which behavior is 'right' or 'wrong'. Interestingly, and against the common notion that our own moral values always refer to universal moral principles (e.g., "Be reliable and sincere"), they can also be influenced by the social groups we belong to (Ellemers, 2017; Ellemers et al., 2013; Ellemers & Van der Toorn, 2015).

Moreover, social groups, religions, or societies can have different ways in what it means to act in line with such universal moral principles and therefore how specific 'moral behavior' is defined. For instance, in one group "being reliable and sincere" might translate into a social norm where people should voice their own opinions and stand points, whereas in another group this might translate into having to be loyal to the group and therefore conforming to these group opinions (Ellemers & Van der Toorn, 2015). Or, for one group it is moral to help refugees in need by letting them into the country and providing them with food and shelter, whereas another group might see this as morally wrong because resources should be allocated first to people in need in their own society (e.g., to the homeless or poor). Disagreement on such sensitive topics between groups can easily lead to value conflicts. Value conflicts and 'othering' (i.e., seeing an outgroup as essentially different than the self), on the other hand, can feed polarization and dispersion in a society (e.g., Finkel et al., 2020). As our societies become more and more diverse, it becomes increasingly important to find new ways of how we can peacefully communicate with each other, even if group members communicate about differing moral values. We need to find ways to be open to the perspectives of others, regardless of their differing moral convictions. To prevent conflict and stimulate cooperation, it is therefore essential to understand how we can be more open to receive criticism even if it refers to our own morality or if it is coming from outgroup members.

1.4 | Interdisciplinary research methods

With the research described in this dissertation I aimed to contribute to societal issues of diversity and polarization and to find new ways to increase communication between groups by investigating how we can make people more open to receive negative feedback. For this, I combined theory, methods, and measures from social psychology and neuroscience. Previous research, to my current knowledge, on intergroup criticism and criticism on people's morality has only been conducted in the field of social psychology (Esposito et al., 2013; Täuber et al., 2018; Täuber & van Zomeren, 2013). More specifically, this means that we have some knowledge about how people consciously respond to being criticized by outgroups and for their morality. However, we do not know whether such social context factors (e.g., being

an outgroup member) also influence unconscious processes that take place before we can even deliberately reflect on being criticized. This is where using neuroscientific methods can help as they can give new insights into the cognitive processes that occur before we are able to consciously respond to a situation (Ellemers & van Nunspeet, 2020). For instance, it can shed light onto what type of criticism receives and holds our attention and which type is not attended to (also see an elaboration on this in the next paragraph). Thus, the combination of these fields is promising as it adds another level to the results previously established in social psychological literature.

Another field that influenced my research was economics. Economic research developed a very precise and mathematical way to measure human decision making (i.e., game theory, for a review on how it is applied in social psychology see Krueger et al., 2020). The behavior in these games is believed to be indicative of how humans make decisions in everyday life. However, these games have also been criticized for their external validity (i.e., in which way they reflect how people make decisions in the real world, e.g., Galizzi & Navarro-Martinez, 2019; Levitt & List, 2007). Real life decision making or social behavior, especially if it concerns moral topics, might be influenced by more factors than what is traditionally tested in economic games. It is often influenced by 'invisible' characteristics of the social context, group processes, or group norms (Ellemers, 2017; Ellemers & van Nunspeet, 2020, but see Otten et al., 2020 for an elegant way of how norms can be implemented in economic games). In social psychology, researchers, therefore, make use of cover stories to make an experiment as close as possible to a real-life situation, to achieve greater external validity. In the research of this dissertation, I made use of a combination, creating new experimental paradigms that measure economic behavior (e.g., monetary contributions to a common good in Chapter 3) but embedded such paradigms in a relevant context by using cover stories. Moreover, in Chapter 2, participants reflected on their own autobiographical experiences of being criticized at the workplace, ensuring external validity.

In the following, I will shortly elaborate on the process of combining approaches from different disciplines and finding similarities in their assumptions and

scientific arguments. For the current dissertation, the common denominator is the argument that the social context in which feedback is embedded can influence how we respond to the feedback. It influences not only our deliberate responses but can shape our attentional processing of feedback messages before we even become aware of their meaning.

1.5 | The social context influences the cognitive processing of and responses to feedback

Most of the questions we are interested in today have been already discussed thousands of years ago. For instance, the ancient philosopher Aristotle discussed already around 350 BC how humans can become ethical and how they should ideally live together in a community (e.g., *Nicomachean Ethics*, 350 BCE, Aristotle et al., 2009). In his famous work 'Rhetorik' (Aristotle & Roberts, 2004) he argued that to influence others with your speech, you have to be perceived as a good person. Moreover, he states that it is difficult to motivate others to become a good person using only speech (350 BCE, Aristotle et al., 2009). Thus, this work offers a first insight into why it is important to consider social context information such as being perceived as a good person when trying to motivate others to behave morally. But he also already speaks of the difficulties to do so. Much later, but building on Aristotle's work, an expert on rhetorical theory, Kenneth Burke takes this argument further (Burke, 1969). He discusses the importance of identifying with a speaker. Identifying with the (needs of a) speaker might increase the influence the speech has on you as an audience member. In communication science, early models of communication argue something similar: Sender characteristics (i.e., who is sending a message) and message content (i.e., what is said) influence the communication process (Berlo, 1960; Lasswell, 1948). These early works demonstrate that people have long thought about how the identity of a feedback sender might influence how we respond to the feedback. In social psychology, this has been investigated with literature on the intergroup sensitivity effect (Hornsey & Imani, 2004; Hornsey et al., 2002). However, as previously mentioned, this line of research investigated groups (e.g., outgroups) criticizing other groups (e.g., ingroup) and it is not known whether similar processes occur at the individual level, when our own behavior is criticized, rather than our ingroup.

Early work in cognitive psychology has demonstrated that context information can influence the way we perceive and visually process things. For example, in an experiment (Palmer, 1975) participants were first presented with a contextual scene (e.g., kitchen) and afterwards with an object that was either appropriate for this context (e.g., bread loaf) or not appropriate for the contextual scene (e.g., drum). There was also a condition where objects were presented without seeing a contextual scene. Participants recognized ‘appropriate’ objects more easily than ‘inappropriate’ objects (as indicated by a higher probability to identify objects as correct), demonstrating that contextual information can influence how we perceptually process objects. More recently, neuroscientific research has investigated whether such context effects also occur in social feedback situations (Schindler et al., 2014, 2020; Schindler & Kissler, 2018). Schindler and colleagues (2014, 2018, 2020) investigated how participants attentionally and visually process feedback messages delivered by different types of feedback senders. For instance, whether feedback was delivered by a computer or by a human. In these studies, participants’ immediate brain responses to receiving feedback was measured by recording an electroencephalography (EEG). This method can be used to record spontaneous electrical activity on the scalp while participants view stimuli in an experiment. From this recorded activity, researchers can then extract event-related potentials (ERPs) which represent averaged EEG signal and refer to stimulus-(or response-)locked fluctuations in the electrical activity. These ERPs are believed to represent cognitive processes associated with the processing of a stimulus (or response to a stimulus, depending on the experimental paradigm).

In the studies on the effects of social context on the processing of feedback, the authors investigated how the type of feedback sender affected early ERPs related to perceptual attention (i.e., P200, e.g., Carretié, Martín-Loeches, et al., 2001; Kissler et al., 2006) and later ERPs related to higher level processing of stimuli such as paying sustained attention to a stimulus (i.e., P300, LPP, Lang et al., 1997; Polich, 2007; Schupp et al., 2004, 2006). According to the framework of motivated attention (Lang, Bradley, & Cuthbert, 1997; Lang & Bradley, 2013) these ERPs can give insight into differences of self-relevance attached to stimuli. In other words, these neuroscientific

indicators can tell us whether participants choose to pay sustained attention to certain stimuli over others because these stimuli are more relevant to the self. Using these neuroscientific markers, Schindler and colleagues (2018, 2020) demonstrated that participants paid more sustained attention to feedback coming from more relevant senders (i.e., from a human) than from less relevant senders (i.e., from a computer).

For the research described in Chapters 3 and 4, I made use of these neuroscientific measures to gain new insights into how social context factors influence how people process and respond to feedback. A well-known methodological problem in social psychological research is that participants' responses in a study can be biased because they are responding in a way that they believe is socially acceptable, rather than sharing how they truly think or feel about an experience (i.e., social desirability, e.g., Nederhof, 1985). This is especially relevant for the current research, where participants report on being criticized at the workplace (Chapter 2) and receive negative evaluations of their moral character (Chapters 3 and 4). As previously described, these situations are very painful to experience for people and often elicit defensive responses such as dismissing feedback (Gausel, 2013; Gausel & Brown, 2012; Gausel & Leach, 2011; Täuber & van Zomeren, 2013; Van der Lee et al., 2016). It is therefore difficult to capture actual and truthful responses when investigating how people respond to being criticized for their morality. Using ERPs constitutes an elegant way to resolve these problems because it can give insight into how stimuli are cognitively processed. This may offer insights into experiences that people may not be aware of or are reluctant to report (e.g., Ellemers & van Nunspeet, 2020).

Using neuroscientific markers for investigating the social psychological questions that inspired the research in this dissertation led to new insights which might spur exciting new research avenues. For instance, findings of Chapter 3 demonstrate that not just contextual scenes or the 'humanness' of a feedback sender (rather than a computer as feedback sender) influence how we attentionally process objects or feedback, but that also more subtle context information such as the social group-membership or the social dimension addressed in a feedback message can influence our unconscious

processing. For a more detailed discussion on such future research avenues please see Chapter 5.

In conclusion, using perspectives and research methods from different disciplines can give new and exciting insights into how social information unconsciously influences what draws and holds our attention. In the following I will give an overview of how this combination of different methods was translated into the different empirical chapters of this dissertation.

1.6 | Overview of dissertation

The aim of the research in this dissertation was to understand how two social context factors influence different psychological processes that occur when people receive feedback. More specifically, I investigated how the social group-membership of a feedback sender and the social dimension addressed in a feedback message influence *information processing* of feedback, *behavioral responses* to feedback, and *reflection* on receiving feedback in the past (see Figure 1.1). Across chapters, I made use of a variety of different methods: Qualitative (i.e., content coding of autobiographical data), quantitative (i.e., analyzing self-report data and behavior in the experiment), and neuroscientific (i.e., analyzing ERPs) measures. Figure 1.1 shows a schematic illustration of which measures are used in each chapter to investigate the relationship between social context factors and the psychological processes (under ‘level of examination’) of receiving feedback. They represent different stages of receiving, reacting to, and reflecting on feedback. In the following, I will use an example to demonstrate how all these stages may play a role in the reception and response to outgroup criticism on people’s morality.

Each person has their *own values and beliefs* (e.g., opinions about an outgroup) that they bring to a social feedback situation. For example, a participant from Chapter 2 describes an outgroup member from whom they received criticism at the workplace as follows:

“This person was a programmer. He looked down on me [...] a glorified graphics designer. What he was working on was far more complicated than what I was doing. He was very intelligent albeit arrogant and unforgiving.” – Participant Chapter 2 Study 1.1

Holding such beliefs about a person, and their group-membership, might influence reactions to receiving criticism from a person. For example, they may influence how *information* coming from feedback senders *is received (or processed)*. One can imagine that the participant was reluctant to listen to the criticism by the programmer (i.e., outgroup member) and might not have even paid attention to what was said. The participant might have paid more attention to the same feedback if it was coming from their colleague who holds the same profession and belongs to an ingroup.

Not listening and not paying attention to feedback will most likely result in the participant not remembering what was exactly said. Moreover, the participant will probably also not *change their behavior* in line with what the critic suggested. In other words, the feedback is dismissed and loses its effectiveness. In the case the critic had helpful suggestions, the participant therefore misses out on an opportunity to grow (e.g., Hornsey & Imani, 2004).

To conclude, it is important to consider all these psychological processes when investigating how people respond to feedback. Each chapter in this dissertation therefore had a slightly different focus on which psychological processes it investigates. The measures used in each chapter were selected to give the best insight into these processes. For example, to investigate participants' reflections of past criticism experiences, I took a typical social psychological approach by asking people to reflect on their autobiographical experiences in Chapter 2. In Chapter 3 and 4, I had a closer look at the information processing stage by investigating how participants cognitively process feedback in the actual moment of receiving it.

In the following, I will give an overview of the three empirical chapters of this dissertation, which measures were used in each chapter, and how each measure contributes to the understanding of how social context factors influence responses to receiving and reflecting on feedback. This overview is divided into two sections. The first section describes the research which tested the first strategy to increase the receipt of negative feedback: Making one's competence (vs. one's morality) the focal point addressed in a feedback message (Chapters 2 and 3). The second section addresses the second strategy: Communicating the intention to help when delivering negative

feedback messages on one's morality (Chapter 4).

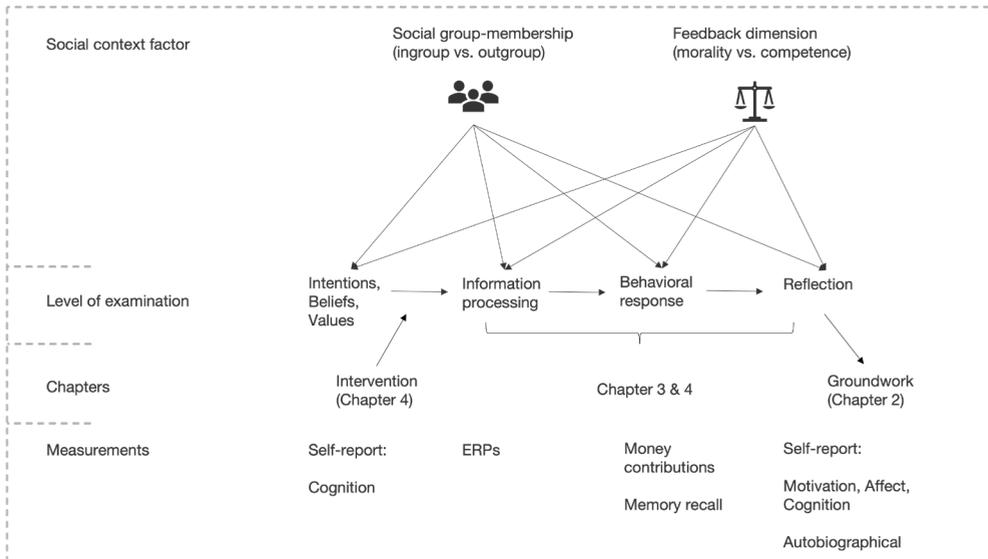


Figure 1.1. Schematic illustration of the mixed method approach of the research in this dissertation. Combining insights from different stages of information processing can lead to innovative insights into how and why some feedback messages are more effective and receive more attention compared to others. The figure shows how the two social context factors ‘social group-membership of feedback sender’ and ‘feedback dimension’ examined in this dissertation relate to different psychological stages of how people process and react to feedback. The figure displays the psychological processes investigated in each chapter in this dissertation and which measures were used. Whereas self-report measures can give insight into how people explicitly respond to and reflect on receiving feedback, ERP measures can give insight into how social context factors affect people’s attention more implicitly. Combining ERP measures with self-report and behavioral measures can give insight into how one stage (e.g., information processing) affects later stages (e.g., behavioral response).

1.7 | Improving the impact of outgroup criticism on people's morality

1.7.1 | Strategy 1: Focusing criticism on competence (vs. morality)

1.7.1.1 | Chapter 2: Reflection on past criticism at the workplace

Chapter 2 represents the starting point and groundwork for the following two empirical chapters. I started investigating how people reflect on being criticized at the workplace ('reflection' stage, see Figure 1.1). I aimed to get a first overview of how people deal with criticism at the workplace and which underlying psychological (i.e., motivational, cognitive, affective) mechanisms influence the effectiveness of criticism. Specifically, I examined whether people are more reluctant to accept criticism from an outgroup member at the workplace (e.g., a colleague from another department) compared to from an ingroup member (e.g., a colleague from the same department). Moreover, as a first step to examine whether people are more receptive when a criticism message is addressing their competence rather than their morality, I compared how people respond to these two types of criticism at their workplace. An example for criticizing someone's morality at the workplace would be to call an employee *disloyal* for letting colleagues do all the work. An example for criticizing someone's competence would be to call an employee *lazy* for letting colleagues do all the work. Both criticism messages address the same behavior: An employee is not putting enough effort into their work.

Reframing criticism messages slightly so they do not refer to the morality of a person, but to their competence, might make people more open to such messages. Research has shown that people are better able to cope with criticism referring to their competence (vs. morality, Van der Lee et al., 2016). This may be caused by a decreased concern about their social image for showing incompetent (vs. immoral) behavior (e.g., Martijn et al., 1992; Skowronski & Carlston, 1987; Täuber et al., 2018; Täuber & van Zomeren, 2013; Tetlock, 2002; Van der Lee et al., 2016, 2017). In other words, people may be more open to—and less defensive towards—criticism messages addressing their competence, instead of their morality at the workplace.

I thus compared feedback addressing the two fundamental dimensions people use to judge others (i.e., the Vertical dimension for 'getting along',

including communion/warmth and morality; and the Horizontal dimension for 'getting ahead', including agency and competence; Abele-Brehm, Ellemers, Fiske, Koch, & Yzerbyt, 2020; Koch, Yzerbyt, Abele, Ellemers, & Fiske, 2020). For this, I used traditional social psychological methods of self-report questionnaires, analyzing both qualitative (i.e., autobiographical memories) and quantitative (i.e., self-report) data.

For the analyses of the qualitative data, I content coded participants' answers to open-ended questions. Content coding is a qualitative method, where the researcher develops a coding scheme and categorizes instances (e.g., a participants' answer on an open-ended question) into different codes (Krippendorff, 2004). Using this type of method can give insights into what kind of criticism situations people experience at their workplace. When I asked people to recall a situation in which they were criticized for their morality, they most often reported an event in which they failed to show cooperative behavior (e.g., being disloyal by letting colleagues do all the work). When criticized for their competence, these types of situations were most often related to the work outcome such as making mistakes or forgetting to complete a task. Content coding is a descriptive, rather than predictive research method. This means that it is difficult to detect patterns in the data such as underlying mechanism that influence reactions to being criticized. I, therefore, combined the content codings with self-report questionnaires and quantitative analyses.

To understand what kind of underlying mechanisms influence whether criticism is effective, meaning whether people accept it and change their behavior accordingly, participants were asked to fill out survey questions using Likert scales. To be able to capture psychological mechanisms related to the questions in this dissertation, I developed new scale measures. For this, I used a combination of items used in previous research (Hornsey et al., 2004) and creating new items. I then used factor analyses to find relevant underlying factors. I ended up with three main scales that stand for the psychological constructs of motivation (i.e., motivation to improve based on criticism), cognition (i.e., making negative assumptions about a critic), and affect (i.e., divided into negative, positive, and moral emotions after being

criticized). I then used these measures in mediation analyses. This method can help to understand how the effects of our two independent variables of interest (i.e., the group-membership of a critic and the dimension [morality vs. competence] of criticism) on whether criticism was effective are caused by an affective, cognitive, or motivational response. These analyses were the groundwork for the other experiments as they gave insight into the underlying mechanisms of our effects of interests. I found that criticism delivered by outgroup (vs. ingroup) members at the workplace was less effective as participants less often accepted this criticism and less often changed their behavior based on this criticism. This relationship was mediated by the defensive response (i.e., making negative assumptions about the person criticizing) and the motivational response (i.e., motivation to improve based on criticism) towards being criticized. In other words, because people are more defensive towards, and less motivated to improve based on, criticism coming from outgroup (vs. ingroup) members, they are more reluctant to accept the criticism and they change their behavior less often after receiving the criticism. Moreover, I found that criticism on participants' morality (vs. competence) is less effective because people get more defensive and are less motivated to improve when receiving this type of criticism. Please note that for full causality claims in mediation analyses, mediators have to be experimentally manipulated rather than merely measured as in this first empirical chapter (Fiedler et al., 2018). This is what was done in Chapter 4 by manipulating one of the mediators: The negative assumptions participants make about the feedback senders. The insights from this first chapter then guided the development of the follow up studies in Chapters 3 and 4.

1.7.1.2 | Chapter 3: Receiving feedback in the actual moment

Asking participants to reflect on their own experiences is a good starting point, but it cannot capture how participants respond to being criticized in the actual moment. The main aim of the third chapter of this dissertation was thus to find out whether similar processes as in Chapter 2 occur when being criticized in the actual moment. Here, I investigated how participants react to being criticized on the spot, now experimentally manipulating feedback on one's morality and one's competence (in these studies participants receive both negative [i.e., criticism] and positive feedback) and from either ingroup

or outgroup members. Building on the results of Chapter 2, I tested whether 1) people (unconsciously) pay more attention to feedback messages from ingroup (v. outgroup) and feedback on their competence (vs. morality). To this end, I analyzed self-report (using scales from Chapter 2), behavioral, and neuroscientific (i.e., EEG) data from one online and one lab study.

The most important problem of using self-report measures is that people's answers to the questions can be influenced by social desirability (i.e., responding in a more socially accepted way). For my research, this means that participants might not report criticism situations that were especially painful to them (because they might be ashamed to reveal these) or that they downplay their affective reaction. I therefore created a new experimental paradigm in Chapter 3 that could capture participants' responses more indirectly by recording an EEG while participants received criticism in the experiment. To be able to extract ERPs from the EEG, the main measures of this method, an experimental paradigm needs to include many more trials than in a typical social psychological paradigm. That's because we need to ensure that we measure actual brain responses to our manipulations rather than noise (we need to increase the signal-to-noise ratio, e.g., Luck, 2005). Thus, I built on a repeated measure design from previous ERP-research (Schindler & Kissler, 2018) and adapted it for my purposes, combining it with methods and approaches from economics and social psychology. Firstly, I added a cover story to increase external validity. And secondly, I added a money contribution task that is similar to social dilemma situations (Komorita & Parks, 1994) where participants are asked to make either selfish (i.e., take money for themselves) or prosocial decisions (by contributing money to a common good). These two adaptations of the original task ensured that participants would find it believable to receive feedback on their monetary decisions by others (i.e., ingroup and outgroup members) and feedback on their morality (e.g., selfish monetary decisions) or their competence (e.g., smart monetary decisions).

In line with the findings of Chapter 2, I found that ingroup (vs. outgroup) feedback had more impact on participants' information processing of the feedback, their affective responses towards receiving the feedback, and their

reflection on receiving the feedback. That is, when receiving feedback in the actual moment, participants were more emotionally affected by ingroup compared to outgroup feedback. This was also reflected in the information processing stage: Participants initially paid more attention to feedback messages coming from ingroup (vs. outgroup) senders (as measured with EEG). For the second social context factor, the feedback dimension, I found more evidence for the ineffectiveness of feedback on people's morality. Participants paid less sustained attention to feedback messages concerning their morality (vs. competence). This was complemented by self-report findings which showed that participants felt worse after receiving negative feedback on their morality (vs. competence). Finally, participants also remembered fewer negative feedback messages when these related to their morality (vs. competence).

To conclude, Chapters 2 and 3 showed that ingroup (vs. outgroup) feedback and feedback on participants' competence (vs. morality) is more effective. However, framing criticism as being about someone's competence might not always be desirable. For example, in some cases a person giving the feedback might want to stress the ethical implications of immoral behavior or enforce a specific moral norm (e.g., no fraudulent behavior on the financial market). In the final chapter, I, therefore, aimed to test a second and more broadly to be applied strategy to increase the receipt of feedback from outgroup members and concerning one's morality: Communicating the intention to help when delivering feedback.

1.7.2 | Strategy 2: Communicating the intention to help

1.7.2.1 | Chapter 4: Intervention

Chapter 4 is introducing a new intervention, set in place before the information processing stage (see Figure 1.1): Communicating wanting to help when delivering feedback to make people more open to being criticized for their morality. This intervention builds both on social psychological theory and research findings (Brambilla et al., 2021; Cuddy et al., 2008; Hornsey et al., 2004; Hornsey & Imani, 2004) as well as on the findings from the first two empirical chapters of this dissertation. The past research and theory suggest that the morality of a person can give insights into whether a person's

intentions might be harmful or beneficial. In Chapter 2, participants indeed made more negative assumptions about critics (e.g., that their criticism was unconstructive) when they were criticized for their morality compared to for their competence at the workplace. This then resulted in the criticism to be less effective. Participants reported that they accepted this type of criticism less and that they less often changed their behavior based on the criticism. Making such negative assumptions, for example about the intentions of a critic is therefore a key psychological mechanism that prevents people from accepting moral criticism (as compared to competence criticism). In Chapter 4, I aimed to find an intervention that could prevent people from making negative assumptions about the person criticizing them. This intervention entailed having feedback senders communicate their helpful intentions when giving feedback. Or in other words, having them communicate that they provided the criticism to help the other improve (i.e., other-enhance intention). This intervention was also aimed to increase the receipt of outgroup criticism. Outgroup criticism is less effective than ingroup criticism because people are inclined to think the negative feedback is provided with more hostile intentions (e.g., Esposito et al., 2013; Hornsey & Imani, 2004; Hornsey et al., 2004; Thürmer et al., 2019). Thus, the strategy was aimed to intervene in both processes: Getting defensive towards *outgroup criticism* and when criticized on one's *morality*. To this end, I analyzed self-report and neuroscientific (i.e., EEG) data from one online and one lab study.

The experimental paradigm used in these studies is an improved version of the one I created for Chapter 3. That is, rather than passively viewing feedback messages, in this study participants were actively making decisions and received feedback after each decision. This method increased the self-relevance of the feedback (as can be seen, for example, in increased ERP amplitudes compared to the ERP-study of Chapter 3, see Figures 3.2, 3.3, 4.4, and 4.5).

In this chapter, I demonstrated that communicating intentions can indeed prevent participants from making more negative assumptions about outgroup (vs. ingroup) feedback senders and increase the perceived fairness of such feedback (reflection stage). Moreover, communicating helpful intentions

seemed to reduce threat associated with receiving negative feedback on people's morality. This was indicated by reduced attentional vigilance when participants cognitively processed negative feedback messages relating to their morality (as indicated by decreased P200-amplitudes, measured with EEG, information processing stage).

1.7.2.1 | Chapter 5: General Discussion

In Chapter 5, I integrate the findings from all three empirical chapters to form a new understanding of how social context factors influence the way we react to being criticized. I embed the findings in the social psychological and neuroscientific literature and show how this dissertation contributes to this literature. I discuss practical implications of the two strategies tested to make people more receptive to being criticized on their morality and by outgroup members as well as limitations and avenues for future research. Please note that Chapters 2-4 are prepared as separate journal articles. There might be some overlap in the theoretical background, but this also implies that these chapters can be read independently.

The background is a teal watercolor wash with a white torn paper effect at the bottom. The text is centered in white.

CHAPTER

2

Don't Tell Me About My Moral Failures but Motivate Me to Improve: Increasing Effectiveness of Outgroup Criticism by Criticizing One's Competence

This chapter is based on:

Rösler, I. K., Van Nunspeet, F., Ellemers, N. (2021). Don't tell me about my moral failures but motivate me to improve: Increasing effectiveness of outgroup criticism by criticizing one's competence. *European Journal of Social Psychology*, 51, 597-609. <https://doi.org/10.1002/ejsp.2764>

Author contributions:

IKR, FVN, and NE designed the studies. IKR performed the data analysis and wrote a first draft of the manuscript. FVN and NE provided substantial comments and suggestions for improvement.

Abstract

Criticizing people for their prior moral failures often causes them to react defensively, especially when this is done by an outgroup. In the current research, we tested whether people become more receptive to such outgroup criticism when it refers to (failures of) their competence, rather than their morality. We conducted two studies, using a 2: Critic's group-membership (receiving criticism from an ingroup vs. outgroup) x 2: Dimension (competence vs. morality as focal concern addressed with the criticism) mixed design. Findings showed that, regardless of source, participants made fewer negative attributions, were more motivated to improve, and more often indicated they changed their behavior after they had been criticized for their competence, instead of for their morality. Thus, criticizing past behavior for failing to show competence instead of morality might be a way to reduce defensive responses and to stimulate behavior change, even for outgroup critics.

In a Dutch interview in 2007, a journalist questioned the former CEO of ABN AMRO for receiving 26 million euros while the bank was taken over by a consortium bank. In the video of this interview, it can be observed that the former CEO had a very defensive response when being accused of acting disloyally towards his former colleagues at ABN. Moreover, he repeatedly referred to agreements that had been made about these types of payments in the financial world during a takeover. According to his group, receiving that much money did not correspond to violating any norms. However, the media and the general public saw this as a moral scandal. In the current article, we will present research findings indicating that a more successful strategy for the journalist (being an outsider) might have been to criticize the former CEO's competencies as a banker, rather than questioning his integrity. This might have stimulated a more open-minded discussion on how things can be changed in the future, rather than the defensive responses from the former CEO dominating the conversation.

2.1 | Criticism from an ingroup versus outgroup source

The degree of the banker's defensive reaction might have been influenced by the journalist not being a banker himself. Whereas criticism from our ingroups (e.g., fellow bankers; Tajfel & Turner, 1979) is seen as helpful for the self and the ingroup, criticism from outgroups (e.g., journalists) is often not accepted (Esposito et al., 2013; Hornsey et al., 2002; Hornsey & Imani, 2004). The defensive response towards an outgroup criticizing the ingroup (i.e., group-directed criticism) is often referred to as the intergroup sensitivity effect (e.g., Hornsey & Esposito, 2009). The underlying process here has proven to be related to making negative attributions about both the critic (i.e., the person criticizing) and the criticism (i.e., the message, Hornsey & Imani, 2004). When someone is criticized by an outgroup member, people often attribute negative motives to the critic and perceive the criticism as less constructive.

However, group-directed criticism is only loosely related to any kind of individual-level improvement attempts (e.g., behavioral change). An individual can hardly change the behavior of the whole group and is always presented with only the negative side of criticism (e.g., negative affect, threat) and not

the positive side (e.g., opportunity to grow). Thus, to be able to stimulate behavioral change, which is the ultimate goal when criticizing, we should also look at criticism at the individual level. Specifically, the differences between group- and individual-level criticism come from the different paths that a person can follow when responding to outgroup criticism. In group-directed criticism, the relevance of people's own past and future actions is not clear and they are not presented with an opportunity to make up for past failures (Van der Toorn et al., 2015). However, in individual-level criticism situations, this opportunity can make people motivated to improve. In the current research, we examined three underlying processes (i.e., emotional, cognitive, and motivational) that influence people's reactions to being criticized for their individual behavior by an outgroup. Thereby, we offer a novel perspective on why (and how strongly) the intergroup sensitivity effect occurs in individual-level situations. Additionally, we add to the literature by refining the social dimensions that are addressed with the focal concern of a critical message (i.e., morality vs. competence).

2.2 | Criticism of someone's morality versus competence

Morality and competence are two basic social dimensions that people use to form impressions of others and themselves. These two dimensions have been labeled in different ways, but there is a consensus about the content. One dimension refers to task ability (e.g., competence, agency) and the other to interpersonal intentions (e.g., morality, communion, warmth, Abele et al., 2016; Brambilla & Leach, 2014; Fiske et al., 2007). In previous research on intergroup criticism these two dimensions were not made explicit, nor directly compared. For example, Hornsey and colleagues (Hornsey et al., 2002; Hornsey & Imani, 2004) investigated group-directed outgroup criticism such as being called a "racist" or "living in an ivory tower". We argue that being called a racist can be considered criticism of one's morality, whereas living in an ivory tower can be considered criticism of one's competence, and that these types of focal concerns cause different reactions to the criticism. Being (seen as) a moral person is important to people, often even more important than being (perceived as) competent (Brambilla & Leach, 2014; Ellemers et al., 2008; Goodwin et al., 2014; Leach et al., 2007). Specifically, morality is more important than competence when it comes to the groups

we want to belong to (Leach et al., 2007), the norms we adhere to (Ellemers et al., 2008), and when we form impressions about the character of people (Brambilla & Leach, 2014; Goodwin et al., 2014). Since being moral is so relevant to people, they typically show highly emotional and threatened responses when light is shed on their moral failures (Täuber et al., 2015, 2018; Täuber & van Zomeren, 2013; Van der Lee et al., 2016). We argue that this can result in a self-defensive reaction, for instance by making negative attributions about the person voicing this criticism. A self-defensive reaction (discrediting the source of the criticism) can in turn cause people to ignore concerns raised (Gausel et al., 2012; Gausel & Leach, 2011; Giner-Sorolla, 2012). We extend prior work by examining whether this also prevents people from being motivated to improve their own behavior. When one's competence (vs. morality) is criticized, we predict this will result in less self-defensiveness and more motivation towards self-improvement. This prediction builds on prior research showing that people generally feel more capable of coping with their past shortcomings in the competence domain than with past moral failures (Van der Lee et al., 2016). Furthermore, whereas positive (vs. negative) judgments weigh in more when the competence of another person is judged, negative (vs. positive) judgments are more important when the morality of a person is judged (Martijn et al., 1992; Skowronski & Carlston, 1987). Thus, people might get more defensive and are inclined to hide immoral (vs. incompetent) behavior since the former has a higher potential to hurt their image. Even behavior and motivational processes are affected when potential implications of behavior are framed in terms of one's morality versus competence. In past research, participants were being evaluated on a task that could assess their social bias, but the implications of their performance were either framed in terms of their morality (i.e., their moral values concerning egalitarianism) or their competence (i.e., their ability to quickly learn new tasks, Van Nunspeet et al., 2014). Participants showed less negative social bias towards a target group in the morality (vs. competence) condition. This behavioral effect was complemented by an effect in the same direction for cognitive processes related to unconscious response monitoring (i.e., measured with EEG) which are indicators for the motivation to give the correct responses on a task. Participants were more motivated to respond

correctly when they were evaluated on their morality (vs. competence). In other words, in a situation that has a moral frame, people are more concerned about their moral image. This might make people more defensive towards criticism of their morality (vs. competence).

Our research differs from previous studies contrasting morality and competence criticism (Täuber et al., 2018; Täuber & van Zomeren, 2013) by having the criticism delivered by ingroup and outgroup members (i.e., as in the literature on the intergroup sensitivity effect) rather than by a neutral third party. Moreover, the current research investigated cognitive attributions people make about (the motives of) critics and how these impact on their motivation to improve their own behavior, rather than capturing the emotion of “moral outrage” directed against an outgroup that is portrayed as displaying superior morality (Täuber & van Zomeren, 2013).

2.3 | Current research

The reviewed findings suggest that criticism of one’s competence (vs. morality) might make people less defensive and more motivated to improve. In turn, this might make people more often accept criticism and change their behavior, even if the criticism comes from the outgroup.

To test this, we designed two studies in which we included the critic’s group-membership (i.e., ingroup vs. outgroup) and dimension of criticism (i.e., morality, competence). Balancing the trade-off between higher power and controlling for individual differences (e.g., a priori sensitivity to either ingroup vs. outgroup or morality vs. competence criticism) in within-participant designs and the feasibility of people having experienced a situation for all our four conditions, we varied the design between the two studies. In Study 1.1, the critic’s group-membership was included as a within-participant factor (dimension as a between-factor), since we aimed to test whether the intergroup sensitivity effect also applies in situations where individual behavior is criticized by an outgroup member. Here, we also aimed at getting a full understanding of the underlying (i.e., emotional, motivational, and cognitive) processes in individual-level criticism situations. In Study 1.2, our main aim was to test whether competence (vs. morality) criticism is a way

to decrease defensiveness and increase effectiveness of outgroup criticism (e.g., making people more often change their behavior based on criticism). Thus, we switched the dimension from a between-participants to a within-participant factor (and group-membership to a between-factor).

For both studies, we hypothesized that:

Cognitive attributions about the critic (e.g., motives, trustworthiness, credibility) would be more negative for outgroup critics compared to ingroup critics (Hornsey et al., 2002; Ilgen et al., 1979) and that criticism of the participant's competence, rather than of their morality, would decrease these negative attributions (Martijn et al., 1992; van der Lee et al., 2016; Van Nunspeet et al., 2014).

Participants would experience a more negative emotional response when being criticized by an outgroup member compared to an ingroup member (Hornsey et al., 2002), and that this emotional reaction would be smaller when criticism concerned their competence rather than their morality (Martijn et al., 1992; Van der Lee et al., 2016; Van Nunspeet et al., 2014).

Participants would be more motivated to improve when criticized by the ingroup (vs. outgroup, Hornsey et al., 2002) and criticism of participants' competence (vs. morality) would make people more motivated to improve (Martijn et al., 1992; Van der Lee et al., 2016; Van Nunspeet et al., 2014).

Criticism would therefore be more effective when given by the ingroup (vs. outgroup) and on participants' competence (vs. morality, see Figure 2.1).

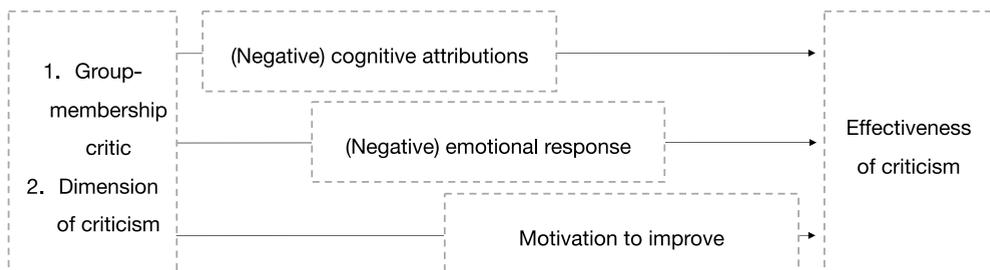


Figure 2.1. Conceptual model of hypothesized underlying mechanisms of the effect of group-membership and dimension on the effectiveness of criticism.

2.4 | Study 1.1

2.4.1 | Methods

2.4.1.1 | Participants and design

A total of 202 participants completed the questionnaire on the online research platform Prolific in exchange for €6. We based our sample size on past (lab) research in which participants recalled evaluations of ingroup members on moral and competence failures (69 participants, Van der Lee et al., 2016) and increased the sample size due to the extended design (i.e., we added a between-participants factor: Critic's group-membership), increased anonymity, and to be able to exclude participants who did not meet attention checks and other controls for serious participation, as usual for online studies. The study was approved by the local ethics committee.

The study had a 2: Critic's group-membership (in- vs. outgroup, within-participants factor) x 2: Dimension (morality vs. competence, between-participants factor) design. The order of the critic's group-membership (i.e., whether criticism from an ingroup or outgroup member was recalled first) was counterbalanced between participants. Participants answered both quantitative (i.e., closed questions about their experiences being criticized) and qualitative questions (e.g., open-ended questions, in which they could elaborate on their experiences). We focus on the quantitative part in the current article and report on some of the qualitative data for demonstration purposes.

Six participants were excluded from the data analyses because they were unable to recall a situation in which they had received criticism as specified in one of our conditions (i.e., ingroup or outgroup criticism of the participant's morality or competence). The remaining sample consisted of 196 participants, 103 in the competence condition, 93 in the morality condition ($M_{age} = 35.5$ years, $SD = 9.71$, 101 females, one not indicating sex). Sensitivity analyses in G*power ($\alpha = .05$, $1-\beta = .80$, $N = 196$) indicated that we were able to detect effect sizes as small as $\eta^2_p = .04$ (small effect) for within-factors using repeated measure analysis of variance (RM ANOVA).

Participants were employed in a broad spectrum of career sectors, holding

different types of industry roles, and most participants were from the UK (58.3%, see Supplementary Materials for more details).

2.4.1.2 | Procedure

Participants were asked to recall two situations in which they had received criticism at their workplace. Depending on the experimental condition, participants were requested to recall a situation where the criticism had addressed either their competence or morality (i.e., between-participants) and was provided either by an ingroup or an outgroup member (i.e., within-participants). We explained that moral criticism could concern, for example, behavior that was considered to be dishonest, insincere, or untrustworthy, and competence criticism could concern behavior that was considered to be incompetent, unintelligent, or unskilled. To emphasize the group membership of the person providing the criticism, we first explained that “individuals represent different groups, for instance, your work team versus other work teams, colleagues with similar skills or different skills, or individuals who share distinctive features (nationality, education, etc.), or differ from each other in these features” and asked participants to think of different groups at their workplace. We then instructed participants to think of a specific situation in which they had received criticism from someone at the workplace (excluding their boss/supervisor). Most often mentioned ingroups were colleagues (same position 42%, same work team 32%) and most mentioned outgroups were different work teams (32%) or other departments (27%, also see Supplementary Materials).

2.4.1.3 | Measures

To help participants recall the experience, we asked them about the context of the situations, their (negative) attributions about the critic and the criticism, the identification with the group-membership that defined the critic as an ingroup versus outgroup member, their reaction to the criticism (i.e., emotional response, perceived motivation to improve), and asked them to indicate whether they accepted, had the intention to change their behavior based on, and changed their behavior based on, the criticism. Except for guided open-ended questions about the context of the situation recalled, and when otherwise indicated, all items were presented on 7-point scales (1 = not at all,

7 = very much, reliability indices for all measures sufficient, Supplementary Materials Table S2.2).

2.4.1.1.1 | Checks

After describing each of the two situations, we asked participants to specify the critic's group-membership (ingroup vs. outgroup) and the primary dimension of the criticism (morality vs. competence). In case participants indicated that the incident they recalled did not meet the criteria, we asked them to think of another situation that did match the experimental condition they were in. Additionally, we asked them to indicate to what extent the criticism referred to their morality (1) or their competence (7, in the actual questionnaire this was labeled "performance") on a 7-point bipolar scale. We used three items to measure group identification (e.g., "I identify with other members of this group") and four items to assess collective self-esteem (e.g., "I feel good about this group") pertaining to the group that defined the critic as an ingroup versus outgroup member. These items were adapted from the subscales developed by Ellemers et al. (1999). The items could be answered on an 8-point scale, ranging from 1 (not at all) to 8 (very much).

We measured fear of exclusion from the ingroup with seven items (adapted from Cavazza et al., 2014, "I was afraid that my group would exclude me if they would hear about the criticism") to be able to exclude this as an alternative explanation for our anticipated effects. For other potential confounds, participants were presented with guided questions in which they had to choose between several preset answering options to specify the nature of the situation they had recalled. Participants were asked about their interpersonal relationship with the critic (i.e., "I liked/did not like this person"), the work related status of the critic compared to the participant (i.e., "Lower, Equal, Higher"), the delivery form of the criticism (e.g., "Quick, clear, and friendly"), the circumstances under which the criticism was provided (e.g., "The feedback was offered spontaneously by the other person"), and the time of the event (e.g., "1–3 months ago").

2.4.1.1.2 | Negative attributions

We measured how participants evaluated the person criticizing them (i.e.,

the critic) and the critical message itself (i.e., the criticism). We measured perceived intention of the critic (two items, e.g., “I think the person who criticized me did this in my best interest”, adapted from Hornsey et al., 2004), trust in the critic (two items, e.g., “I now trust the person”), and credibility of the critic (two items, e.g., “I think the person is credible”). To investigate how the message was evaluated by participants, we measured perceived constructiveness of the criticism (two items, adapted from Hornsey et al., 2004, e.g., “The criticism was intended to be constructive”), the credibility of the message (one item, “The criticism was credible”), and perceived threat (two items, e.g., “I perceived the criticism as threatening”). Maximum likelihood factor analysis (MFLA) using varimax rotation revealed that all items loaded on one factor (for excluded items and factor analyses see Supplementary Materials), rather than several factors, explaining at least 68.5% (lowest score of the two repeated measures) of the variance with factor loadings ≥ 0.69 . We thus combined the items into a single scale, negative attributions about critic (nine items), and recoded positively framed items so higher scores on this scale indicate more negative attributions.

2.4.1.1.3 | Emotional response

We measured participants’ emotional responses to being criticized with three items from the Positive and Negative Affect Schedule PANAS, Watson et al., 1988, e.g. “The criticism upset me”), 11 items of more specific emotions related to performance feedback (i.e., achievement success, avoidance failure, adapted from Brockner & Higgins, 2001, e.g., “The criticism made me feel tense”) and one item related to perceived control (i.e., “I felt the criticism had to do with something outside of my control, something I am unable to influence”). MFLA using varimax rotation yielded a three-factor solution for participants’ emotional responses, which together explained 63.9% of the variance in the individual items and was related more to general valence rather than specific emotional responses to the criticism (e.g., avoidance of failure). The first factor comprised negative emotion items (at least 42.4% explained variance with factor loadings ≥ 0.53), the second factor comprised positive emotion items (at least 16% explained variance, factor loadings ≥ 0.61), and the third factor comprised the two moral emotions guilt and shame (at least 7.8% explained variance, factor loadings ≥ 0.70). We thus

used three scales to measure the emotional responses to the criticism, negative emotions (eight items), positive emotions (four items), and moral emotions (two items) for our analyses.

2.4.1.1.4 | Motivation to improve

The motivation to improve based on the criticism was measured with four items (e.g., “The criticism made me feel energized about improving myself”, “I felt like I could do something with the criticism”).

2.4.1.1.5 | Effectiveness of criticism

To investigate whether participants accepted and used the criticism to change their behavior, we measured acceptance of criticism (three items, e.g., “I accepted the criticism”), behavioral intention to adapt their behavior based on the criticism (three items, e.g., “I planned to use the criticism to improve”), and actual behavioral change based on the criticism (three items, e.g., “I changed my behavior based on what was criticized”). Correlational analysis revealed that all three measures were highly correlated (see Supplementary Materials Table S2.3). Furthermore, MFLA with varimax rotation revealed that all three measures loaded on one rather than three distinct factors. The one-factor solution explained at least 81.5% of the variance with factor loadings ≥ 0.74 . We thus combined the items to one measure, effectiveness of criticism (nine items).

We tested whether effectiveness was distinctive of motivation to improve by specifying two factors in another MFLA analysis with varimax rotation. Results confirmed this. In the two-factor solution, effectiveness explained at least 73.5% of the variance with factor loadings ≥ 0.50 . Motivation to improve explained at least 6% of the variance with factor loadings ≥ 0.67 .¹ Additionally, we specified two factors to test whether effectiveness was distinctive of negative attributions, which was also the case. Effectiveness explained at least 65.4% of the variance, with factor loadings ≥ 0.62 and negative attributions explained at least 11.6% of the variance with factor

¹ There was overlap between the acceptance part (three items) of effectiveness and motivation to improve (see Supplementary Materials Table S2.7 [Study 1] and Table S2.14 [Study 2]). Taking out these items improves the distinction of these two factors but does not change mediation (see Supplementary Materials Figure S2.1 [Study 1] and Figure S2.4 [Study 2]).

loadings ≥ 0.64 .

2.4.2 | Results

2.4.2.1 | Checks

We examined identification with the critic's group-membership with paired sample t-tests. As intended, participants reported significantly higher levels of identification with the ingroup ($M = 5.61$, $SD = 1.62$) than with the outgroup ($M = 3.57$, $SD = 1.74$), $t(195) = 13.57$, $p < .001$. Analysis of collective self-esteem yielded very similar results (see Supplementary Materials).

Additionally, we compared the means from both the ingroup and outgroup condition on our bipolar dimension manipulation check (1 = morality, 7 = competence) to the midpoint of the scale (4) with one-sample t tests. Participants in the morality condition reported to have experienced moral criticism ($M = 3.14$, $SD = 1.34$), $t(92) = -6.17$, $p < .001$. Participants in the competence condition reported to have experienced competence criticism ($M = 5.44$, $SD = 1.33$), $t(102) = 10.94$, $p < .001$. Additionally, both means were significantly different from each other, $t(194) = 12.00$, $p < .001$. Thus, our manipulations were successful.

We did not find evidence for alternative explanations for our reported effects and criticism situations were comparable (see Supplementary Materials). Fear of exclusion from the ingroup, interpersonal liking between the participant and the critic, delivery of the criticism by the critic (e.g., unfriendly), the status of the critic, time of delivery, and context of criticism situations did not vary systematically with our manipulation. Furthermore, criticism concerned participants' (specific) behavior rather than their (abstract) identity.

2.4.2.2 | Content coding criticism situations

We content-coded criticism situations participants ($N = 392$) reported on from one of our open-ended questions (i.e., "First, we would like to ask you to think about the actual criticism. What was the essence of the criticism you received?"). We first developed an initial coding scheme with blinded conditions. Then, we formed super-categories including initial codes that were related to each other, separately for the morality and the competence

condition.

The coding revealed that criticism situations recalled by participants referred to very common everyday shortcomings people may display in the workplace. The behaviors mentioned most often in the competence condition were related to lack of competence displayed in the “work outcome” (69%), the “work attitude” (11%), and “care in completing work-related tasks” (7%). Examples of work outcomes seen as indicating lack of competence are, for example, “not doing a task good enough” or “working too slow”. Work attitudes criticized for lack of competence include behaviors such as “behaving unprofessionally” or “taking too many breaks”. Lack of care in completing work-related tasks include behaviors such as “making mistakes” or “forgetting a task”. Likewise, the behaviors that were mentioned most often in the morality condition referred to behavior that did not seem to meet moral standards pertaining to “cooperative behavior” (23%), “honesty” (21%), and “work ethic” (18%). Examples of behaviors indicating lack of cooperation are, for example, “letting other people do more work” or “not being part of the team”. Examples of failures to meet standards for honesty include behaviors criticized for “not being truthful about having performed a task” or “cheating”. Behavior that was criticized as indicating lack of appropriate work ethics includes “accepting gifts from customers” or “not following company guidelines”.

2.4.2.3 | Negative attributions

We predicted that participants would make more negative attributions about outgroup (vs. ingroup) critics. Furthermore, we hypothesized that receiving criticism of one’s competence (vs. morality) would decrease negative attributions. To test this, we used a 2: Group-membership Critic (in- vs. outgroup, within-participants factor) x 2: Dimension (morality vs. competence, between-participants factor) mixed RM ANOVA.

As predicted, there was a main effect of the critic’s group-membership on negative attributions. Participants made more negative attributions about outgroup ($M = 4.82$, $SD = 1.62$) than about ingroup critics ($M = 3.90$, $SD = 1.78$), $F(1, 194) = 28.60$, $p < .001$, $\eta^2_p = .13$, 95% CI = [.58, 1.25]. However,

there were no other significant effects, $F_s < 1$.

2.4.2.4 | Motivation to improve

We predicted that participants would report more motivation to improve when the ingroup (vs. outgroup) criticized them and that receiving criticism of one's competence (vs. morality) would increase motivation to improve. As predicted, a mixed RM ANOVA (Critic's group-membership [within], Dimension [between]) revealed that participants reported more motivation to improve when criticism came from the ingroup ($M = 3.46$, $SD = 1.95$) than from the outgroup ($M = 2.83$, $SD = 1.72$), $F(1, 194) = 14.09$, $p < .001$, $\eta^2_p = .07$, 95% CI = [.30, .96]. There was also a main effect of dimension: Participants reported more motivation to improve when they were criticized for their competence ($M = 3.36$, $SD = 1.44$) as compared to their morality ($M = 2.91$, $SD = 1.36$), $F(1, 194) = 5.12$, $p = .025$, $\eta^2_p = .03$, 95% CI = [.06, .85]. There was no interaction effect between the critic's group-membership and dimension, $F < 1$.

2.4.2.5 | Emotional response

For the participant's emotional responses, we predicted an effect of the critic's group-membership and dimension in the same direction (i.e., a greater negative emotional response to outgroup [vs. ingroup] criticism, lower in case the criticism was about one's competence [vs. morality]).

A mixed RM MANOVA (Critic's group-membership [within], Dimension [between]) with negative emotions, positive emotions, and moral emotions as dependent factors and using an alpha level of .016 (Bonferroni correction, .05/3) to correct for multiple comparisons revealed an effect of the critic's group-membership on the emotional response at the multivariate level, Pillai's Trace = .07, $F(3, 192) = 4.80$, $p = .003$, $\eta^2_p = .07$. At the univariate level, participants reported stronger negative emotions when the criticism came from the outgroup ($M = 4.27$, $SD = 1.52$) as compared to the ingroup ($M = 3.82$, $SD = 1.60$), $F(1, 194) = 13.79$, $p < .001$, $\eta^2_p = .07$, 95% CI = [.21, .70]. Similarly, they indicated stronger positive emotions when the criticism came from the ingroup ($M = 2.37$, $SD = 1.39$) as compared to the outgroup ($M = 2.10$, $SD = 1.30$), $F(1, 194) = 6.59$, $p = .011$, $\eta^2_p = .03$, 95% CI = [.06,

.48]. There were no other significant effects ($F_s \leq 2.40$).

2.4.2.6 | Effectiveness of criticism

We tested whether the critic's group-membership and dimension influenced the effectiveness of the criticism received. As predicted, the criticism was more effective when it came from the ingroup ($M = 3.93$, $SD = 2.07$) than from the outgroup ($M = 3.24$, $SD = 2.03$), $F(1, 194) = 11.93$, $p = .001$, $\eta_p^2 = .06$, 95% CI = [.29, 1.07].² However, there was no significant effect of the dimension on effectiveness (competence $M = 3.70$, $SD = 1.43$, morality $M = 3.47$, $SD = 1.62$), $F(1, 194) = 1.15$, $p = .286$, and no interaction effect, $F < 1$.³

2.4.2.7 | Mediation

Lastly, we examined mediation effects. We had anticipated that the cognitive (i.e., negative attributions about the critic), the emotional (i.e., affective responses), and the motivational pathway (i.e., the motivation to improve) would mediate the effects of the critic's group-membership and dimension on effectiveness of the criticism. We found direct effects of the critic's group-membership (within-participant factor) for all our DV's (i.e., negative attributions, emotional response, motivation to improve, effectiveness). However, we did not find an effect of dimension on the main dependent variable, effectiveness of the criticism. Therefore, we only tested for mediation of the critic's group-membership effect (for indirect effects for dimension see Supplementary Materials). We used the SPSS Macro MEMORE (Montoya & Hayes, 2017) for two-condition within-participants mediation. The relationship between the critic's group-membership and effectiveness of criticism was fully mediated by negative attributions about the critic and motivation to improve (parallel mediation, see Figure 2.2), with the overall model explaining 71.7% of the variance. The total unstandardized bootstrapped (5000 samples) indirect effect was significant, .78, 95% CI [.45, 1.11]. When adding the

² Significant effects in the same direction were found when using the three initial scales (i.e., acceptance, intention to change, actual behavioral change) independently (acceptance: $\eta_p^2 = .10$, intention to change: $\eta_p^2 = .04$, actual behavioral change: $\eta_p^2 = .04$).

³ Interestingly, whereas negative emotions correlated negatively with effectiveness of the criticism, moral emotions (even though also negative in valence) correlated positively with effectiveness. This suggests that reporting more moral emotions was associated with more effectiveness of the criticism (see Supplementary Materials Table S2.3 [Study 1] and Table S2.10 [Study 2]).

emotional responses (i.e., positive, negative, moral) to this mediation model, they did not significantly mediate the relationship between the critic's group-membership and effectiveness of the criticism.

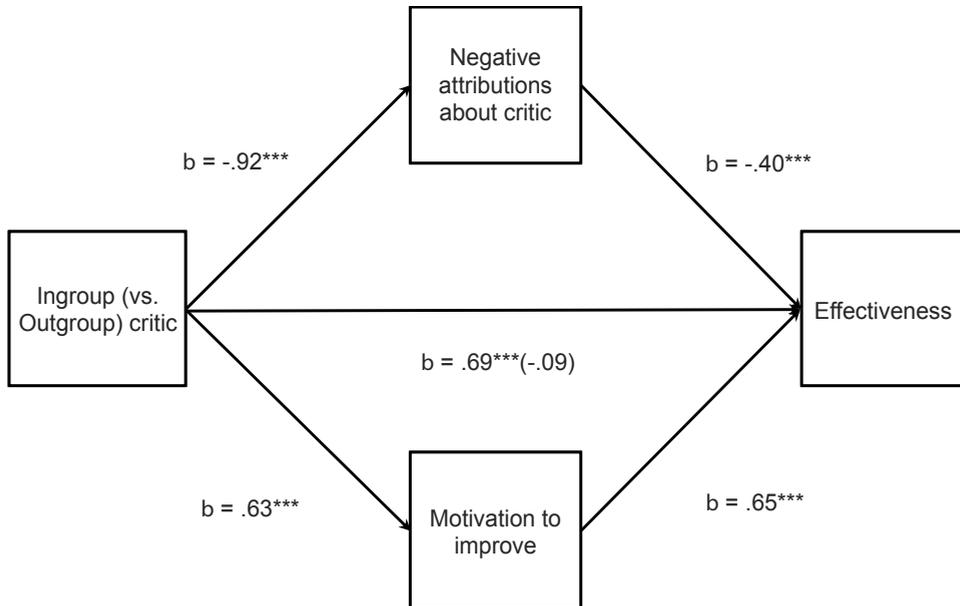


Figure 2.2. Mediation model Study 1.1 with unstandardized regression coefficients (b) depicting the relationship between the within-participants factor 'critic's group-membership' and 'effectiveness of the criticism', as mediated by 'negative attributions about critic' and 'motivation to improve'. The coefficient for the direct effect of 'group-membership' on 'effectiveness of criticism' is in parentheses. *Note:* $***p < .001$

2.5 | Study 1.2

In Study 1.1, we demonstrated that the intergroup sensitivity effect also applies when individual (vs. group) behavior is criticized. The second goal of the current research was to test whether competence (vs. morality) criticism can increase effectiveness of (outgroup) criticism. We therefore switched the within versus between nature of the two main factors in our experimental design. In Study 1.2 the dimension of the criticism is no longer a between-

participants but a within-participant factor (and the critic's group-membership now a between-participants factor). This way, we can control for individual differences and increase statistical power while examining the effects of the dimension of criticism.

A possible alternative explanation for our effect of group-membership on effectiveness of criticism is that outgroup criticism is perceived as more severe. People usually interact more with ingroup members at the workplace and they might receive more specific and concrete types of criticism from them, which therefore may seem less severe ("I saw you made a mistake on this task") than if it came from outgroup members ("I feel I can't trust you"). To be able to investigate and exclude this alternative explanation, we added a measure to Study 1.2 to check for perceived severity of the criticism.

2.5.1 | Methods

2.5.1.1 | Participants, design, and procedure

A total of 402 participants completed this study on Prolific in exchange for €6. Participants ($M_{\text{age}} = 34.5$ years, $SD = 10.94$, 266 females, one not indicating sex⁴) were randomly assigned to two of four conditions of the 2: Dimension (morality vs. competence, within-participants factor) x 2: Critic's group-membership (in- vs. outgroup member, between-participants factor) design. Whether participants were first asked to recall criticism of their morality or competence (or vice versa) was counterbalanced between participants. We excluded 17 participants from data analyses because they could not think of a situation that we asked for. The remaining sample size consisted of 385 participants, 197 participants in the ingroup condition, 188 in the outgroup condition. Sensitivity-analyses in G*power ($\alpha = .05$, $1-\beta = .80$, $N = 385$) indicated that we can detect effect sizes as small as $\eta_p^2 = .02$ (small effect) for within-factors using repeated measure analysis of variance (RM ANOVA). The study was approved by the local ethics committee.

The sample was similarly diverse as Study 1.1 (Supplementary Materials). As in Study 1.1, experiences with ingroup members mostly referred to

⁴ Four participants failed to insert their identification number and are not included here since we could not identify their demographic data; however, we included them in the main analyses.

people from the same work team (60%) or from the same department (33%). Outgroup experiences mostly indicated people from different departments (47%) and other work teams (35%).

We used the same instructions and manipulations as in Study 1.1.

2.5.1.2 | Measures and dependent variables

We added items to our measures of moral emotions (adapted from Gausel & Brown, 2012, four items, e.g., “I feel guilty when I think about being criticized this way”) to be able to construct a more reliable measure of moral emotions. We added one item to be able to investigate perceived severity of the situation (“How severe/extreme was the situation for you?”), both presented on 7-point scales (1 = not at all, 7 = very much). We excluded measures that were found to be unrelated to our effects (e.g., fear of exclusion) and items that were excluded from our measures (see Supplementary Materials) in Study 1.2. Other than that, we used the same scales as in Study 1.1 (reliability indices sufficient [Table S2.9] and factor analyses Supplementary Materials).

2.5.2 | Results

2.5.2.1 | Checks

We examined identification with ingroup and outgroup critics with independent sample t-tests. As intended, both in the morality and competence criticism condition respectively, participants reported higher levels of identification with the ingroup ($M = 4.88$, $SD = 1.57$; $M = 4.91$, $SD = 1.47$) than with the outgroup ($M = 3.72$, $SD = 1.78$; $M = 3.79$, $SD = 1.57$), $t(383) = 6.79$, $p < .001$ and $t(383) = 7.23$, $p < .001$. Collective self-esteem yielded similar results (see Supplementary Materials).

We compared the means from our bipolar dimension manipulation check (morality condition: 1 = morality, 7 = competence) to the mid-point of the scale (4) with one-sample t tests. Participants in the morality condition reported that the criticism was about morality ($M = 2.68$, $SD = 1.81$), $t(384) = -14.30$, $p < .001$ and in the competence condition that the criticism was about competence ($M = 5.68$, $SD = 1.68$), $t(384) = 19.61$, $p < .001$. Thus, our manipulations were successful.

As in Study 1.1, we checked whether criticism situations were comparable, and whether criticism concerned participants' behavior rather than their identity, which was found to be the case (see Supplementary Materials).

To check for an additional alternative explanation that we could not exclude in Study 1.1, we also asked participants to rate the perceived severity of the criticism received. This allowed us to examine whether the perceived severity of the situation varied between our conditions. A mixed RM ANOVA (Dimension [within], Critic's group-membership [between]) revealed no effects of how severely participants rated the criticism situation either for the group-membership of the critic, or for the dimension addressed with the criticism, or for their interaction, $F_s \leq 1.92$, $p_s \geq .167$. Thus, our effects cannot be explained by a difference in perceived severity between experimental conditions.

2.5.2.2 | Content coding criticism situations

We content-coded criticism situations ($N = 770$). The coding revealed that, as in Study 1.1, most participants in the competence condition reflected on behaviors that were criticized for failing to display competence in "work outcomes" (69%), "work attitudes" (9%), and "care in work-related tasks" (9%). In the morality condition, behaviors participants recalled were criticized for failing to meet standards for "honesty" (32%), "cooperation" (18%), and "work ethics" (11%).

2.5.2.3 | Negative attributions

We predicted that outgroup (vs. ingroup) and competence (vs. morality) criticism would lead to fewer negative attributions about the critic. Indeed, a mixed RM ANOVA (Dimension [within], Critic's group-membership [between]) revealed that participants made more negative attributions when receiving criticism of their morality ($M = 4.73$, $SD = 1.60$) as compared to their competence ($M = 4.34$, $SD = 1.71$), $F(1, 383) = 13.31$, $p < .001$, $\eta_p^2 = .03$, 95% CI = [.18, .60]. The expected main effect of the critic's group-membership on negative attributions was also significant, $F(1, 383) = 14.50$, $p < .001$, $\eta_p^2 = .04$, 95% CI = [.24, .74]. Participants made more negative attributions about the outgroup ($M = 4.79$, $SD = 1.13$) compared to the ingroup ($M = 4.30$, SD

= 1.37). There was no interaction effect, $F < 1$.

2.5.2.4 | Motivation to improve

In line with our expectation that competence (vs. morality) criticism would increase participants' motivation to improve, a mixed RM ANOVA (Dimension [within], Critic's group-membership [between]) yielded a main effect of dimension on motivation to improve. Participants were more motivated to improve in the competence ($M = 3.15$, $SD = 1.93$) than in the morality condition ($M = 2.60$, $SD = 1.69$), $F(1, 383) = 25.10$, $p < .001$, $\eta^2_p = .06$, 95% CI = [.34, .78]. However, motivation to improve did not significantly differ between the ingroup ($M = 2.99$, $SD = 1.52$) and outgroup ($M = 2.74$, $SD = 1.36$), $F(1, 383) = 2.91$, $p = .089$. There was no interaction effect, $F < 1$.

2.5.2.5 | Emotional response

We hypothesized that outgroup (vs. ingroup) criticism and morality (vs. competence) criticism would elicit more negative emotions. A mixed RM ANOVA (Dimension [within], Critic's group-membership [between]), negative emotions, positive emotions, and moral emotions as dependent factors, and an alpha level of .016 (Bonferroni correction, $.05/3$), revealed an effect of dimension on the emotional response at the multivariate level, Pillai's Trace = .03, $F(3, 381) = 4.40$, $p = .005$, $\eta^2_p = .03$. At the univariate level, participants reported stronger positive emotions in the competence ($M = 1.93$, $SD = 1.02$) than in the morality condition ($M = 1.76$, $SD = 1.01$), $F(1, 383) = 7.62$, $p = .006$, $\eta^2_p = .02$, 95% CI = [.05, .28]. Furthermore, participants indicated stronger moral emotions in the morality ($M = 3.09$, $SD = 1.75$) as compared to the competence condition ($M = 2.88$, $SD = 1.63$), $F(1, 383) = 6.05$, $p = .014$, $\eta^2_p = .02$, 95% CI = [.04, .28]. There were no other significant effects ($F_s \leq 1.38$).

2.5.2.6 | Effectiveness of criticism

We tested our prediction whether competence criticism can increase the effectiveness of outgroup criticism with a mixed RM ANOVA (Dimension [within], Critic's group-membership [between]). Criticism was more effective in the competence ($M = 3.42$, $SD = 1.97$) compared to the morality condition ($M = 2.86$, $SD = 1.85$), $F(1, 383) = 20.97$, $p < .001$, $\eta^2_p = .05$, 95% CI = [.32,

.80].⁵ Criticism in the ingroup condition ($M = 3.28$, $SD = 1.54$) was somewhat more effective compared to the outgroup condition ($M = 3.00$, $SD = 1.44$), $F(1, 383) = 3.48$, $p = .063$, $\eta^2_p = .01$, 95% CI = $[-.02, .58]$. There was no interaction effect, $F < 1$.

2.5.2.7 | Mediation

We expected that negative attributions and motivation to improve would mediate the effects of the critic's group-membership and dimension on effectiveness of the criticism. We found only marginally significant effects of group-membership on effectiveness and thus only investigated mediation for dimension of criticism (but see Supplementary Materials for these analyses). As predicted, the relationship between dimension and effectiveness of criticism was fully mediated by the negative attributions about the critic and the motivation to improve. Similar to our results in Study 1.1, where emotional responses did not mediate the relationship between the critic's group-membership and effectiveness, emotional responses did not mediate the relationship between dimension and effectiveness. However, moral emotions (i.e., shame, guilt) did mediate the relationship between dimension and effectiveness (parallel mediation, see Figure 2.3)—with competence (vs. morality) criticism being negatively associated with moral emotions, but moral emotions being positively associated with effectiveness. The full model explained 74.8% of the total variance, with the total unstandardized bootstrapped (5000 samples) indirect effect being significant, .46, 95% CI $[.25, .67]$.

⁵ Significant effects in the same direction were found when using the three measures independently (acceptance: $\eta^2_p = .06$; intention to change: $\eta^2_p = .05$; actual behavioral change: $\eta^2_p = .03$).

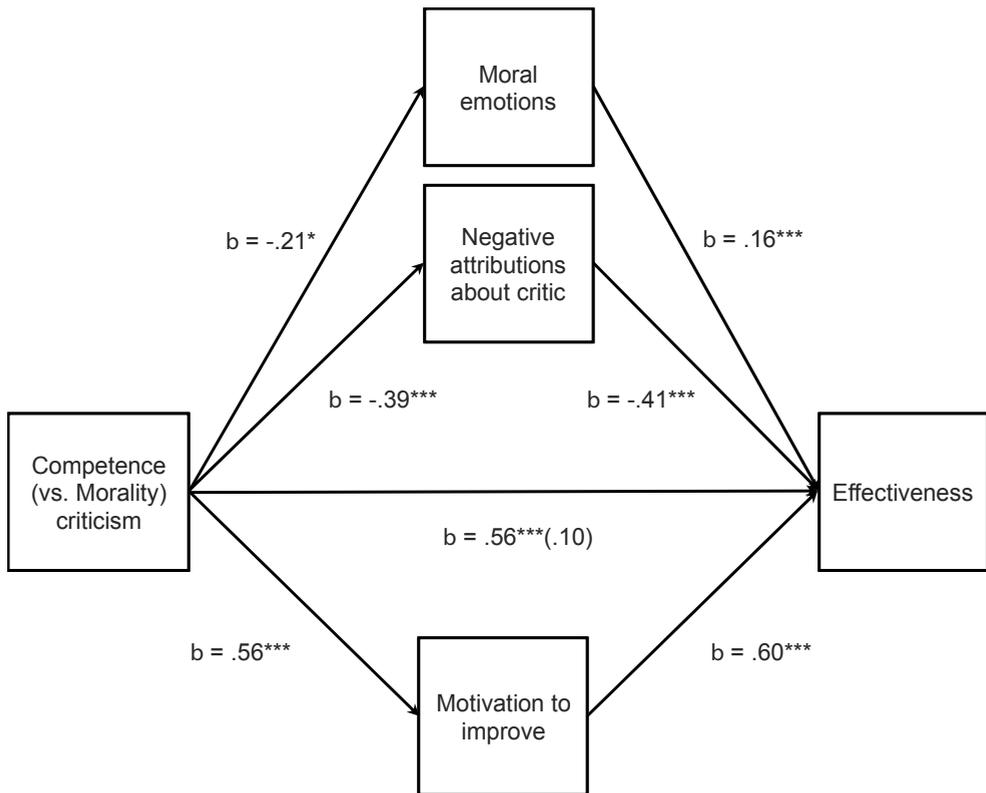


Figure 2.3. Mediation model Study 1.2 with unstandardized regression coefficients (b) depicting the relationship between the within-participants factor ‘dimension’ and ‘effectiveness of the criticism’, as mediated by ‘moral emotions’, ‘negative attributions about critic’, and the ‘motivation to improve’. The coefficient for the direct effect of ‘dimension’ on ‘effectiveness of criticism’ is in parentheses.

Note: $*p < .05$, $***p < .001$

2.5.3 | Discussion

Past research on intergroup criticism has focused mostly on group-directed criticism and the negative motives ingroup members attribute to outgroups criticizing the ingroup (intergroup sensitivity effect, Hornsey et al., 2002; Hornsey & Imani, 2004). With the present research, we aimed to extend previous findings by demonstrating that the intergroup sensitivity effect also applies when individual (vs. group) behavior is criticized by an outgroup

(Study 1.1) and that the focal concern addressed with the criticism (i.e., competence instead of morality) can increase effectiveness of outgroup criticism (Study 1.2). Additionally, we examined which underlying processes (i.e., emotional, cognitive, and motivational) influence people's reactions to being criticized from either an ingroup or outgroup member and on either their morality or competence.

In Study 1.1, we investigated whether a critic's group-membership and the social dimension addressed with a criticism influence its effectiveness. In line with our predictions, participants made more negative attributions about outgroup critics compared to ingroup critics. Additionally, they reported feeling stronger negative and weaker positive emotions and were less motivated to improve when criticism came from the outgroup. Most importantly, ingroup criticism was more effective: Participants accepted, were more likely to change, and actually changed their behavior more often when being criticized by an ingroup (vs. outgroup) member. Additionally, we found a main effect of dimension on motivation to improve in the hypothesized direction. That is, participants reported more motivation to improve when receiving criticism of their competence (vs. morality). This suggests that, even though criticism from an outgroup member is less effective overall, receiving criticism of one's competence (vs. morality) makes a criticism message more effective. There was no mediating role of emotional response (e.g., negative emotions) between the critic's group-membership and the effectiveness of the criticism. However, negative attributions and the motivation to improve fully mediated the relationship between the critic's group-membership and effectiveness of the criticism. Thus, the underlying process of why outgroup criticism is perceived as less effective does not seem to be a more intense negative emotional reaction, but a more negative evaluation of the critic (i.e., defensiveness). This finding reveals interesting implications. Whereas an intense negative emotional response of the person criticized might be hard to prevent, it might be easier to stimulate a cognitive re-evaluation of the situation. For example, the critic could communicate his/ her good intentions or emphasize the opportunity to grow from the criticism. Decreasing defensiveness towards outgroups criticizing is especially needed considering recent research showing that defensiveness can translate into behavioral

reactions, for example defending the ingroup or showing hostile actions towards the outgroup (Thürmer et al., 2019; Thürmer & McCrea, 2018).

In Study 1.2 we switched dimension from a between- to a within-participants factor to control for individual differences. We were able to show strong support for our prediction that criticism of participants' competence (vs. morality) reduces participants' negative attributions about the critic, increases motivation to improve, and is overall more effective (e.g., participants reported having changed their behavior more often). Furthermore, participants reported stronger moral emotions in the morality compared to the competence condition, but not stronger negative emotions. Thus, even though this was not statistically tested here, we can speculate that both types of criticism elicit a similar negative emotional response. Competence criticism, on the other hand, made participants feel stronger positive emotions (e.g., feeling confident or happy) and more motivated to improve (replicating Study 1.1). Our hypothesized mediation effects were also confirmed. The effect of dimension on effectiveness of criticism was mediated by negative attributions about the critic and motivation to improve (not by negative or positive emotions). Competence criticism decreased negative attributions about the critic and increased motivation to improve, which, in turn, made the criticism more effective (i.e., more reported behavioral change). The combined results of both studies suggest that competence criticism is more effective than moral criticism, not because it lessens negative affect, but because it changes how people interpret the situation. This finding is promising and offers a relatively easy alternative approach for outgroup critics who aim at evoking behavioral change in the people they are criticizing. To illustrate, a critic could choose to focus the focal concern of a criticism message on the recipient's moral character (e.g., not being loyal to other coworkers by slacking at work) or on situational performance (e.g., being lazy). According to our findings, the latter might evoke more behavioral change by making the recipient less defensive and more motivated to improve.

Please note that moral emotions also mediated the relationship between dimension and effectiveness. People reported stronger moral emotions in the morality (vs. competence) condition (in line with Van der Lee et al.,

2016) and stronger moral emotions led to more effectiveness. Thus, in contrast to more motivation to improve and fewer negative attributions in the competence condition, which led to more effectiveness, fewer moral emotions in the competence condition inhibited part of this effectiveness. This might indicate that, in the morality condition, participants felt pressured to change their behavior by their guilt or shame (e.g., stress response). However, more importantly, in the competence condition, driving the overall effect of dimension on effectiveness, participants made fewer negative attributions about the critic and might have interpreted the criticism more as an opportunity to grow.

We replicated the effect of the critic's group-membership on negative attributions about the critic that we found in Study 1.1, but not for motivation to improve and effectiveness (only marginally significant). However, we had anticipated that the power for the critic's group-membership might not be sufficient to detect the effects of the between-participants factor in this type of research (i.e., recalling autobiographical events).

Since we investigated autobiographical situations there are some limitations to our research. Firstly, we only have self-report measures of effectiveness of criticism. A better approach to this would be to investigate actual behavioral change based on criticism, for example by experimentally manipulating the dimensions. Furthermore, we did not ask participants to reflect on evaluations of moral and competence failures but asked them to reflect on criticism which could have been given unjustly. However, we did this intentionally since people might engage in coping mechanisms (e.g., avoidance, covering up mistakes), restricting the number of situations people can think of as well as narrowing the range of reported criticism situations (e.g., only very small or insignificant failures of morality or competence). Secondly, research has shown that people are often not able to predict how they actually emotionally respond in a situation (Wilson & Gilbert, 2003) or to correctly report on past events (Shiffman et al., 1997). However, even in the absence of these emotional responses, we find consistent effects relating to the cognitive evaluation (i.e., making negative attributions about a source) and motivational tendencies (having seen the ability to improve based on the

criticism). In theory, these cognitive responses might have been triggered by immediate emotional responses, and we cannot exclude the possibility that this was the case. Importantly, this does not distract from the fact that the way people interpreted the comments they received—depending on the group-membership of the source and the focal dimension—predicted the effectiveness of the criticism. Future research might expand on this finding by also capturing immediate emotional responses in order to explore whether and how these might contribute to the cognitive effects we observed or have additional independent effects.

Thirdly, we specifically asked participants to reflect on criticism at their workplace—where competence, as well as morality, are relevant dimensions for behavioral judgment. One could argue that this context might highlight the relevance of competence over the morality dimension. However, there is a large body of research from social cognition suggesting that morality generally dominates in person perception and evaluation (Brambilla & Leach, 2014; Goodwin et al., 2014), as well as research on organizations suggesting that morality is also highly relevant (often even more than competence and sociability) at the workplace and in task performance contexts (Ellemers et al., 2011; Leach et al., 2007; van Prooijen & Ellemers, 2015). Nevertheless, future research might examine whether these effects generalize to other contexts. Perhaps people are generally less concerned, and hence also less responsive, to competence-related criticism in a context where task performance is less relevant—for instance when in a group of friends or a family setting.

A potential confound for our effects of group-membership on our dependent variables could be related to the delivery of the criticism as we found that ingroup (vs. outgroup) criticism was more often (perceived to be) delivered in a “friendly and clear” manner. However, when submitting delivery of criticism as a covariate to our main analyses, the effect of group-membership remained (see Supplementary Materials).

Another potential confound for our effects of dimension (i.e., morality vs. competence) on effectiveness of criticism could be the concreteness/abstractness of the criticism that was delivered (Moscatelli et al., 2019).

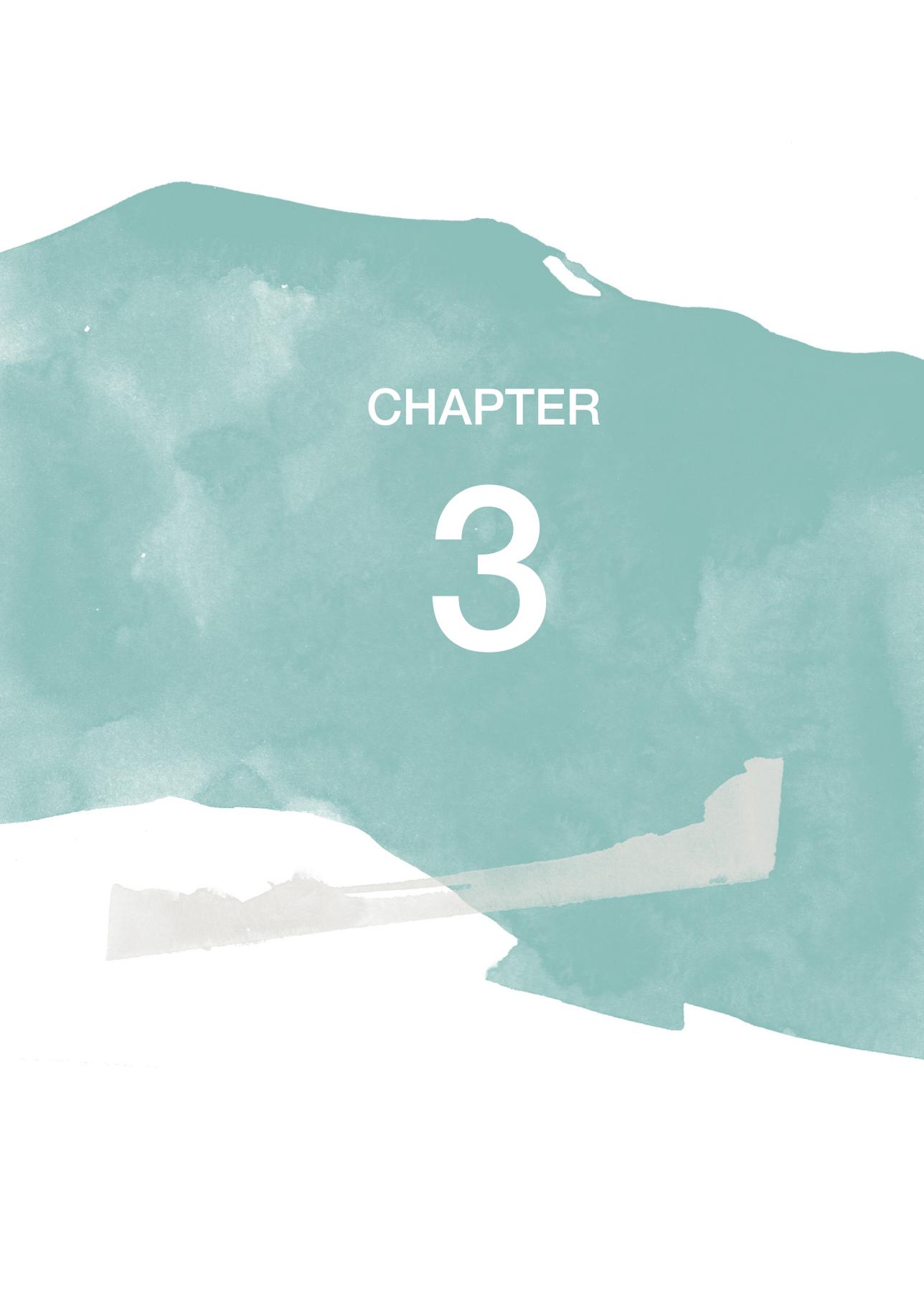
Competence criticism could, for example, be more about concrete behaviors (e.g., working faster), whereas moral criticism could be more abstract (e.g., not being trustworthy in general). To check for this possibility, we investigated whether the criticism referred to participants' concrete behavior or their abstract identity (see Supplementary Materials). This investigation confirmed that most of the criticism situations reported by participants were about concrete behaviors (e.g., not working fast enough), rather than about abstract identity (e.g., not being a trustworthy person). Further, we checked and confirmed that the observed effects of criticism dimension hold when controlling for abstractness/concreteness by adding it as a covariate to our analyses. We therefore conclude that it is unlikely that abstractness of criticism received confounds our results.

In conclusion, we show that also in individual-level criticism situations, people make more negative attributions about the critic and they see less motivation to improve when they are criticized by an outgroup member. This, in turn, negatively affects whether people change their behavior based on criticism. However, regardless of this group-membership effect, we demonstrate that being criticized on one's competence, compared to one's morality, makes people less defensive towards the criticism, more motivated to improve, and finally, makes them change their behavior more often in line with the criticism.

Please note that we do not claim that all moral failures can be framed as addressing someone's competence (or should be). Rather, we argue that in situations where both competence and moral aspects of problematic behavior can be made the focal point addressed with a criticism message (e.g., lying about having performed a work task while actually having failed to do so), focusing on lack of competence (failing to do the work task) versus lack of morality (lying about having performed the work task) displayed is more likely to get the critic "a foot in the door" towards implementing change, by making people less defensive. For example, because people are especially motivated to be seen as moral, critics often frame the primary focus of a criticism message as being about someone's morality in order to increase motivation to change. However, this can make people perceive the task of changing their behavior as too difficult and threatening (Van der

Lee et al., 2016), which is why they may more easily “give up” this task and get defensive or disengage (Bandura, 1982; Gausel & Leach, 2011; Täuber et al., 2015). Thus, emphasizing the moral dimension of criticism often backfires. In comparison, we show in the current research that the competence dimension of criticism can offer people an opportunity to grow and motivate them to change their own behavior.

Taken together, our findings suggest that criticizing someone’s competence rather than someone’s morality can make people more open to criticism, even when this criticism comes from the outgroup. Making people more susceptible to outgroup criticism would be very beneficial in all kinds of intergroup situations, for example, in a political discussion, where people with different political standpoints negotiate agreements that will apply to the whole nation. An open ear to other people’s criticism might result in a better outcome for all.

The background is a teal watercolor wash with a white torn paper effect at the bottom. The text is centered in the upper half of the image.

CHAPTER

3

Falling on Deaf Ears: The Effects of Sender Identity and Feedback Dimension on How People Process and Respond to Negative Feedback – an ERP Study

This chapter is based on:

Rösler, I. K., Van Nunspeet, F., Ellemers, N.. Falling on deaf ears: The effects of sender identity and feedback dimension on how people process and respond to negative feedback – an ERP study. *Under review*.

Author contributions:

IKR, FVN, and NE designed the studies. IKR performed the data analysis and wrote a first draft of the manuscript. FVN and NE provided substantial comments and suggestions for improvement.

Abstract

Social contexts can affect how people respond to feedback from others. We investigated how context information modulates the cognitive processing of feedback messages (i.e., external evaluations of one's character). We manipulated two aspects of (positive and negative) feedback messages: The identity of the sender (ingroup vs. outgroup member), and the dimension (one's competence vs. morality) as focal concern addressed in the feedback. We measured affective and behavioral responses after participants received such feedback (Study 2.1, $N = 194$), and additionally recorded an EEG in Study 2.2 ($N = 49$). In both studies, participants reported being more emotionally affected by negative feedback from ingroup compared to outgroup senders. Participants in Study 2.1 also reported to perceive feedback on their morality (vs. competence) as more negative. Complementing these findings, ERP results of Study 2.2 revealed greater preferential processing (i.e., increased P200) of the feedback messages from ingroup rather than outgroup members. Additionally, participants paid less sustained attention to feedback on their morality (vs. competence, as indicated by decreased P300- and LPP-amplitudes), and afterward recalled less morality- (vs. competence-) related feedback messages. The ERP findings were more pronounced for negative compared to positive feedback. These results suggest that subtle cues such as the social group-membership of a sender or the dimension addressed in a feedback message can modulate the cognitive processing of that message. Furthermore, our findings may explain why people are inclined to disregard negative feedback from outgroup senders and on their moral character.

Humans constantly give each other feedback to regulate social behavior and to ensure norm adherence in groups (Ellemers et al., 2013; Ellemers & Van der Toorn, 2015). How successful this feedback is in regulating behavior may depend on the characteristics of both the sender of the message and the message itself. For example, whether the sender is considered an ingroup- or an outgroup member, and whether the message is positive or negative, influences the affective and behavioral responses of the receiver (Hornsey, Oppes, & Svensson, 2002; Tajfel & Turner, 1979). Specifically, people perceive negative feedback from outgroup senders as less constructive than negative feedback from ingroup senders (Esposito, Hornsey, & Spoor, 2013; Hornsey et al., 2002) and are therefore more reluctant to accept such feedback (Hornsey et al., 2004; Rösler et al., 2021a). Positive feedback, however, is perceived favorably regardless of whether it is coming from ingroup or outgroup senders (Hornsey et al., 2002). Another characteristic of a feedback message that may affect its impact is what dimension it addresses. For instance, whether one's morality or one's competence is the focal concern of a feedback message (Abele & Wojciszke, 2007; Brambilla & Leach, 2014; Fiske et al., 2007; Judd et al., 2005). That is, people are less inclined to accept feedback messages on their moral than on their competence failures because feedback on moral failures is perceived as less constructive (Rösler et al., 2021a), more harmful to the self-image (Pagliaro et al., 2016), and because people perceive this type of feedback as more difficult to cope with (Van der Lee, Ellemers, & Scheepers, 2016).

Most of the prior research on how people perceive and respond to feedback messages has relied on participants' self-reported recollections of such events, or their imagination of prescribed scenario's including such events. In the current research, we investigated how feedback is received in the actual moment and measured the cognitive processes revealing the receipt of such messages. Specifically, using Event-Related brain Potentials (ERPs), we examined participants' motivated attention to feedback messages coming from ingroup vs. outgroup members, addressing participants' (lack of) morality or competence. Moreover, we complemented these neuroscientific markers with self-report and behavioral data to understand how people deliberately respond to this feedback. By examining both these explicit and

more implicit measures we can reach a more comprehensive understanding of why certain feedback is (not) effective for regulating behavior and how effectiveness may be improved (Ellemers & van Nunspeet, 2020).

3.1 | Social feedback situations

Communication models suggest that both sender characteristics (i.e., who is sending a message) and message content (i.e., what is said) influence the communication process in social feedback situations (Berlo, 1960; Lasswell, 1948). Neuroscientific research has shown that these characteristics can influence very early perceptual and attentional processes (Klein et al., 2015; Schindler et al., 2020; Schindler & Kissler, 2018; Wieser et al., 2014). An example comes from an ERP experiment in which participants were presented with different target faces as well as context information about the different types of motives of these targets (Klein et al., 2015). The ERP findings showed that targets' socially threatening motives (e.g., criticizing someone) increased participants' sustained attention to the target faces – compared to physically threatening motives (e.g., wanting to hit someone) or neutral motives (e.g., wanting to give someone a bag). Another example comes from ERP-research where participants were provided with positive, negative, and neutral feedback messages about themselves and about unknown others, before they were presented with the faces of the feedback senders (Wieser et al., 2014). Here, early and late processing of the senders' faces was enhanced when the feedback was about the self (vs. another person) and when the feedback messages were negative or positive (vs. neutral). Thus, social context information such as the perceived motives of a person can influence which subsequently viewed stimuli are cognitively favored and receive more sustained attention.

Which stimuli are cognitively favored when being processed? One important factor in the selection is the self-relevance of a stimulus. According to the framework of motivated attention, the self-relevance of a stimulus can increase the motivated attention devoted to it (Lang, Bradley, & Cuthbert, 1997; Lang & Bradley, 2013). The underlying assumption of this framework is that both appetitive (e.g., approach) and defensive (e.g., avoidance) mechanisms modulate whether participants pay sustained attention to stimuli. This means

that people are more motivated to attend to pictures containing high arousing positive and high arousing negative content compared to neutral content. Motivated attention can be measured with two ERPs related to higher level processing and stimulus evaluation: the P300 and the Late Positive Potential (LPP, Lang et al., 1997; Polich, 2007; Schupp et al., 2004; Schupp, Flaisch, Stockburger, & Junghöfer, 2006). In social feedback situations, this means that feedback from a more relevant sender (i.e., a human) is processed more deeply, as indexed by increased P300 and LPP-amplitudes, compared to feedback coming from a less relevant sender (i.e., a computer, Schindler & Kissler, 2018; Schindler, Miller, & Kissler, 2020). These findings thus suggest that, besides contextual information, also the self-relevance of a stimulus and the arousal elicited by it can increase the motivated attention towards the stimulus.

We know from social psychological research that two other contextual factors can also influence whether we perceive feedback as relevant. Firstly, feedback from senders belonging to the same social group (i.e., ingroup, Cikara & Van Bavel, 2014; Tajfel & Turner, 1979) is perceived as more relevant than feedback from outgroup senders (Ellemers et al., 2013; Hornsey et al., 2002). And secondly, feedback on one's own morality is perceived as less relevant than feedback on one's own competence (Rösler et al., 2021a; Täuber & van Zomeren, 2013; Van der Lee et al., 2016). With the current research, we investigated whether these two factors also modulate the motivated attention towards feedback messages. If people perceive feedback from ingroup members as more self-relevant than feedback from outgroup members, and feedback on their competence as more self-relevant than feedback on their morality, they should also pay more attention to feedback messages from ingroup (vs. outgroup) members and to feedback on their competence (vs. morality). By investigating this question, we extend prior research by combining insights from different theoretical research areas and literatures from social psychology and neuroscience.

3.2 | Sender's social group-membership

When interpreting social feedback situations, people usually start by noticing the identity of the feedback sender, thus whether they are an ingroup or

outgroup member. From both social psychological and neuroscientific research, we know that the group-membership of an interaction partner can shape how people perceive a person and react to them (Cikara & Van Bavel, 2014; Ellemers, 2012; Tajfel & Turner, 1979). This means that people are often more responsive to behavioral recommendations voiced by ingroup (vs. outgroup) members (Ellemers et al., 2013; Pagliaro et al., 2011). That is, because following these recommendations can potentially secure respect and inclusion from fellow group members. Moreover, negative feedback from outgroups often make receivers defensive because the feedback is perceived as less constructive than ingroup feedback and potentially harmful to the self or ingroup (Esposito et al., 2013; Hornsey et al., 2002, 2004; Rösler et al., 2021a). Feedback from ingroup (vs. outgroup) members is, therefore, more self-relevant and potentially more beneficial in terms of information gathering (e.g., how one can earn respect, inclusion, and be a good group-member, Ellemers et al., 2013). Thus, people might also be more motivated to pay attention to ingroup compared to outgroup feedback. Support for this claim comes from research investigating how people cognitively process feedback they receive on their character by examining particular brain potentials (Schindler et al., 2020; Schindler & Kissler, 2016, 2018). Results showed that feedback from a relevant rather than a less relevant sender (i.e., a human [expert] vs. computer) increased participants' motivated attention to the feedback. More specifically, the processing of feedback messages from a human (vs. computer) sender was associated with increased amplitudes of the P200, the P300 and the LPP. The P200 potential has been associated with the initial lexical encoding of emotional words (Kissler, Assadollahi, & Herbert, 2006; Trauer, Andersen, Kotz, & Müller, 2011). The P300 and LPP potentials are thought to give an indication about whether participants pay sustained attention to the processing of words (Lang et al., 1997; Polich, 2007; Schupp et al., 2004; Schupp, Fleisch, Stockburger, & Junghöfer, 2006). The increased amplitudes of these potentials thus suggest that the attentional processing of feedback coming from more self-relevant senders was facilitated compared to when such messages were delivered by less self-relevant senders. People paid both more initial and more sustained attention to such feedback messages. Extending this prior research, we aimed to

test whether manipulating the social group-membership of a feedback sender, in terms of ingroup vs. outgroup, modulates attentional processing of feedback messages in a similar vein. We expected that people would be more motivated to attend to feedback from ingroup (vs. outgroup) members, as reflected in higher P200-, P300-, and LPP-amplitudes when viewing the feedback messages.

After noticing sender characteristic such as the identity of a sender, the next step in information processing would be to notice characteristics of the feedback message itself, such as the valence and dimension of the feedback.

3.3 | Valence and dimension of feedback message

Concerning the potential valence of the message, social psychological research would suggest that participants show a general negativity bias in processing feedback (Baumeister et al., 2001; Rozin & Royzman, 2001). This bias is referring to the effect that negative information, even when being of the same intensity as positive information, has a greater impact on people's cognition (e.g., dwelling on negative information about the self, Baumeister et al., 2001). Whether the negativity bias is also reflected in the cognitive processing of positive and negative information (using ERPs), is debated in the literature. Some research finds that the processing of and motivated attention towards negative information is enhanced compared to the processing of positive information (i.e., increased P200 and LPP, Ito, Larsen, Smith, & Cacioppo, 1998; Carretie, Mercado, Tapia, & Hinojosa, 2001). Other research does not find differences for valence (i.e., on the P300 and LPP, Schindler & Kissler, 2016, 2018). In the current research, we aimed to increase the understanding of how negative feedback in particular is perceived and processed. Yet, we included positive feedback in our paradigm as a comparison condition and to increase the credibility of the experiment because participants received multiple rounds of feedback in the experimental task. We expected that participants would be more emotionally affected by negative (vs. positive) information about themselves and investigated whether this was also reflected in brain responses associated with the processing of the feedback (e.g., motivated attention, P300, LPP).

Besides being framed as negative or positive, feedback can refer to different dimensions of social evaluation. Here, we examined how people receive and process negative feedback on their morality (e.g., being perceived as uncooperative) compared to negative feedback on their competence (e.g., being perceived as incompetent). We thus compared feedback addressing the two fundamental dimensions people use to judge others (i.e., the Vertical dimension for 'getting along', including communion/warmth and morality; and the Horizontal dimension for 'getting ahead', including agency and competence; Abele-Brehm, Ellemers, Fiske, Koch, & Yzerbyt, 2020; Koch, Yzerbyt, Abele, Ellemers, & Fiske, 2020). For social evaluations, morality is of special importance as it refers to the norms of our societies and regulates the relations between individuals living in these societies (Ellemers, 2017; Ellemers et al., 2013). Moral information can give us cues about the potentially harmful intentions of others (e.g., Abele & Wojciszke, 2007; Brambilla & Leach, 2014; Cuddy et al., 2008) and is central to how we see ourselves and are seen by others (Goodwin et al., 2014; Strohminger & Nichols, 2014). This importance of morality also makes people motivated to show appropriate behavior when their moral image is at stake, such as when other people are watching (Bateson, Nettle, & Roberts, 2006; Van Nunspeet, Derks, Ellemers, & Nieuwenhuis, 2015). However, when dealing with negative feedback on their own moral character or group, the importance of morality puts people in a self-protective state (Gausel & Leach, 2011). When people are criticized for their morality (vs. competence), they often act defensively (Rösler et al., 2021a; Täuber & van Zomeren, 2013). In fact, people often turn to coping mechanisms to deal with such painful evaluations such as justification of their moral failures (Mazar et al., 2008; Shalvi et al., 2015). This goes as far as misremembering past moral failures (Carlson et al., 2020; Kouchaki & Gino, 2016), distancing the self when reflecting on past moral failures by emphasizing having changed since the occurrence (Stanley et al., 2019), and not seeing the necessity to change one's moral character (Sun & Goodwin, 2020). While this can protect well-being, it impedes behavioral improvement and can hinder people from perceiving and using feedback as an opportunity to grow (e.g., Hornsey & Esposito, 2009). This is why researchers have tried to investigate alternative ways of delivering feedback to avoid people engaging

in coping mechanisms and responding defensively. One of these ways is to frame feedback given in response to someone's failures in terms of their competence rather than their morality. Criticism on one's competence (vs. morality) makes it easier for people to cope with being evaluated (Van der Lee et al., 2016) and is perceived as less harmful to the self-image (Pagliaro et al., 2016). Moreover, it can make people more motivated to improve, rather than getting defensive (Rösler et al., 2021a; Täuber & van Zomeren, 2013). Most of the research examining the effects of morality- vs. competence-framed feedback or criticism has either focused on participants' recollection or imagination of events in which they received such feedback. In the current research, we extended this research by investigating how (negative) feedback is received in the actual moment. In line with previous research suggesting that negative feedback on one's morality is less effective and less relevant than feedback on one's competence, we expected that participants would be less motivated to pay sustained attention to moral (vs. competence) feedback, as reflected in decreased P300-, and LPP-amplitudes (Rösler et al., 2021a; Täuber & van Zomeren, 2013)⁶.

Moreover, we expected that the affective and behavioral measures complement these findings and based our predictions on previous social psychological literature. That is, we expected that negative feedback of participants' morality is perceived as more negative than on participants' competence (Pagliaro et al., 2016; Rösler et al., 2021a; van der Lee et al., 2016). Furthermore, we expected that when given an opportunity to make up for immoral behavior, participants would do so by showing moral behavior by increasing money contributions from T1 to T2 (also see below, moral cleansing, Brañas-Garza, Bucheli, Paz Espinosa, & García-Muñoz, 2013; Van der Toorn, Ellemers, & Doosje, 2015). And finally, based on previous research on intergroup criticism, we expected that participants would be more receptive of ingroup compared to outgroup feedback, as indicated by a decreased emotional and defensive response when receiving negative feedback (Ellemers et al., 2013; Pagliaro, Ellemers, & Barreto, 2011; Esposito

⁶Based on the reviewed social-psychological findings we would expect this effect to be more pronounced for negative feedback. However, as the ERP-findings are inconclusive on whether valence modulates ERPs related to motivated attention, we did not include this prediction here.

et al., 2013; Hornsey et al., 2002, 2004; Rösler et al., 2021a).

3.4 | The current research

In the current research, we aimed to investigate immediate brain responses to receiving feedback messages on one's morality or competence, either verbalized by ingroup or outgroup members. We modeled these social feedback situations in an experimental paradigm, in which participants were first asked to perform a behavioral decision-making task. Then participants received feedback, in the form of trait evaluations of their character, from ostensible senders who were presented as ingroup or as outgroup members judging participants' behavioral decisions on the task. We included questionnaires to measure participants' self-reported affective and defensive responses to receiving feedback. And we included a repeated monetary contribution task (T1: Before feedback, T2: After feedback) to measure whether the feedback influenced behavioral decision making.

In Study 2.1, conducted online, we showed that participants perceived negative feedback messages on their morality as more negative than on their competence. Contrary to our initial expectations, we found that participants reported to be more emotionally affected and defensive towards feedback from the ingroup vs. the outgroup. In Study 2.2, conducted in the lab, we replicated the findings of Study 2.1. Moreover, complementing the self-report findings of Study 2.1, ERP-results showed that participants preferentially processed feedback messages delivered from ingroup (vs. outgroup) members. Moreover, participants were less motivated to attend to (negative) feedback messages on their morality compared to their competence and recollected less of these messages afterwards.

3.5 | Study 2.1

3.5.1 | Methods

3.5.1.1 | Design and participants

Participants performed a task in which they made behavioral decisions (see Instruments) and received feedback on their behavior from ostensible ingroup members (i.e., fellow students) or outgroup members (i.e., professional trainees). The feedback was either negative or positive. These two factors

were varied within-participants, using a randomized, rather than blocked experimental design. The third factor of interest (i.e., dimension of feedback) was varied between participants: Feedback was either about participants' morality or about their competence. The experiment, therefore, had a 2: Sender's group-membership (in- vs. outgroup member, within) x 2: Feedback valence (positive vs. negative, within) x 2: Feedback dimension (morality vs. competence, between) design.

We based sample size calculations on previous research (Rösler et al., 2021a) where participants were asked to reflect on receiving negative feedback by either an ingroup or outgroup member, on their morality or competence. Here, the effect of feedback dimension (i.e., morality vs. competence) on whether participants accepted and changed their behavior in accordance with the feedback was $\eta^2_p = .05$. We used this as a proxy for our current research questions where we tested whether the feedback dimension affects people's affective and behavioral responses to feedback. Sample size calculations in G*power ($\alpha = .05$, $1-\beta = .80$) revealed a minimum sample of 154 participants using repeated measure ANOVAs. We oversampled this number to be able to test for all effects of our manipulations (e.g., effect of sender's group-membership on affective responses to feedback).

A total of 201 UK university students took part in our online experiment on Prolific (<https://prolific.ac>) and were paid €4. All participants had agreed on the research platform to being deceived. Seven participants who failed both attention checks were excluded. The final sample size consisted of 194 participants ($M_{age} = 26.57$, $SD_{age} = 7.92$, 120 female, 2 other). The study was approved by the local Ethics Review Board.

With the current research, we aimed to investigate how participants respond to receiving feedback on their character in terms of their morality or competence, by ingroup or outgroup members. To make this feedback credible in an experimental setting, we made use of a 'social dilemma' situation. These situations are qualified by people having to choose between their self-interest and the long-term interests of a group (e.g., Komorita & Parks, 1994). We first asked participants to make monetary contributions to a common good, the realization of which would depend on many people

contributing. These contribution decisions would then be shown to others who were ostensibly asked to judge (and provide feedback on) the participant's character, based on how they evaluated the participant's decisions (not) to contribute to the common good (see Figure 3.1).

After soliciting informed consent, we described this cover story to participants, and explained the common good: A crowdfunding project for new housing. Participants were requested to indicate their willingness to contribute to this common good by, hypothetically, donating part of their monthly income for one year (resulting in twelve contributions). Participants were also told that the contributions they made were presented to several other participants taking part in our experiment (ostensibly in real-time, on Prolific), and that these others (i.e., the 'senders' of the feedback) were asked to evaluate the participant's character based on their contributions by selecting traits they thought were indicative of the participant's character. Finally, we explained that half of these feedback senders were presented as ingroup members (i.e., fellow students) and the other half as outgroup members (i.e., professional trainees).

After reading the instructions, participants were asked to perform the experimental task in which they had to make monetary contributions to the common good (T1). Then, the feedback phase started (see Instruments), where participants were told they would receive the feedback messages ostensibly chosen for them by other participants. The feedback consisted of one character trait per trial and was delivered by one feedback sender that could either belong to the ingroup or the outgroup. After the feedback phase was over, participants were first asked to reflect on the perceived valence of all the character traits they had just received from the ingroup and the outgroup separately (i.e., group-based feedback). Then, they received a message summarizing the valence of the group-based feedback (i.e., averaged amount of negative and positive traits chosen for them), also for the ingroup and outgroup separately (see Instruments for wording of the message). We added these group-feedback messages to the experiment because we delivered the same amount of positive and negative character traits to participants in the feedback phase, but we aimed to investigate the

affective responses to negative feedback in particular. Therefore, both of these group-feedback messages were negative. Then, participants were asked about their affective responses to the group-feedback messages. After this, participants were asked to make the contribution decisions a second time (T2). To increase face validity of the experiment, participants themselves were also asked to judge (the contributions made by) ostensible other participants at the end of the experiment. The experiment lasted 30 minutes. After completion, participants were fully debriefed and thanked for taking part in the study.

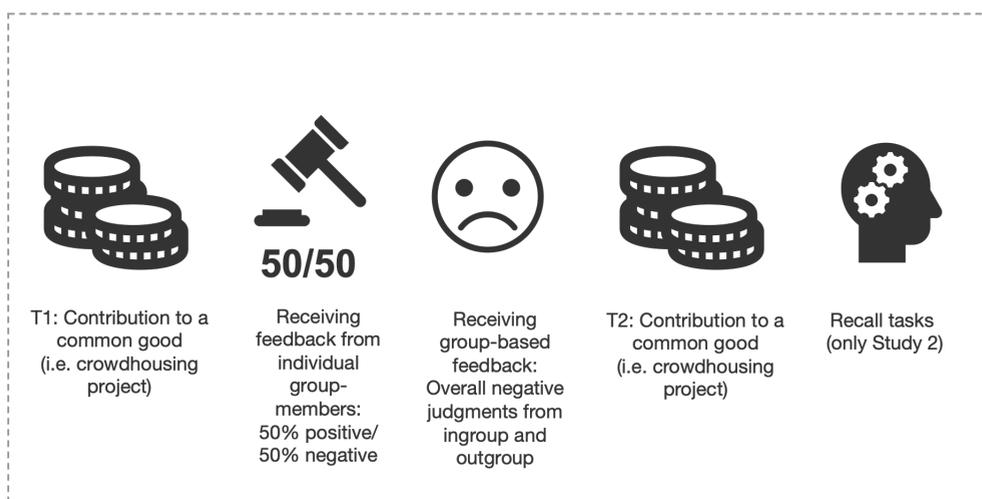


Figure 3.1. Overview of experimental procedure.

3.5.1.2 | Instruments and stimuli

The experimental task was presented using the experiment builder Gorilla (<https://gorilla.sc>). In the following, we go into detail about the different phases of the experimental paradigm. As explained above, each participant first made monetary contribution decisions (T1), then went through the feedback phase where individual senders delivered feedback. They then received two group-feedback messages which summarized the previous feedback messages as an overall more negative than positive judgment, and then made the contributions decisions a second time (T2).

3.5.1.2.1 | Contribution to common good T1 & T2

Participants made 12 contribution decisions to a common good (i.e., a crowdfunding project for new housing): One for each month of one hypothetical year. They were asked to indicate how much of their monthly income (after paying all bills), they would be willing to contribute to the common good. They could choose between 0-30% of their income. We added a monetary incentive (i.e., one €20 voucher) for the person with the lowest contributions to highlight they might personally benefit by contributing less instead of more to the common good. We hoped this would ensure variability in participant's contribution decisions and enhance credibility of receiving both negative and positive feedback on their behavioral decisions.

3.5.1.2.2 | Feedback phase: Feedback messages from individuals

Participants were presented with feedback in response to the decisions they made. This was provided in the form of character traits, ostensibly selected by other participants in response to participants' behavior, but which we preprogrammed (experimental paradigm adapted from Schindler & Kissler, 2018; Schindler, Wegrzyn, Steppacher, & Kissler, 2014). Each trial consisted of a fixation cross (250ms), followed by a sender's face and background color alternating ingroup and outgroup membership (1000ms), and the character trait indicating the feedback message (either related to competence or morality, depending on condition). The feedback was presented to participants in a random order.

Each participant received 32 feedback messages, consisting of character traits (see Supplementary Materials for list). Depending on the experimental condition, all traits presented to a participant were either related to their morality (e.g., [un]cooperative, [im]moral) or to their competence (e.g., [un]intelligent, [in]competent). Fifty percent of the feedback messages were positive, and fifty percent negative. Specifically, participants were presented with 16 different traits (eight positive, eight negative) and each trait was presented twice.

For the feedback senders, we selected sixteen Caucasian faces (8 female, 8 male) from the Radboud faces database (Langner et al., 2010). All (faced-

forward) faces showed a neutral expression. Each participant received feedback from fifty percent ingroup members (i.e., fellow students, indicated by a blue background color) and fifty percent outgroup members (i.e., professional trainees, indicated by a yellow background color). To learn these group-memberships, participants were asked to perform a categorization task before the feedback phase. They had to correctly categorize each face twice.

3.5.1.2.3 | Group-based feedback

In the feedback phase, participants were presented with fifty percent positive and fifty percent negative feedback messages (i.e., character traits) from individual ingroup and outgroup members. As a next step in our paradigm, we asked participants to briefly reflect on the overall valence of the messages they had received during the feedback phase. We asked them to indicate whether they perceived the feedback messages (from the ingroup and from the outgroup, i.e., two questions) as more positive (answer = 1) or negative (answer = 7)⁷.

Additionally, because of our focus on responses to negative feedback in particular, we then presented participants with the (manipulated) group-based feedback. Specifically, we explained that the valence of all delivered feedback messages was averaged for both the ingroup and the outgroup and that both groups evaluated their moral or competent character (depending on condition) as more negative than positive (e.g., “The students, on average, judged your morality as: Negative (= Immoral)”).

3.5.1.2.3 | Self-report measures

All self-report measures were presented on a 7-point scale (1 = not at all, 7 = very much), unless indicated otherwise.

3.5.1.2.3.1 | Manipulation checks

To check our between-participants manipulation of dimension, we asked participants to indicate the nature of the feedback they had received. They

⁷ The attentive reader might wonder whether this scale was scored correctly, with higher numbers indicating more negativity. That was indeed the case, and the labels were made clear to the participant.

could indicate whether they received feedback on their morality or competence with a bipolar scale (“To what extent did the feedback messages refer to your morality rather than your competence?”, 1 = morality, 7 = competence). Besides including a categorization task where participants had to correctly categorize each face twice, we also checked for self-relevance of the group-memberships. We asked participants whether they identified with the ingroup (i.e., students) and outgroup (i.e., trainees) with a three-item measure of identification (e.g., “I identify with other members of this group”, Ellemers, Kortekaas, & Ouwerkerk, 1999, ingroup, $\alpha = .90$, outgroup, $\alpha = .89$).

3.5.1.2.3.2 | Affective and defensive responses to group-based feedback

We measured participants’ affective responses to the two negative group-based feedback messages (for averaged ingroup judgments and outgroup judgment separately) with a selection of items of past research (PANAS, Brockner & Higgins, 2001; Rösler et al., 2021a; Watson et al., 1988). The basis of the current cluster of items was based on previous research (Rösler et al., 2021a), in which we had found evidence for a meaningful distinction between these measures. We selected items that were relevant for the context of the current research which slightly differed from the previous research (see Supplementary Materials for items). For affective responses to negative feedback we combined four items to the measure *negative emotions after group-based feedback* (e.g., “The feedback from the [students/trainees] made me feel upset”, ingroup $\alpha = .90$, outgroup $\alpha = .92$), three items to the measure *positive emotions after group-based feedback* (e.g., “The feedback from the [students/trainees] made me feel confident”, ingroup $\alpha = .79$, outgroup $\alpha = .85$), and two items to the measure *moral emotions after group-based feedback* (e.g., “The feedback from the [students/trainees] made me feel guilty”, ingroup $r[192] = .75$, $p < .001$, outgroup $r[192] = .82$, $p < .001$).

We measured defensiveness, that is whether participants held negative assumptions about feedback senders, with seven items (e.g., “I think the [students/trainees] who gave feedback to me did this in my best interest”, Hornsey, Trembath, & Gunthorpe, 2004). To improve reliability, we excluded two items (see Supplementary Materials). We recoded positively framed items, so higher scores on this scale indicate more defensiveness. We combined

the five remaining items into an overall measure assessing *defensiveness after group-based feedback* (ingroup $\alpha = .79$, outgroup $\alpha = .78$).

3.5.2 | Results and Discussion

In line with the timeline of our paradigm (see Figure 3.1), we first report manipulation checks, then the self-report results, and finally the behavioral results.

3.5.2.1 | Checks

As intended, a comparison of the manipulation check for feedback dimension (bipolar scale, i.e., 1 = morality, 7 = competence) to the midpoint of the scale (i.e., 4) showed that in the morality condition, participants were more inclined to report having received feedback on their morality ($M = 3.02$, $SD = 1.62$), $t(92) = -5.84$, $p < .001$, $CI = [-1.31, -.65]$. In the competence condition they indicated the feedback applied more to their competence ($M = 4.89$, $SD = 1.83$), $t(100) = 4.89$, $p < .001$, $CI = [.53, 1.25]$. Both means were also significantly different from each other, $t(192) = -7.51$, $p < .001$, $CI = [-2.36, -1.38]$.

A paired-sample t-test revealed that participants identified more with ingroup ($M = 4.03$, $SD = 1.71$) as compared to outgroup senders ($M = 2.75$, $SD = 1.27$), $t(193) = 9.15$, $p < .001$, $CI = [1.00, 1.55]$, as intended. Moreover, participants were able to correctly categorize the sender faces as belonging to the ingroup and outgroup (as checked with our categorization task where participants had to correctly categorize each face twice).

3.5.2.2 | Self-report and behavioral results

3.5.2.2.1 | Perceived valence of feedback from individuals

As an indirect check of the impact of feedback messages delivered during the feedback phase, we asked participants to reflect on the perceived valence of the character judgments (i.e., traits) they received (i.e., bipolar measure, 1 = positive, 7 = negative⁷). In line with prior social-psychological theory (e.g., Rozin & Royzman, 2001) and reflective of the overall negative judgment they received to summarize the feedback (right after this measure in the experiment), a one-sample t-test against the midpoint of the scale (i.e.,

4) showed a general negativity bias. Participants perceived the feedback as more negative than positive ($M = 4.15$, $SD = .73$), $t(193) = 2.96$, $p = .003$, $CI = [.05, .26]$. More interestingly and as predicted, morality (vs. competence) feedback was perceived as more negative. A mixed RM (M)ANOVAs (Feedback dimension [between], Sender's group-membership [within]) showed a main effect of feedback dimension on perceived valence, $F(1, 192) = 6.37$, $p = .012$, $\eta_p^2 = .03$. Even though the valence of the feedback was equated in the task, participants thought the feedback was more negative when they had received feedback on their morality ($M = 4.29$, $SD = .76$) compared to on their competence ($M = 4.03$, $SD = .68$, between-participants factor). There were no other effects, $F_s < 1$.

3.5.2.2.2 | Affective and defensive response to group-based feedback from the ingroup and outgroup

For all affective and defensive responses towards group-based feedback, we computed mixed RM (M)ANOVAs (Feedback dimension [between], Sender's group-membership [within]). For negative, positive, and moral emotions, we computed a MANOVA and used an alpha level of .016 (Bonferroni correction, $.05/3$) to correct for multiple comparisons. Results revealed an effect of sender's group-membership at the multivariate level, Pillai's Trace = .11, $F(3, 190) = 7.43$, $p < .001$, $\eta_p^2 = .11$. Contrary to our expectations, participants reported stronger negative emotions after an overall negative feedback message given by the ingroup ($M = 3.74$, $SD = 1.61$) than by the outgroup ($M = 3.48$, $SD = 1.66$), $F(1, 192) = 20.03$, $p < .001$, $\eta_p^2 = .09$, at the univariate level. There were no other effects, $F_s < 1$. Similarly, for defensiveness, participants made more negative assumptions about the group after receiving an overall negative feedback message from the ingroup ($M = 4.81$, $SD = 1.16$) than from the outgroup ($M = 4.54$, $SD = 1.15$), $F(1, 192) = 26.06$, $p < .001$, $\eta_p^2 = .12$. There were no other effects, $F_s < 1$.

3.5.2.2.3 | Contribution to common good T1 & T2

We checked how much money participants contributed to the common good and whether the contributions varied between T1 & T2. For this, we averaged the contribution decisions before further analysis. We used these averages as dependent variables in a mixed RM ANOVA (Feedback dimension [between],

Time-point contribution decision: T1-T2 [within]). Overall, there were no effects of dimension or time-point contribution, $F_s < 1$. Thus, participants did not seem to contribute more money in the morality condition than in the competence condition (as a form of moral cleansing or opportunity to repair their moral image).

In conclusion, as predicted, participants perceived feedback on their morality as more negative than feedback on their competence, even though the valence of feedback was equated. Contrary to our predictions, we found that negative group-based feedback from the ingroup (vs. outgroup) seemed to elicit a stronger, rather than a less strong, affective and defensive response in participants. This may suggest that participants were more emotionally impacted by being negatively evaluated by the ingroup, as compared to the outgroup. Since we did not find that participants engaged in moral cleansing by contributing more money at T2 after receiving negative feedback on their morality (vs. competence), we do not discuss this measure here further.

3.6 | Study 2.2

With Study 2.1, we found that participants judged feedback on their morality (vs. competence) as more negative and that they were more emotionally impacted by negative group-judgments by the ingroup (vs. outgroup). In Study 2.2, we aimed to replicate these findings and investigate whether participants are less motivated to pay sustained attention to feedback messages coming from outgroup compared to ingroup members, and addressing their morality, as compared to their competence. Moreover, we addressed two potential drawbacks of Study 2.1 related to the believability of the task and the relevance of the social group-membership of the feedback senders. More specifically, in Study 2.2 we eliminated potential doubts of receiving feedback messages from other people in an online setting by testing participants in the lab. Additionally, we changed the social group-memberships from 'students' (i.e., ingroup) and 'trainees' (i.e., outgroup) to 'Dutch students' and 'international students' (group-memberships respectively). In Study 2.2, both ingroup and outgroup members were therefore students with the same amount of professional experience. This not only increased relevance of group-membership to our student participants, but also addresses the

potential confound that either group is more equipped to give feedback and thus is more attended to.

3.6.1 | Method

3.6.1.1 | Design and participants

The experimental design was the same as in Study 2.1. Participants were asked to complete a contribution task making use of a social dilemma situation in which they had to make contributions to a common good (i.e., crowdfunding housing) and received feedback on their character from ostensible others based on their decisions. We switched feedback dimension (morality vs. competence) from a between-participants to a within-participants factor to increase power (to decrease the total number of participants needed for this ERP study), and because ERP-research needs a high number of trials presented to participants (e.g., Woodman, 2010).

Sample size calculation was based on previous ERP-research that used a similar experimental paradigm (Schindler & Kissler, 2018). This research found effects of the sender's identity (i.e., computer vs. human) and feedback valence (i.e., positive and negative vs. neutral feedback) of a feedback message on the P300, $\eta_p^2 = .45$, early LPPs (400-650 ms), $\eta_p^2 = .31$ and $\eta_p^2 = .23$, and on the late LPP (650-900 ms), $\eta_p^2 = .18$. Sample size calculation with G*power ($\alpha = .05$, $1-\beta = .80$) revealed a required sample size between 13–39 for the main effects of sender and valence.

We tested 49 right-handed Dutch students ($M_{\text{age}} = 21.20$, $SD_{\text{age}} = 2.15$, two participants not indicating age, 39 female) with no history of neurological or psychiatric problems, and normal to corrected vision. For the ERP analyses, we had to exclude two participants due to technical problems and six participants due to poor data quality (i.e., not enough trials left after pre-processing), resulting in an ERP sample size of 41. Behavioral and self-report analyses are based on all 49 participants. The study was approved by the local Ethics Review Board.

3.6.1.2 | Procedure

The procedure was similar to Study 2.1 (see Supplementary Materials for

details about minor changes in the contribution task). We slightly adapted our cover story as this study took place in the lab (vs. online). In Study 2.1, we increased credibility of the feedback by explaining to participants that their monetary contributions would be presented to and evaluated by several other participants taking part in the online experiment. In the current study, we explained to participants that the feedback messages they would receive were based on a previous experiment with a similar set-up. Specifically, we explained to participants that they would be presented with feedback that was given to prior participants who had made similar contribution decisions as they did, and that these feedback messages were therefore also indicative of how their character would have been evaluated. The experiment lasted 90 minutes. After completion, participants were fully debriefed, thanked, and compensated with course credit or €18.

3.6.1.3 | Instruments and Stimuli

The experimental task was presented on E-prime, instructions and self-report measures were presented using Gorilla.

3.6.1.3.1 | Feedback phase: Feedback messages from individuals

Participants were presented with 448 trials, 56 trials per condition of our experimental design (i.e., 2: Sender's group-membership x 2: Feedback valence x 2: Feedback dimension). All factors were varied within-participants and participants were presented with the same, randomized, order of trials. In line with related prior ERP studies (e.g., Schindler & Kissler, 2018), we added an affirmation phase in the trials of this study to ensure that our ERP analyses would focus on activity related to receiving the feedback rather than the language processing of the word. We explained to participants that the feedback senders were presented with several traits and had to select which traits they thought were representative of the character of the participant. Specifically, after the presentation of the trait, a change of color indicated whether the feedback sender affirmed they thought this trait was indicative of the participant's character (see also Figure 3.2). Each trial thus started with a fixation cross (jittered duration, 350-750 ms), followed by a sender's face and background color alternating ingroup and outgroup membership of the sender (500 ms), a second fixation cross (jittered duration, 350-750 ms),

the presentation of the character trait (500 ms), and the affirmation decisions indicating whether this was the feedback that pertained to their decision (i.e., feedback message turned orange [vs. stayed black if it was not affirmed]; min. 1000 ms).

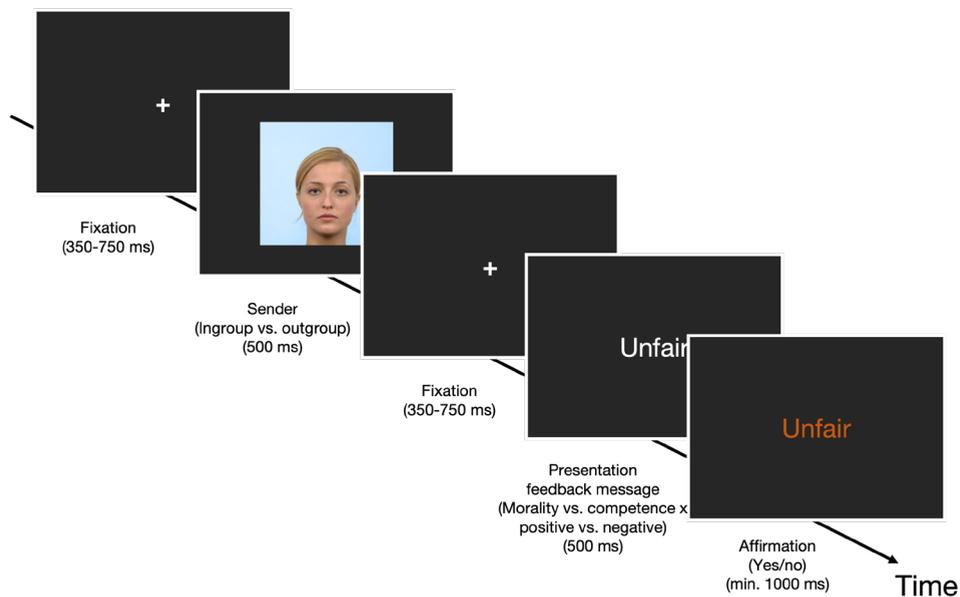


Figure 3.2. Example of trial feedback-phase.

For the feedback messages, we translated and extended the trait list (see Supplementary Materials) and pretested the resulting trait terms with a Dutch convenience sample ($N = 100$, $M_{\text{age}} = 26.4$ years, $SD = 9.43$, 65 females, 1 other). Participants rated arousal and valence associated with traits using the Self-Assessment Manikins (SAM, Bradley & Lang, 1994, 9-point scale), familiarity of traits (9-point scale, 1 = not at all, 9 = very much), and dimension of trait (bipolar scale, 1 = morality, 9 = competence). Frequency (per million) and length of trait words were taken from the database of Dutch word frequencies 'SUBTLEX-NL' (Keuleers, Brysbaert, & New, 2010, <http://crr.ugent.be/isubtlex/>). We selected 64 traits: 16 positive competence-related traits, 16 negative competence-related traits, 16 positive morality-related traits, and 16 negative morality-related traits. As intended, the results only revealed differences between positive and negative words for valence and between competence and morality words for dimension (see Table 3.1).

Table 3.1: Means, Standard deviations, and F-statistics for trait stimuli per condition for Study 2.2

		Morality Negative (N = 16)	Morality Positive (N = 16)	Competence Negative (N = 16)	Competence Positive (N = 16)	F(1, 60)
Valence	Dimension	2.39 (.55)	7.39 (.67)	2.68 (.55)	7.49 (.47)	1.80
	Valence					1201.18***
	Dimension*					.45
Dimension	Valence					2.85
	Dimension	3.12 (.87)	3.37 (.76)	6.58 (1.07)	7.08 (.82)	260.93***
	Dimension*					.30
Arousal	Valence					1.23
	Dimension	5.72 (.46)	5.71 (.55)	5.43 (.53)	5.70 (.26)	1.68
	Dimension*					1.46
Familiarity	Valence					2.38
	Dimension	7.50 (.36)	7.59 (.51)	7.53 (.32)	7.75 (.36)	.90
	Dimension*					.34
Frequency (million)	Valence					.29
	Dimension	21.60 (66.56)	17.67 (43.11)	14.09 (30.83)	6.50 (8.96)	.76
	Dimension*					.03
Length	Valence					.95
	Dimension	10.38 (3.10)	8.69 (2.39)	8.75 (2.86)	9.13 (2.33)	.78
	Dimension*					2.35

Note: Each trait was rated by at least 20 participants, *** $p < .001$

3.6.1.3.2 | Sender categorization check

To check whether participants differentiated between ingroup and outgroup senders, we included two checks. At the more explicit level, we included the social categorization task at the end of the experiment, this time without the background colors, to check whether participants could still correctly categorize the faces. In addition, we examined ERPs indicating the social categorization of sender faces (i.e., N100, N170, P200, N200) before receiving feedback from the senders (see Supplementary Materials). To be able to investigate these ERPS associated with social categorization, we made some changes to the face stimuli relative to the stimuli we used in Study 2.1. Specifically, for Study 2.2, we matched the sender's gender to the participant's gender and controlled for the face's attractiveness (no effects of the sender's group-membership and gender on attractiveness, all $F_s < 1$, $p_s \geq .896$). Therefore, we enlarged the face stimuli set and chose a slightly different selection (16 Caucasian faces [i.e., 8 female, 8 male], Langner et al., 2010). We selected four faces for the ingroup (i.e., Dutch students, blue background) and four faces for the outgroup (i.e., international students, yellow background).

3.6.1.3.3 | Recall task and self-report measures

As a validation and behavioral measure of motivated attention to the feedback messages, we added a recall task. If participants were to pay more attention to feedback on their competence, compared to their morality, they should also remember these feedback messages better afterward. We asked participants to list as many traits they could remember after the feedback phase. Participants were not informed about being asked to perform this task beforehand. We used the same self-report measures as in Study 2.1⁸.

3.6.1.3.4 | EEG acquisition

EEG was recorded from 23 electrodes embedded in a stretch head cap, positioned according to the 10-10 system: F7, F3, Fz, F4, F8, FC3, FCz, FC4, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P7, P3, Pz, P4, P8, POz, and Oz.

⁸ For two measures there was lower reliability in one of the two within-participants conditions, possibly due to translation of the items from English to Dutch and a decreased sample size (see Supplementary Materials). We kept the same items as in Study 2.1 for comparability

We used the Biosemi active-electrode recording system and a sampling rate of 256 Hz. BioSemi applies an analog hardware filter at 1/3 of the sample frequency to prevent aliasing. Electrode impedance was kept below 5 k Ω . CMS and DLR were used as initial reference and ground for recording voltages. Horizontal and vertical eye movements were recorded to correct for eye movements. EEG activity was recorded with ActiView software.

Offline, EEG was re-referenced to the average of the left and right mastoids using Brain Vision Analyzer 2.1 (Brain Products GmbH, Munich, Germany). The signal was corrected for ocular artifacts using the regression approach (Gratton et al., 1983), filtered (0.01-30Hz), and trials with movement artifacts were rejected. We created epochs for the feedback stimulus (i.e., trait word) ranging from -200 prior to 800 ms after the event. Epochs were averaged and baseline corrected with the average signal 200 to 0 ms before the event. Separate epochs were created for each of the eight conditions of our experimental design (i.e., 2: Sender's group-membership x 2: Feedback valence x 2: Feedback dimension).

3.6.1.4 | EEG analyses

3.6.1.4.1 | P200

P200 peaks were largest at fronto-central electrodes FCz, Cz, and CPz, consistent with prior studies (Herbert et al., 2006; León et al., 2010). The P200 was scored within 220–300 ms post-stimulus-onset. Peak-amplitude values were then submitted to a 3 (Electrode site) x 2 (Sender's group-membership) x 2 (Feedback valence) x 2 (Feedback dimension) RM ANOVA.

3.6.1.4.2 | P300

Consistent with prior studies (Polich, 2007; Schindler & Kissler, 2018; Schupp et al., 2004), P300-amplitudes were scored at centro-parietal electrodes Pz, CPz, and Cz. Mean amplitudes for the P300 were averaged between 300–400ms post stimulus-onset. Mean amplitudes were then submitted to a 3 (Electrode site) x 2 (Sender's group-membership) x 2 (Feedback valence) x 2 (Feedback dimension) RM ANOVA.

3.6.1.4.3 | LPP

Consistent with prior studies (Klein, Iffland, Schindler, Wabnitz, & Neuner, 2015; Schindler & Kissler, 2018; Schupp, Flaisch, Stockburger, Jungho, & Junghöfer, 2006), LPP-amplitudes were scored at fronto-central and centro-parietal midline electrodes Fz, FCz, Cz, CPz, and Pz. Mean amplitudes were averaged between 400–700ms post stimulus-onset. Mean amplitudes were then submitted to a 5 (Electrode site) x 2 (Sender's group-membership) x 2 (Feedback valence) x 2 (Feedback dimension) RM ANOVA.

In case of violation of sphericity, p-values and effect sizes corrected according to Greenhouse-Geisser are reported. For readability, the uncorrected degrees of freedom are reported. Post hoc comparison tests for interactions were set up using Bonferroni adjustment to adjust for multiple comparisons.

3.6.2 | Behavioral and self-report results

3.6.2.1 | Checks

We checked our within-participants manipulations of sender's group-membership and feedback dimension. As intended, paired-sample t-tests revealed that participants identified more with ingroup senders (i.e., Dutch students, $M = 3.97$, $SD = 1.32$) than outgroup senders (i.e., international students, $M = 2.86$, $SD = 1.10$), $t(48) = 5.47$, $p < .001$, $CI = [.70, 1.52]$. For the feedback dimension, we asked participants to indicate whether the feedback they received addressed their morality and their competence (separately, i.e., 1 = fully disagree, 7 = fully agree). As intended, a one-sample t-test against the midpoint of the two scales (i.e., 4) for our manipulation checks of feedback dimension confirmed that participants reported having received feedback on both their morality: $M = 5.69$, $SD = 1.05$, $t(48) = 11.35$, $p < .001$, $CI = [1.39, 1.99]$, and on their competence: $M = 5.73$, $SD = 1.13$, $t(48) = 10.72$, $p < .001$, $CI = [1.41, 2.06]$.

Results of the free social categorization task (without the background color cue for ingroup and outgroup faces) showed that participants were highly accurate in categorizing faces (93%) as belonging to the ingroup or outgroup. This indicates that participants focused on the faces, rather than just the background color, and that they were aware of the group-membership of

the senders. Additionally, we found that participants perceived ingroup and outgroup faces differently, as indicated by a significant difference in amplitudes (when viewing ingroup compared to outgroup faces) in ERPs associated with social categorization (i.e., N200, see Supplementary Materials for further details).

3.6.2.2 | Perceived valence of feedback from individuals

A one-sample t-test against the midpoint of the scale (i.e., 4, 1 = positive, 7 = negative⁷) showed that participants perceived the feedback as more negative than positive, $M = 4.33$, $SD = .75$, $t(48) = 3.05$, $p = .004$, $CI = [.11, .54]$, replicating the results of Study 2.1. A RM ANOVA (Sender's group-membership [within], Feedback dimension [within]) revealed no significant effects of dimension and sender's group-membership on perceived valence, $F_s \leq 3.61$, $p_s \geq .063$.

3.6.2.3 | Affective and defensive response to negative group-based feedback⁹

Also in line with Study 2.1, a RM MANOVA (Sender's group-membership [within]; negative, positive, and moral emotions as dependent variables; Bonferroni correction, $\alpha = .05/3$) showed that participants reported stronger negative emotions after receiving an overall negative feedback message from the ingroup ($M = 3.14$, $SD = 1.25$) than from the outgroup ($M = 2.59$, $SD = 1.11$), multivariate: Pillai's Trace = .43, $F(3, 46) = 11.55$, $p < .001$, $\eta_p^2 = .43$, univariate: $F(1, 48) = 35.05$, $p < .001$, $\eta_p^2 = .42$. As in Study 2.1, the other affective responses (i.e., positive and moral) did not show any univariate effects ($F_s \leq 1.85$, $p_s \geq .180$).

Finally, we also replicated the findings of Study 2.1 for defensiveness after receiving feedback, using a paired-sample t-test. Participants made more negative assumptions after receiving an overall negative feedback message from the ingroup ($M = 5.10$, $SD = .82$) than from the outgroup ($M = 4.95$, $SD = .95$), $t(48) = 2.07$, $p = .044$, $CI = [.004, .29]$.

⁹ Please note that feedback dimension was measured as a within-participants factor in this study, which is why we cannot investigate the effect on the dependent variable here.

3.6.3 | Event-related brain potentials (ERPs)

3.6.3.1 | P200

Participants paid more initial attention to feedback that was delivered by ingroup than by outgroup members. In line with our expectations that participants would attend more to feedback messages coming from the ingroup (vs. outgroup), there was a main effect of sender's group-membership on the P200 elicited by feedback messages. That is, amplitudes in response to the feedback messages were larger when an ingroup sender had delivered the feedback message ($M = 5.57$, $SE = .30$) compared to an outgroup sender ($M = 5.16$, $SE = .29$), $F(1, 40) = 7.31$, $p = .010$, $\eta_p^2 = .16$ (see Figure 3.3). Additionally, we found a three-way interaction effect between electrode, sender's group-membership, and valence, $F(2, 80) = 4.35$, $p = .027$, $\eta_p^2 = .10$. To break down this interaction, we examined effects for each electrode. The main effect of sender's group membership emerged on all three electrodes, FCz: $F(1, 40) = 4.36$, $p = .043$, $\eta_p^2 = .10$, Cz: $F(1, 40) = 8.95$, $p = .005$, $\eta_p^2 = .18$, and CPz: $F(1, 40) = 7.26$, $p = .010$, $\eta_p^2 = .15$. In addition, there was a two-way interaction between sender's group-membership and valence on FCz, $F(1, 40) = 4.64$, $p = .037$, $\eta_p^2 = .10$. When viewing negative feedback, the P200 was significantly larger when it was delivered by ingroup senders ($M = 6.18$, $SE = .34$) compared to outgroup senders ($M = 5.31$, $SE = .38$), $p = .006$. When viewing positive feedback, this difference was not significant, $p = .949$. When viewing feedback delivered by ingroup senders, the P200 was significantly larger for negative ($M = 6.18$, $SE = .34$) compared to positive feedback ($M = 5.54$, $SE = .39$), $p = .037$. This difference was not significant when viewing feedback from outgroup senders, $p = .337$. There was a main effect of electrode, with largest amplitudes at FCz ($M = 5.65$, $SE = .32$), $F(1, 40) = 23.06$, $p < .001$, $\eta_p^2 = .37$. There were no other effects, $F_s \leq 2.89$, $p_s \geq .079$.

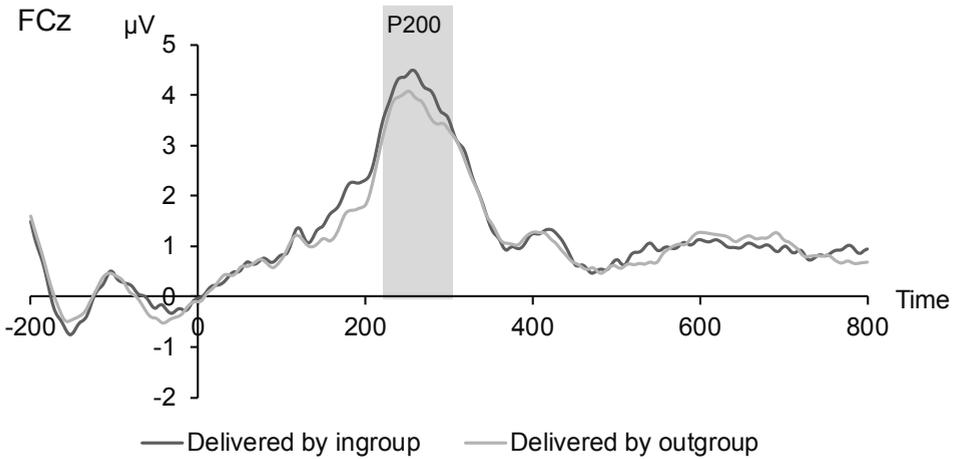


Figure 3.3. ERP-waveforms at fronto-central midline electrode FCz for feedback messages delivered by ingroup and outgroup senders.

3.6.3.2 | P300

Participants were more motivated to attend to feedback related to their competence rather than their morality. That is, P300-amplitudes in response to feedback messages were larger when the feedback concerned participant's competence ($M = 1.64$, $SE = .25$) compared to their morality ($M = 1.35$, $SE = .23$), $F(1, 40) = 5.00$, $p = .032$, $\eta_p^2 = .11$ (see Figure 3.4). Interestingly and similar to the results on the P200, this main effect was qualified by an interaction effect between feedback dimension and valence, $F(1, 40) = 4.62$, $p = .038$, $\eta_p^2 = .10$. When viewing negative feedback, the P300 was larger when feedback concerned participant's competence ($M = 1.85$, $SE = .23$) as compared to their morality ($M = 1.32$, $SE = .25$), $p = .003$. There was no main effect of dimension for positive feedback, $p = .817$. When viewing feedback concerning participant's competence, the P300 was significantly larger for negative feedback messages ($M = 1.85$, $SE = .23$) compared to positive messages ($M = 1.42$, $SE = .24$), $p = .006$, $\eta_p^2 = .18$. There was no main effect of valence for feedback on morality, $p = .698$. There was also a main effect of electrode, P300-amplitudes were largest at CPz ($M = 1.85$, $SE = .27$), $F(2, 80) = 35.39$, $p < .001$, $\eta_p^2 = .47$. There were no other effects, $F_s \leq 2.79$, p_s

$\geq .102$.

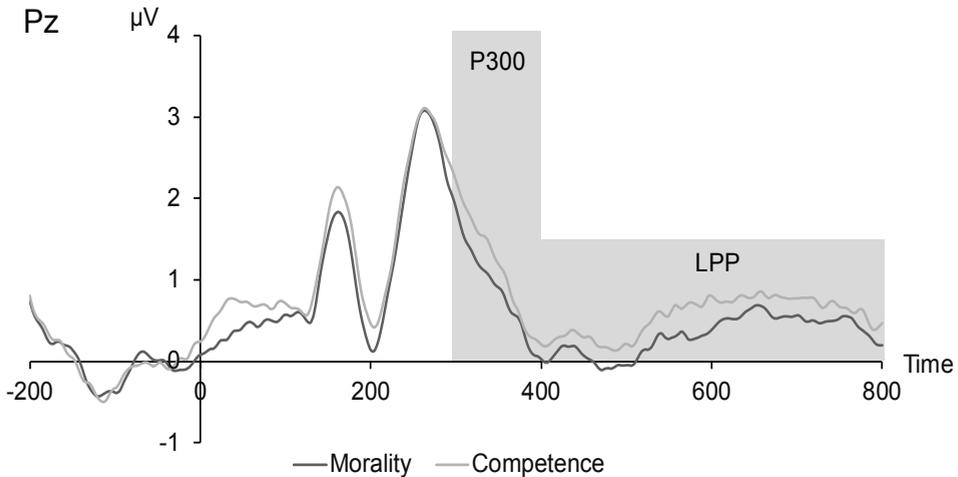


Figure 3.4. ERP-waveforms at parietal midline electrode Pz for feedback messages (i.e., character traits) on participant's morality and competence.

3.6.3.3 | LPP

Consistent with our reasoning and the findings for the P300, we found that participants were more motivated to attend to feedback related to competence rather than morality (see Figure 3.4): LPP-amplitudes were larger when feedback concerned participant's competence ($M = 1.01$, $SE = .23$) compared to their morality ($M = .72$, $SE = .22$), $F(1, 40) = 12.17$, $p = .001$, $\eta_p^2 = .23$. In line with the results on the P200 and P300, we checked whether the effect of feedback dimension was stronger for negative (vs. positive) feedback, even though there was no significant interaction. For negative feedback, the LPP was larger for feedback concerning participant's competence ($M = 1.15$, $SE = .24$) as compared to their morality ($M = .75$, $SE = .24$), $p = .026$. This difference was not present for positive feedback, $p = .275$. There were no effects of valence for feedback on morality and competence, $ps \geq .165$. There was a main effect of electrode with the largest amplitudes at Cz ($M = 1.10$, $SE = .25$), $F(4, 160) = 6.37$, $p = .005$, $\eta_p^2 = .14$. There were no other effects, $F_s \leq 1.78$, $ps \geq .190$.

3.6.3.4 | Validation ERP effects: Recall task

In line with our results for the P300 and LPP which suggested that participants paid less motivated attention to feedback on their morality compared to their competence, a RM ANOVA (Feedback dimension [within], Feedback valence [within]) revealed a main effect of dimension on recalled feedback messages, $F(1, 48) = 8.77, p = .005, \eta_p^2 = .15$ (see Figure 3.5). On average, participants recalled (or reported) more competence- ($M = 2.85, SD = 1.10$) compared to morality-related feedback messages ($M = 2.16, SD = 1.34$). The main effect of valence was not significant (positive [$M = 2.66, SE = 1.52$], negative [$M = 2.36, SE = 1.55$]), $F(1, 48) = 3.44, p = .070, \eta_p^2 = .07$. There was, however, a significant interaction between dimension of the feedback and valence, $F(1, 48) = 4.84, p = .033, \eta_p^2 = .09$. For negative feedback, messages related to competence were recalled significantly more often ($M = 2.88, SE = .19$) than messages related to morality ($M = 1.84, SE = .21$), $p < .001$. For positive feedback, there was no significant difference for dimension, $p = .283$. For feedback on participant's morality, positive feedback was recalled more often ($M = 2.50, SE = .22$) than negative feedback ($M = 1.84, SE = .21$), $p = .004$. For feedback on participant's competence, there was no difference for valence, $p = .801$. Thus, negative feedback messages on participants' morality were recalled the least.



Figure 3.5. Averages of recalled feedback messages on recall task per condition (Feedback dimension x feedback valence)

3.6.4 | Discussion

Prior neuroscientific research suggests that social context information, such as the threatening motive of a sender of a feedback message, can influence the cognitive processing of the sender's face and their message (Klein et al., 2015; Schindler et al., 2020; Schindler & Kissler, 2018; Schwarz et al., 2013; Wieser et al., 2014). The current research extends prior findings by examining whether social information, such as the sender's social group-membership (ingroup vs. outgroup, Tajfel & Turner, 1979) and the dimension addressed with a negative feedback message (morality vs. competence, Abele-Brehm et al., 2020) can modulate the cognitive processing of feedback messages in a similar vein. Our hypotheses were informed by social-psychological research suggesting that feedback from ingroup senders is more effective than from outgroup senders (Ellemers et al., 2013; Esposito et al., 2013; Hornsey et al., 2002, 2004; Pagliaro et al., 2011; Rösler et al., 2021a). Furthermore, they were informed by research suggesting that feedback on people's morality is less effective and more difficult to cope with than feedback on people's competence (Täuber & van Zomeren, 2013; Van der Lee et al., 2016).

Which is why in the case of negative moral feedback people often engage in defensive responses such as making negative assumptions about the sender (e.g., having no expertise) or misremembering moral failures (Carlson et al., 2020; Gausel & Leach, 2011; Rösler et al., 2021a).

Extending past research and as expected, participants in Study 2.1 perceived negative feedback on their morality as more negative than negative feedback on their competence. Complementing this finding, participants in Study 2.2 were less motivated to attend to feedback messages on their morality compared to their competence and were less likely to recall such information. Moreover, they recalled negative evaluations of their morality the least. Contrary to our expectations, participants responded more defensively to and reported to be more negatively affected by negative feedback from the ingroup (vs. outgroup), possibly suggesting that participants were more emotionally affected by such feedback. With Study 2.2, we replicate this group effect and show how this social-psychological phenomenon is reflected in participants' immediate brain responses. That is, feedback messages from ingroup (vs. outgroup) senders initially received more attentional resources.

3.6.4.1 | Social group-membership of a sender

Our ERP-findings demonstrate that the social group-membership of a feedback sender can modulate early brain responses when processing feedback messages. More specifically, we found that feedback messages coming from ingroup senders elicited larger P200-amplitudes compared to messages coming from outgroup senders. This extends previous ERP-research which showed that feedback messages coming from more vs. less relevant senders (e.g., from a human vs. a computer) are associated with increased P200 amplitudes (Schindler et al., 2015, 2020). Our effects were more pronounced for negative (vs. positive) feedback. That is, compared to positive feedback and feedback from outgroup senders, negative feedback from ingroup members received the most initial attentional resources (i.e., largest P200 amplitudes), indicating that participants are especially vigilant when receiving such feedback. This is in line with research showing that people are particularly motivated to form a positive impression in the eyes of fellow ingroup members (Pagliaro et al., 2011). Moreover, prior ERP-

research demonstrated that participants are especially motivated to perform well on a task when they are observed by a (minimal) ingroup vs. outgroup member (Pfabigan, Holzner, & Lamm, 2016; Van Nunspeet et al., 2015). Being more vigilant towards ingroup feedback is further supported by our self-report data. Whereas previous research showed that people respond more emotional and defensive to negative feedback from outgroup (vs. ingroup) members when reflecting on such feedback retrospectively (e.g., Rösler et al., 2021a), we find an opposite pattern when participants receive such feedback in the actual moment. Across two studies, using samples from the UK (Study 1) and the Netherlands (Study 2.2) and different social group-memberships (Study 2.1: Students/non-students, Study 2.2: Dutch students/international students), participants were more emotionally affected by and more defensive towards negative feedback from the ingroup compared to the outgroup. Even though this may initially seem contradictory to previous research, it is in line with the theoretical reasoning that feedback coming from the ingroup is less effective than feedback coming from the outgroup. To explain, the stronger emotional and defensive reaction to receiving negative feedback from the ingroup (vs. outgroup) *in the actual moment* (rather than reflecting on it at a later stage) may suggest that such negative feedback is perceived as more self-relevant (as also reflected in the EEG findings of Study 2.2).

A limitation of our research might be that we found an effect of sender's group-membership only for the P200 and not at later stages of stimulus evaluation (i.e., on the P300 and LPP), as initially predicted. This may, however, be explained by the fact that we used a more subtle manipulation of sender's group-membership compared to previous research. That is, instead of comparing feedback from humans vs. computers (Schindler & Kissler, 2016, 2018), we introduced the senders of the feedback as fellow national (ingroup) vs. international students (outgroup) and presented them by photographs of human faces —manipulating whether they belonged to the ingroup or outgroup by changing the background color. Recent research using similar social group-memberships, such as a sender being a therapist vs. a layperson, has been shown to also modulate only the early P200 (Schindler et al., 2020). Thus, strong manipulations of group-membership

might modulate ERPs such as the P300 and LPP, whereas more subtle and social group-memberships primarily seem to modulate earlier ERPs such as the P200.

3.6.4.2 | Valence of and feedback dimension addressed in feedback

In line with previous research (Carretié, Mercado, et al., 2001; Ito et al., 1998; Rozin & Royzman, 2001) we found evidence for a negativity bias for processing feedback messages. In both studies, participants perceived feedback as more negative than positive, after having received an equal amount of positive and negative feedback messages (and having tested stimulus information for equivalence, see Table 3.1). In Study 2.2, negative feedback elicited greater P200-amplitudes than positive feedback when delivered by ingroup members (as discussed above). P300-amplitudes were also larger for negative (vs. positive) feedback when feedback concerned participant's competence (vs. their morality).

As for the dimension addressed in a message, we found that, as predicted, feedback on participant's competence elicited greater P300- and LPP-amplitudes than feedback on their morality. The effect was more pronounced for negative (vs. positive) feedback. Complementing this finding, our self-report and behavioral results showed that participants perceived moral (vs. competence) feedback messages as more negative (Study 2.1), remembered morality (vs. competence) messages about themselves less often (Study 2.2), and remembered (or reported) negative feedback on their morality the least (Study 2.2). Our findings may imply that negative feedback on people's morality is disregarded because it puts people in a self-defensive state. This interpretation fits with research suggesting that giving people negative feedback on their morality (vs. on their competence) makes them respond defensively because they perceive this type of feedback as difficult to cope with (e.g., Gausel & Leach, 2011; Täuber & van Zomeren, 2013; Tetlock, 2002; Van der Lee et al., 2016).

3.6.4.3 | Conclusion

Whenever we receive feedback from someone in person, this situation often starts with recognizing the social group-membership of the person

and subsequently cognitively processing the content of their feedback. The current research demonstrates that a sender's social group-membership and the social dimension addressed with the feedback modulates how we process the feedback. When (negative) feedback came from the ingroup, participants showed enhanced cognitive processing of the feedback message and they were more emotionally affected by it compared to when it came from the outgroup. The feedback dimension also modulated the processing of the feedback: Receiving feedback on their morality (vs. competence), made participants judge the feedback as more negative. Additionally, they were less motivated to attend to and remembered fewer of such (negative) feedback messages. These results may offer new insights into why feedback from outgroup members and on one's morality is often ineffective. People may dismiss such feedback as early as when they attentionally process the message and they may hide receiving negative feedback on their morality in front of others to guard their social image.

The background is a teal watercolor wash with a white torn paper effect at the bottom. The text is centered in the upper half of the image.

CHAPTER

4

Show Me That You Mean Well: The Beneficial Effects of Communicating the Intention to Help when Delivering Moral Critique as an Outgroup Member

This chapter is based on:

Rösler, I.K., Van Nunspeet, F., Ellemers, N.. Show me that you mean well: The beneficial effects of communicating the intention to help when delivering moral critique as an outgroup member. *In prep.*

Author contributions:

IKR, FVN, and NE designed the studies. IKR performed the data analysis and wrote a first draft of the manuscript. FVN and NE provided substantial comments and suggestions for improvement.

Abstract

Negative feedback delivered by outgroup members and on people's morality is often not accepted. In two preregistered studies, we investigated whether people become more receptive to such feedback when feedback senders make their (good) intentions explicit. In an experimental paradigm, participants received negative feedback from ostensible others on their selfish (rather than altruistic) decisions in a donation task. We manipulated the identity of a feedback sender (ingroup vs. outgroup member) and the intention they provided for giving the feedback (i.e., 'other-enhancement' by communicating the intention to encourage the person criticized to improve vs. 'self-enhancement' by communicating the intention to demonstrate their own moral superiority vs. control [no intention communicated]). We measured participants' self-reported responses to the feedback (Study 3.1, $N = 50$), and additionally recorded an EEG in Study 3.2 ($N = 34$). Consistent with the literature, participants made more negative assumptions about the intentions of outgroup (vs. ingroup) feedback senders and perceived such feedback as less fair when intentions were not made explicit. However, we demonstrate that this group effect can be eliminated when outgroup members make their intentions explicit (Study 3.1). Additionally, we show that when outgroup members communicate that they provide negative feedback with the intention to help the participant (other-enhancement), compared to wanting to show their own moral superiority (self-enhancement) or when they do not communicate their intentions, participants perceive the negative feedback on their moral behavior as more fair. Complementing these self-report findings, ERP results of Study 3.2 seemed to suggest that communicating the intention to help reduced participants' attentional vigilance towards the negative feedback messages (i.e., decreased P200-amplitudes). This may suggest that communicating the intention to help can buffer against the negative effects of receiving such painful evaluations of the self.

“Now if arguments were in themselves enough to make men good, they would [...] have won very great rewards [...]; but as things are, [...] they are not able to encourage the many to nobility and goodness.” (EN X.9. 1179b4–10)

People often try to motivate others to act more morally by criticizing them. However, already around 350 years BC the ancient philosopher Aristotle called into question whether telling someone to act morally is a successful strategy to motivate them to become a good person. In line with this argument, recent social psychological research has demonstrated that when people aim to encourage others to show moral behavior by criticizing them, this often backfires (Gausel & Leach, 2011; Rösler et al., 2021a; Täuber et al., 2018; Täuber & van Zomeren, 2013). It often makes people defensive, angry, and annoyed. Part of this defensive reaction towards moral criticism might entail that the person being criticized makes assumptions about the underlying motives about why the criticism was given. This might happen, because even though critics often have a specific intention or goal in mind when delivering feedback, they often do not communicate their intentions to the person they are criticizing. In these cases, when the actual intentions are unknown, people might use the group-membership of the critic as a cue to infer whether a critic is trying to convey constructive feedback or aims to advance themselves (e.g., by downward comparison, Festinger, 1954; Hornsey et al., 2004; Tajfel & Turner, 1979; Wills, 1981). This is problematic, because when a critic is perceived as an outgroup member, people are inclined to think that any negative feedback is provided with hostile intentions and are therefore less willing to accept the feedback message (Esposito et al., 2013; Hornsey & Imani, 2004; Hornsey et al., 2004; Thürmer et al., 2019). More specifically, a feedback receiver may assume that an outgroup critic is motivated to advance their own self-image by condemning the person they are criticizing, rather than helping the receiver of the feedback by giving constructive feedback (i.e., engaging in social comparison, Festinger, 1954; Wills, 1981). Thus, being an outgroup member and criticizing others with the intention to promote behavioral change can be challenging.

This implies that outgroup members in particular (more so than ingroup members), face three obstacles when trying to encourage others to improve their behavior by criticizing them for their morality. First of all, people who receive such criticism might react defensively because it is their morality that is condemned, which constitutes a threat to one's sense of self (Gausel & Leach, 2011; Rösler et al., 2021a; Täuber et al., 2018; Täuber & van Zomeren, 2013). Such feedback threatens one's self-view of being 'a good' person and it may give rise to the fear of being ostracized for showing immoral behavior (e.g., Gausel & Leach, 2011; van der Lee et al., 2017). Secondly, people receiving such moral criticism may react defensively because the message is coming from an outgroup (rather than an ingroup) member, which causes them to judge the intentions of the critic as less constructive (Esposito et al., 2013; Hornsey & Imani, 2004; Hornsey et al., 2004; Thürmer et al., 2019). Thirdly, people might make specific assumptions about such hostile intentions, for instance, they may infer that someone criticizing them for their morality has the intention to increase their own self-image rather than wanting to help the other improve (e.g., Festinger, 1954; Wills, 1981). It is important to overcome these three obstacles, because feedback on people's moral behavior (also in the form of criticism) coming from outgroup members needs to be heard. Outgroup feedback is essential for preventing any type of group culture, where group members are afraid to speak up about morally questionable group behaviors in fear of being ostracized (e.g., Cavazza et al., 2014; Ellemers & van Nunspeet, 2020; Thau et al., 2015; Van der Lee et al., 2017). Without an outsider perspective, such questionable behaviors easily become normalized and encourage unethical group behavior in the future. Moreover, in the current political climate, where polarization and 'othering' increase (e.g., Finkel et al., 2020), it is important to be able to criticize each other on moral issues and values without causing more potential conflicts between groups.

In the current research, we therefore investigated a new intervention to help outgroup members getting their voice heard when they aim to criticize others for their morality: Explicitly communicating their helpful intentions when providing such feedback. To test whether this intervention is successful, we asked participants to reflect on their experiences of receiving feedback in an

experimental paradigm and, as a measure of their defensiveness, asked them to indicate how fair they perceived the negative feedback to be. However, since defensive responses to feedback are difficult to capture with self-report measures, we additionally examined whether and how participants' brain activity was affected when receiving feedback. To this end, we recorded an electroencephalography (EEG) while participants took part in an experiment and measured their brain responses while they received feedback on their morality (i.e., in the actual moment). Using such methods, we show how communicating helpful intentions can help outgroup feedback senders with the three reasons why receivers of their feedback may respond defensively.

4.1 | Why do people get defensive when their morality is criticized?

The first problem that (outgroup) feedback senders face is related to what receiving a negative evaluation of the self entails for the person who is being criticized. That is, being criticized for our morality means that a central part of our identity (i.e., seeing ourselves and being seen by others as being a moral person) is under attack (Aquino & Reed, 2002; Brambilla & Leach, 2014; Pagliaro et al., 2016; Strohminger & Nichols, 2014). To demonstrate, Van Nunspeet, Ellemers, Derks, and Nieuwenhuis (2014) asked participants to perform a task that was either framed as being indicative of their morality or their competence. Results showed that participants were more inclined to perform well on the task and showed enhanced response-monitoring (measured with EEG), in the morality compared to the competence condition. These findings seem to suggest that people are more intrinsically motivated to show that they are moral (vs. competent). Moreover, other research has demonstrated that when being criticized for their morality, people show defensive mechanisms such as making negative assumptions about the intentions of a feedback sender or getting angry (Gausel & Leach, 2011; Rösler et al., 2021a; Täuber & van Zomeren, 2013; Van der Lee et al., 2016). To still maintain a positive self-image, people therefore engage in coping mechanisms such as justifying and misremembering their moral shortcomings or emphasizing having changed since the moral failure (Carlson et al., 2020; Kouchaki & Gino, 2016; Mazar et al., 2008; Shalvi et al., 2015; Stanley et al., 2019). In other words, feedback on people's morality is often not very effective because people deceive themselves about their

own (im)moral behavior. To increase this effectiveness, prior research has contrasted negative feedback on one's morality with negative feedback on one's competence (Rösler et al., 2021a; Täuber & van Zomeren, 2013; van der Lee et al., 2016). Addressing someone's competence as the focal concern addressed in a feedback message, rather than someone's morality, is able to increase effectiveness through decreasing defensiveness and raising motivation to improve based on the feedback (Rösler et al., 2021a; Täuber & van Zomeren, 2013). For example, an employee might be lying about having performed a task, while not having done so. Criticizing this action might evoke less defensive responses if the competence failure (i.e., not performing the task) is the main focus of the message, rather than the morality failure (i.e., the lying). However, framing criticism as being about someone's competence might not always be desirable. For example, a critic might also want to stress the ethical implications of immoral behavior or enforce a specific moral norm (e.g., no fraudulent behavior on the financial market). In the current research, we therefore tested an intervention that can be applied to many more situations: Communicating the intention to help with feedback (discussed in more detail below).

The second obstacle outgroup critics face when trying to deliver feedback on people's morality is related to them being perceived as belonging to a different social group by the person they criticize.

4.2 | Why criticism from outgroup members is often not effective

In 1969, building on Aristotle's work, the rhetoric expert Kenneth Burke argued for the importance of social contexts in communication and persuasion (Burke, 1969). To influence others with speech, speakers should persuade the audience to identify with the speakers' interests. The problem with this rhetorical strategy is, however, that in many cases, people do not (want to) identify with (the interests of) an interaction partner. For example, because the speaker belongs to a different social group (Ellemers, 2012; Tajfel & Turner, 1979). Social-psychological research has shown that negative feedback from outgroups is not accepted (intergroup sensitivity effect, Hornsey & Imani, 2004; Hornsey, Oppes, & Svensson, 2002). People might infer that feedback from outgroup critics is driven by hostile intentions

and that it is less constructive than when it is coming from critics who they perceive as members of their own (in)group (Hornsey & Imani, 2004). The issue with making such negative assumptions about outgroup feedback senders, that is, perceiving feedback as too harsh or not constructive and (thus) not accepting it, is that it can rob one of the opportunity to grow and improve (e.g., Hornsey & Esposito, 2009). Moreover, it can potentially lead to (intergroup) conflict and increase polarization between groups. In the current research, we thus investigated whether we can intervene in the process of people making negative assumptions about the intentions of outgroup feedback senders by having feedback senders communicate their intentions alongside their feedback messages.

4.3 | Why people may infer that feedback senders want to enhance themselves

The third problem (outgroup) feedback senders face is related to the annoyance people experience when they see others showing off their moral behavior (Cramwinckel et al., 2013; Minson & Monin, 2011; Monin et al., 2008). To understand why people get annoyed and how this is related to the ineffectiveness of outgroup feedback, we first need to look at research on social comparison processes. According to the social comparison theory (Festinger, 1954; Wills, 1981), people compare themselves to others to either increase their well-being or to self-improve. Increasing well-being can be achieved by comparing oneself to others who are (performing) worse than the self (i.e., downward comparison, Wills, 1981). To self-improve, people compare themselves to others who are (performing) better than the self (i.e., upward comparison, Wills, 1981). Although both upward and downward comparison can benefit the person engaging in it, downward comparison is often found to annoy the people who are targeted or witness such behavior (Minson & Monin, 2011; Monin et al., 2008). For example, in two studies where non-vegetarians were asked to judge vegetarians, vegetarians were judged more negatively when non-vegetarians assumed that vegetarians wanted to show that they are morally superior (Minson & Monin, 2011). In the current research, we applied these comparison processes to feedback situations. Negative feedback on one's morality might make receivers of negative feedback assume that the feedback sender wants to engage in

downward comparison and show moral superiority. This would then likely make the feedback receiver annoyed and have them judge the feedback sender and message as more negative, for example as not being fair. However, if the feedback sender would communicate that their feedback is intended to help the receiver, this might increase the receipt of the feedback, making receivers perceive the feedback as more fair.

4.4 | Current research

In the current research, we aimed to find a way to make criticism from outgroup members on people's morality more effective. To this end, we investigated whether people become more receptive to negative feedback when the feedback senders communicate their intention to help by providing the feedback (see Figure 4.1). For our predictions, we draw on the insights from three different streams of research. Firstly, we draw on research on morality showing that people dislike others who show off their moral superiority and that they get defensive when their own morality is criticized (Monin et al., 2008; Rösler et al., 2021a; Täuber & van Zomeren, 2013). Secondly, we draw on research on the intergroup sensitivity effect showing that outgroup (vs. ingroup) criticism is met with defensiveness (Hornsey & Imani, 2004; Hornsey et al., 2002). And thirdly, we draw on research on social comparison processes (Festinger, 1954; Wills, 1981) that might play a role in the receiver's assumptions about the intentions of feedback senders, such as the feedback sender wanting to demonstrate moral superiority. The common denominator between these research lines is that people make assumptions about the feedback sender's intentions of giving feedback and because of that get defensive and disregard the feedback. Our intervention, communicating intentions of why feedback was given, was aimed at reducing defensiveness towards feedback from the outgroup. We tested two main predictions¹⁰: 1) Whether communicating intentions can prevent people from

¹⁰ Based on a (unexpected) finding of a pilot study, we additionally specified a hypothesis in our preregistration about how communicated intentions might affect actual (task-related) behavior. However, we did not find evidence for this in the current studies and therefore did not include it in the paper. Moreover, the preregistration included two studies, each testing a slightly different dependent variable (i.e., perceived fairness of the feedback message and taking the feedback to heart) as we aimed to investigate which of these two are more relevant for our current research. Although we found similar results for both studies we chose perceived fairness as the more relevant dependent variable.

responding defensively when receiving negative feedback from outgroup (vs. ingroup) members, and 2) whether communicating the intention to help (i.e., other-enhancing intention), rather than the intention to demonstrate moral superiority (i.e., self-enhancing intention) or communicating no intentions, can make people perceive negative feedback on their morality as more fair.

We tested these predictions with two studies, using the same experimental paradigm and very similar research designs. In both studies, we asked participants to perform a donation task where they made repeated monetary allocation decisions. Since the main focus of the current research was to investigate responses to negative feedback on people's morality, we examined decisions where people made selfish (rather than altruistic) decisions in the donation task. After each decision, they received feedback in the form of trait evaluations of their character (paradigm adapted from Schindler & Kissler, 2018), from ostensible senders who were presented as ingroup or as outgroup members judging participants' behavioral decisions on the task. Most importantly, the feedback senders' intentions when delivering feedback (i.e., other-enhancement vs. self-enhancement) were also presented to participants. Study 3.1 was conducted online. Study 3.2 was conducted in the lab, and we added EEG recordings to our experiment to investigate how communicating intentions affects the cognitive processing of feedback messages (also see Introduction Study 3.2).

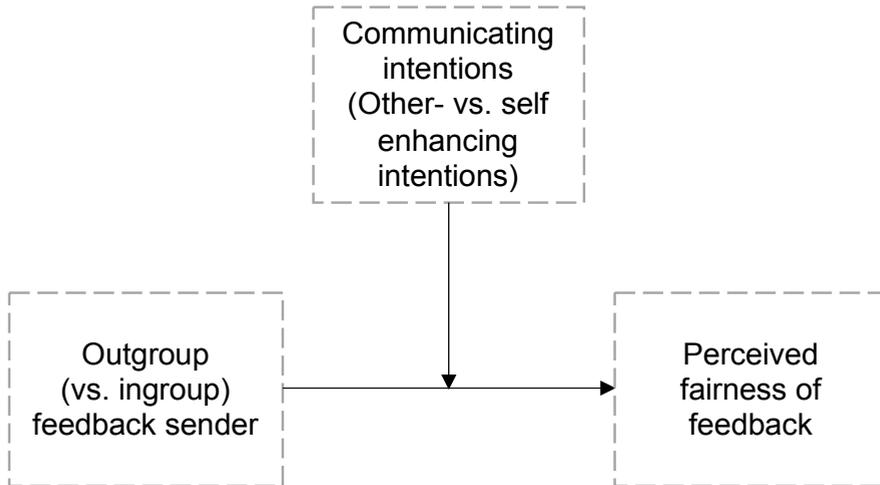


Figure 4.1. Overview of intervention of communicating intentions (other- vs. self-enhancing intention)

4.5 | Study 3.1

4.5.1 | Method

4.5.1.1 | Design and participants

The experiment had the following design: 2: Sender's group-membership (in- vs. outgroup, within-participant factor) x 2: Communicated intentions (other- vs. self-enhancing intention, within-participant factor) x 3: Timepoint of communicated intention (before the sender is presented vs. before the feedback message is presented vs. control [no intentions communicated], between-participants factor) factorial design. We included the factor 'timepoint of communicated intention' to check if it matters whether intentions are made explicit before or after the feedback sender is presented. The factor 'communicated intentions' was only manipulated within the two conditions that presented intentions to participants and not in the control condition.

Trials were presented to participants in a previous to the experiment randomized order. The valence of the feedback was dependent on participants' allocation decisions but preprogrammed (for details see Procedure, instruments, and stimuli).

Hypotheses, exclusion criteria for participants, outliers, and sample size were preregistered before data collection at aspredicted.org and can be found at <https://osf.io/vhjx7/>.

Since we examined whether communicating other-enhancing (vs. self-enhancing) intentions increases the effectiveness of negative feedback on people's morality, we used an estimator of prior research that tested another strategy to increase the effectiveness of feedback to determine a relevant effect size for our sample size calculation. In this past research, where the effectiveness of feedback on one's morality was compared to feedback on one's competence (Rösler et al., 2021a), participants reflected on autobiographical situations in which they had received negative feedback from either an ingroup or outgroup member, similar to our current design. We based sample size calculations on the effect size of the social dimension addressed in the negative feedback message (i.e., morality vs. competence), $\eta^2_p = .06$, on the acceptance of feedback. We used sample size calculation for repeated measure designs as a proxy for linear mixed models. Sample size calculation in G*power for the main effect of our main interest within-participants factor (i.e., communicated intentions: Other- vs. self-enhancing intention) in a repeated measure ANOVA ($\alpha = .05$, $1-\beta = .80$, $\eta^2_p = .08$, 10 repeated measures; see Supplementary Materials on how this was calculated) suggested testing 33 participants. We added 17 more participants to this sample size for several reasons. Firstly, we anticipated that some participants would indicate that it wasn't important to them to give money to people in need (i.e., requirement to make our experimental paradigm credible, see Instruments). Secondly, we wanted to have enough power for testing the effects of the other two manipulations of our experimental design (i.e., sender's group-membership and time of communicating intentions).

Dutch participants took part in the study on the platform Prolific ($N = 50$). Six participants were excluded because they did not meet the inclusion requirement for our experimental paradigm to find it important to give money to people in need (i.e., a score below 4 on the scale of our measure, see Instruments). The remaining sample consisted of 44 participants ($M_{age} = 26.55$, $SD = 8.17$, 17 females, 1 other): 13 participants were in the condition

where intentions were presented directly before the sender was presented, 16 participants in the condition where intentions were presented after the sender was presented and thus directly before the feedback message, and 15 participants in the control condition where no intentions were presented.

4.5.1.2 | Procedure, instruments, and stimuli

The current research aimed to test whether our intervention to communicate intentions when delivering feedback can decrease defensive responses towards feedback from outgroup members and given on one's morality. To test this intervention, we used an adapted experimental paradigm from previous research (Rösler, Van Nunspeet, et al., 2021b, p. 223; Schindler & Kissler, 2018). As in prior research (Rösler, Van Nunspeet, et al., 2021b, p. 223), we explained to participants that the senders and feedback messages originated from a previous experiment. In this previous experiment, participants were presented with behavioral decisions of targets and asked to judge (and provide feedback on) the character of these targets. We further explained that, in the current experiment, participants would receive the judgments that were given to targets who made similar decisions in the prior experiment as they did. We emphasized that the feedback messages would therefore be representative of how feedback senders from the previous experiment would judge their own character. We further explained, for the two conditions where intentions were communicated, that these previous participants were asked to select (pre-described) intentions of why they gave this feedback and that these intentions were presented in the current task as well.

The experiment lasted 30 minutes and was approved by the local Ethics Review Board. Participants gave informed consent and were fully debriefed and thanked at the end of the experiment. The experimental task was presented using the web-based experiment builder Gorilla (<https://gorilla.sc>). There were 20 trials per condition for the two within-participants factors of our experimental design (i.e., 2: Sender's group-membership x 2: Communicating intentions). Therefore, participants were presented with 80 trials. However, the exact number of trials per condition was dependent on participants' allocation decisions.

For the communicated intentions manipulation to be credible, we first needed to make sure that people saw it as one of their values to donate money to help people in need. Therefore, at the beginning of the experiment, we measured this by asking participants whether they agreed with the following statement: I find it important to give some money to people in need (7-point Likert scale, 1 = fully disagree, 7 = fully agree).

In the following, we first give a short overview of a trial before we go into detail about each part of our experimental task. Each trial started with the presentation of the charity and participants were asked to allocate a hypothetical amount of money between themselves and the charity (see Figure 4.2). This was followed by a picture of a sender's face with the background color indicating ingroup or outgroup membership (2000 ms). Then, the manipulation check *inferred good intentions* was measured. After that, the character trait indicating the moral feedback was shown (2000 ms). Finally, we measured the dependent variable of how *fair participants perceived the feedback* to be. Depending on the condition, a sentence communicating the sender's intention with their feedback was presented in the trial at different time points (i.e., before the sender was presented or before the feedback was presented, for 4000 ms), or no intention sentence was presented.



Figure 4.2. Example of trial

4.5.1.2.1 | Allocation decisions and feedback

For the allocation decisions, participants were asked (on each trial) to divide a hypothetical amount of €20 between themselves and a charity¹¹. For this task, we listed 80 Dutch charitable organizations related to different causes, such as a green environment, medical research, or animal rights. We presented their names together with the logo and a short explanation of the charitable cause. Participants could choose percentage amounts presented on a 10-point scale with a 10-point distance between percentages (e.g., 10%, 20%, 30%). Each choice implied both the amount for the self and the charity. To demonstrate, if a participant would choose to allocate 80% (i.e., €16) to the charity, 20% (i.e., €4) would go on their experiment account. We did not give participants the option to equally divide the money between themselves

¹¹ Analyses of the money allocation task showed that for Study 2.1 participants allocated 51% of all possible donations (i.e., contributing the full amount in each trial) and for Study 2.2 54%.

and the charity¹².

For the feedback, we used 32 Dutch moral traits (16 positive adjectives, 16 negative adjectives) from previous research (Rösler, Van Nunspeet, et al., 2021a, see Supplementary Materials). The stimuli had been pretested on and did not vary between their familiarity, length, frequency (per million), and the level of arousal associated with them. The valence of feedback messages was preprogrammed but depended on participants' allocation choices. That is, when participants allocated most of the money to the charity at the beginning of the trial (> 70% of the €20), we coded this decision as 'prosocial', and participants would receive a positive moral feedback message (e.g., cooperative). When participants did not allocate most of the money to the charity (< 70% of the €20), we coded these selfish decisions as 'proself' and participants would receive a negative moral feedback message (e.g., uncooperative). Since the focus of our research was on negative feedback, we were only interested in 'proself' decisions and thus aimed to increase the number of trials where participants made these 'proself' decisions. That is why we added an incentive, a €20 voucher, for the person who made the most 'proself' decisions and who thus ended up with the highest amount on their own experiment account.

4.5.1.2.2 | Manipulation of the sender's group-membership

To manipulate the senders' group-membership, we first measured participants' own social value orientation (SVO, Messick & McClintock, 1968; Van Lange, 1999) with the SVO slider measure (Murphy et al., 2011). We then explained to participants that the senders of the feedback messages would either belong to the ingroup, meaning the sender has a similar SVO-type as the participant, or to the outgroup, meaning the sender has a different SVO-type than the participant (see Supplementary Materials for SVO types of participants). Participants did not receive feedback on their own SVO-type, and we did not specify the SVO-type of outgroup members. To check this manipulation, we measured identification with people having a similar/different SVO-type with three items (e.g., "I identify with other members of this group", Ellemers,

¹²Due to a coding error participants received a little less negative feedback in one of our conditions and instead received positive feedback. Excluding these cases for all three conditions yielded the same results.

Kortekaas, & Ouwerkerk, 1999, ingroup, $\alpha = .74$, outgroup, $\alpha = .74$), directly after introducing the group-memberships to participants.

To represent ingroup and outgroup members in the experimental task, we selected 20 Caucasian (faced-forward) faces (10 female, 10 male) from the Radboud face database (Langner et al., 2010). We changed the background color of half of the faces to blue, to indicate ingroup membership of the faces, and the other half to yellow, to indicate outgroup membership of the faces. Previous research has shown that participants were able to learn the group-memberships and did not need to rely on the background color in a very similar experimental design (Rösler et al, 2021b, p. 223). As intended, an ANOVA revealed that the attractiveness of faces did not vary between gender and assigned social group-membership and that there was no interaction between the two, $F_s \leq 2.22$, $p_s \geq .156$.

After we introduced the group-memberships, participants were asked to perform a social categorization task to learn the group-membership of the senders. Participants had to correctly categorize each sender's face as belonging to the ingroup or outgroup twice to ensure that their correct responses were robust.

4.5.1.2.3 | Manipulation of communicating intentions

The focus of our research was on negative feedback. However, the valence of the feedback depended on participants' allocation decisions and participants were presented with both negative and positive feedback messages to increase the credibility of feedback. Since (in two of the conditions) we included senders' intentions for providing their feedback, and participants received both positive and negative feedback, the intentions also pertained to both the positive and negative feedback messages. That is, for positive feedback after prosocial decisions, intentions were related to acknowledging that the participant had contributed most of the money to the charity (e.g., "It seems like it is important for you to donate to charities"). More importantly, for negative feedback after proself decisions, these intentions were related to either other-enhancing intentions (e.g., "I want to help you by reminding you that you said it is important for you to donate to charities") or self-enhancing

intentions (showing moral superiority, e.g., “I seem to be a better person than you are”). We pretested intention sentences for their representativeness for feedback senders’ intentions of why feedback was given among a Dutch sample ($N = 20$, $M_{\text{age}} = 31.55$ years, $SD = 9.52$, 8 females, on Prolific) and selected the final set based on this pretest (see Table 4.1). To check whether these manipulations were successful, we measured *inferred good intentions* about senders with one item in each trial, directly after the sender was presented (“Do you think this person had good intentions when (s)he evaluated you based on your decision?”, 7-point scale, 1 = not at all, 7 = very much).

Table 4.1: Means, Standard deviations, and F-statistics for sentence stimuli per intention condition

	Other- enhancing sentences ($N = 5$)	Self- enhancing sentences ($N = 5$)	Acknow- ledgment sentences ($N = 10$)	$F(1, 20)$
Other-enhancing intention	5.15 (.19)	2.01 (.25)	4.30 (.26)	225.77***
Self-enhancing intention	3.98 (.19)	6.64 (.10)	2.22 (.32)	493.22***
Acknowledgment intention	2.98 (.12)	2.71 (1.02)	5.73 (.54)	52.14***

Note: Each sentence was rated by 20 participants, dependent variable is whether a sentence is ‘representative for intention’ measured with 7-point Likert scale (1 = not at all, 7 = very much), Acknowledgment intentions were not relevant to the current research as they were shown after positive feedback, *** $p < .001$

4.5.1.2.4 | Dependent variable: Perceived fairness of negative moral feedback
After the feedback message was shown, we measured the perceived fairness of this message with one item (“Do you think this evaluation of your character, based on your decision, was fair?”, 7-point scale, 1 = not at all, 7 = very much)¹³.

¹³ A one-item measure might seem like a non-reliable measure. However, single-item measures have been argued to work similarly well as more elaborate scales when measuring straight-forward constructs (e.g., such as job satisfaction, Wanous et al., 1997). Moreover, we achieve higher reliability through repeated trials (around 20 trials per condition).

4.5.1.2.5 | Data analyses

To test our hypotheses, we fitted linear mixed models (LMM)¹⁴ to our data and used maximum-likelihood (ML) estimation to be able to compare different models. For these analyses, we used the `lme` function of the `nlme` package (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2020) and the `lmer` function of the `lme4` package (Bates et al., 2015) in R (version 4.0.0). Model assumptions were checked using the `model_check` function of the `performance` package (Lüdtke et al., 2020). We added factors and interactions between factors hierarchically, comparing different models using the Akaike-Information-Criterion (AIC) and Bayesian-Information-Criterion (BIC). For clarity, we only report the final model.

4.5.2 | Results

4.5.2.1 | Checks

4.5.2.1.1 | Ingroup/outgroup identification

As intended, a paired sample t-test revealed that participants identified more with ingroup senders ($M = 4.63$, $SD = 1.88$) than with outgroup senders ($M = 2.85$, $SD = .93$), $t(43) = 9.65$, $p < .001$, 95% CI = [1.41, 2.15].

4.5.2.1.2 | Inferred good intentions of the sender

To check whether we conceptually replicate previous research showing that outgroup feedback is met with more defensiveness than ingroup feedback (Hornsey & Imani, 2004; Hornsey et al., 2004), we selected negative feedback trials from our control condition, where no intentions were communicated. According to this previous research, participants should make more negative assumptions about the intentions of outgroup rather than ingroup senders.

¹⁴We accounted for repeated measures, that is participants making several allocation decisions in the paradigm, by adding the participant number as a random intercept effect to our models. Using LMMs allows to model variance that occurs between individuals in repeated trials. For example, variance stemming from the individual differences such as the SVO-type that makes some participants always donate a lot of money, whereas others only donate very little. In general, modeling this type of variance, rather than putting it in the error term, can result in higher power compared to other models (e.g., Baayen, Davidson, & Bates, 2008). However, the most important benefit and reason for why we chose this analysis over repeated-measure ANOVA in the current research, is that LMMs can account for missing data (Baayen et al., 2008; Tibon & Levy, 2015). Our design was unbalanced, meaning that the number of trials per condition depended on participants' responses. Using LMMs, we do not have to exclude datapoints or participants as when using repeated-measure ANOVA.

We fitted LMMs (ML) predicting inferred good intentions specifying the sender's group-membership (ingroup vs. outgroup) as a fixed effect and the participant number as a random effect. This analysis revealed that, as expected, participants inferred worse intentions from outgroup (vs. ingroup) senders before receiving feedback from them, $B = -.57$, $t = -12.00$, $p < .001$, 95% CI = [-.66, -.48].

To check whether our manipulation of communicated intentions worked, we excluded participants from the control condition in which no intentions were communicated and selected trials in which participants had made a proself decision and thus received negative feedback. We then checked whether, in these types of trials, communicating other- (vs. self-)enhancing intentions affected the inferred good intentions of the sender. If successful, our manipulation should induce better inferred intentions on trials where other- (vs. self-)enhancing intentions were communicated. That was indeed the case: An LMM (ML) specifying participant number as a random effect, communicating intentions (other- vs. self-enhancing intention) as a fixed effect, and inferred good intentions as dependent variable revealed that communicated intentions predicted inferred good intentions, $B = -1.87$, $t = -16.85$, $p < .001$, 95% CI = [-2.09, -1.65]. In this selection of data, participants inferred that senders had better intentions when other-enhancing, rather than self-enhancing, intentions were communicated. Thus, our manipulation was successful.

4.5.2.1.3 | Comparing negative to positive feedback messages

To get a first insight into how participants scored on perceived fairness of moral feedback messages in all trials (i.e., both negative and positive feedback), we averaged the scores across trials. We then computed one-sample t-tests against the midpoint of the two scales (i.e., 4). We checked our anticipated effect of whether negative feedback was perceived as less fair than positive feedback. Indeed, averaged scores of perceived fairness differed significantly from the midpoint of the scale for negative feedback, $M = 1.96$, $SD = .89$, $t(42) = -14.96$, $p < .001$, 95% CI = [1.69, 2.24], indicating that participants generally perceived negative feedback on their morality as not very fair. For positive feedback, they did not differ from the midpoint, M

= 4.33, $SD = 1.26$, $t(40) = 1.69$, $p = .099$, 95% CI = [3.93, 4.73], indicating that participants perceived positive feedback as not especially fair or unfair. Comparing these two averages with an independent t-test revealed that, as anticipated, negative feedback was perceived as less fair than positive feedback $t(72) = -9.89$, $p < .001$, 95% CI = [-2.85, -1.89].

4.5.2.2 | Perceived fairness of negative moral feedback messages

Our two main predictions related to the perceived fairness of the feedback were that 1) Communicating intentions can decrease defensiveness towards negative feedback coming from outgroup (vs. ingroup) members and that 2) communicating the intention to help (i.e., other-enhancing intention), rather than the intention to demonstrate moral superiority (i.e., self-enhancing intention) can make people perceive negative feedback on their proself choices as more fair. To test these predictions, we performed two analysis steps. We first checked whether people perceive negative feedback on their proself decisions as less fair when this feedback was delivered by outgroup (vs. ingroup) senders in the control condition, where no intentions were communicated. Then, in a second step, we tested whether this group effect was eliminated in the two conditions where intentions were communicated and whether communicating other-(vs. self-)enhancing intentions made participants perceive negative feedback on their proself decisions as more fair. We used this two-step analysis approach, rather than performing one analyses including all factors, because the factor of communicated intentions was only manipulated in two of the three between-participants conditions.

For the first step, we selected negative feedback trials from our control condition, where no intentions were communicated. We fitted LMMs (ML) predicting perceived fairness of feedback specifying the sender's group-membership (ingroup vs. outgroup) as a fixed effect and the participant number as a random effect. As expected, participants perceived negative feedback messages on their morality as less fair when these were delivered by outgroup members compared to when they were delivered by ingroup members, $B = -.16$, $t = -2.43$, $p = .015$, 95% CI = [-.30, -.03].

For the second step, we only used negative feedback trials and the two

between conditions where intentions were (vs. were not) communicated. We fitted an LMM (ML) with perceived fairness as the dependent variable. We added participant number as a random effect and communicating intentions (other- vs. self-enhancing intention, within-participants), time of communicating intentions (before sender was presented vs. before message was presented, between-participants), and sender's group-membership (ingroup vs. outgroup, within-participants) as fixed effects. As predicted, the sender's group membership did not show any significant effects, suggesting that when intentions are communicated, the group that was found in the control condition does not occur. Also as predicted, communicating other- (vs. self-)enhancing intentions increased perceived fairness of negative moral feedback messages, $B = -.24$, $t = -3.99$, $p < .001$, 95% CI = $[-.37, -.12]$.

The time-point of when the intention was communicated in the trial did not influence any of our variables and is not further discussed here.

4.5.3 | Discussion

The self-report results of Study 3.1 demonstrate that our intervention can successfully increase the receipt of feedback on peoples' morality that is delivered by outgroup members. Firstly, in our control condition, where no intentions were communicated, people perceived negative moral feedback as less fair when it was delivered by outgroup (vs. ingroup) members. However, in the two conditions where intentions were communicated, this effect was eliminated. Secondly, communicating the intention to help when providing the feedback (rather than demonstrating moral superiority) showed to enhance the perceived fairness of negative feedback on one's morality (regardless of group-membership of the sender). Thus, we can conclude that communicating the intention to help when providing feedback can help outgroup (but also ingroup) critics get their voices heard when they criticize other people's morality.

However, only investigating self-report answers might not reveal how participants truly experienced the feedback and whether they actually pay attention to it (e.g., Ellemers & van Nunspeet, 2020). To explain, previous research has shown that people pay less attention to feedback messages

that are delivered by outgroup (vs. ingroup) members (Rösler, Van Nunspeet, et al., 2021b, p. 223). If we want to motivate people to change their behavior based on the feedback they receive, it is necessary that people attend to the feedback messages. To test whether our intervention also increased the attention paid to the feedback messages—and not only the (self-reported) perceived fairness—we ran a second study and measured participants' brain activity in response to such feedback by recording an EEG.

4.6 | Study 3.2

In Study 3.2, participants were asked to perform the same donation task as in Study 3.1. To test whether our intervention also affected the cognitive processing of negative feedback messages, we measured participants' immediate brain responses to receiving feedback by recording an EEG. From this continuous EEG, we extracted event-related potentials (ERPs) to examine the cognitive processes associated with receiving the feedback. We relied on these neuroscientific markers as they can give insight into whether participants paid initial and sustained attention when cognitively processing feedback messages (i.e., P200, P300/ LPP, Herbert et al., 2006; Lang et al., 1997; Lang & Bradley, 2013; León et al., 2010).

4.6.1 | P200¹⁵

Initial attention to emotional words such as feedback messages can be examined by investigating P200-amplitudes that are elicited by the semantic processing of words (Kissler et al., 2006). Modulations of P200-amplitudes have been associated with the initial lexical encoding of words (Kissler, Assadollahi, & Herbert, 2006; Trauer, Andersen, Kotz, & Müller, 2011). For example, in an experiment where participants were presented with task-irrelevant, negative (vs. neutral) words showed enhanced P200-amplitudes, suggesting that the valence and 'emotionality' of a word can modulate the semantic encoding of it (Trauer et al., 2011). Consequently, decreased

¹⁵ This ERP analysis was not pre-registered. However, visually inspecting the data made us decide to additionally examine the P200-amplitudes. Moreover, we had predicted that people would show enhanced social categorization of outgroup (vs. ingroup) faces as indicated by increased N200 amplitudes when viewing faces (as in previous research, Rösler et al., 2021b, p. 223). However, we didn't find any (reliable) effects on social categorization (see Supplementary Materials).

‘emotionality’ of a word may lead to lower P200-amplitudes.

4.6.2 | P300/LPP

Sustained attention can be measured with ERPs associated with higher-level cognitive processing and stimulus evaluation: The P300 (Lang et al., 1997; Polich, 2007) and Late Positive Potential (LPP, Schupp et al., 2004; Schupp, Flaisch, Stockburger, & Junghöfer, 2006), with higher amplitudes of these ERPs indicating increased and sustained attention. According to the framework of ‘motivated attention’, these two ERPs are indicative of associated arousal with a stimulus and of how motivated people are to pay sustained attention to these stimuli (Lang, Bradley, & Cuthbert, 1997; Lang & Bradley, 2013). These assumptions are based on the finding that people show increased P300- and LPP-amplitudes to highly arousing (positive and negative) compared to less arousing or neutral stimuli (Lang et al., 1997; Lang & Bradley, 2013; Schupp et al., 2004). The specific way in which attention is directed towards a stimulus, however, also depends on the context. Both appetitive (e.g., approach) and defensive (e.g., avoidance) mechanisms might modulate whether participants are motivated to pay sustained attention to stimuli (Lang, Bradley, & Cuthbert, 1997; Lang & Bradley, 2013). Research on feedback processing that has examined such context effects showed the importance of taking the group-membership of a feedback sender into account when investigating attention paid to processing feedback messages (Schindler et al., 2020; Schindler & Kissler, 2018). This research showed that participants attended more to feedback messages delivered by a more self-relevant sender (i.e., a human) compared to a less relevant sender (i.e., a computer), as indicated by larger P200, P300, and LPP amplitudes. In other ERP research, participants were presented with pictures of people’s faces, after they had been informed about the behavioral motivation of each person (Klein et al., 2015). For instance, participants learned that the person in the picture wanted to hit someone (i.e., a physically threatening motive) or criticize someone (i.e., a socially threatening motive). Interestingly, learning about these intentions changed how much participants attended to the subsequently viewed faces. That is, participants paid more sustained attention to faces—as indicated by larger LPP amplitudes—when the faces were embedded in a socially, compared to a physically, threatening

context. Thus, characteristics of the social context, such as the relevance of a feedback sender and the framing of a situation, can influence participants' attentional processing of subsequently viewed stimuli.

Building on this prior work (Klein et al., 2015; Schindler et al., 2020; Schindler & Kissler, 2018), we aimed to test whether communicating other- (vs. self-) enhancing intentions can modulate the processing of feedback messages in a similar vein. We investigated how people attend to feedback messages on their morality by using indicators of 'motivated attention'. We initially predicted that our intervention to communicate other- (vs. self-)enhancing intentions may increase participants' attention to negative feedback messages on their morality (as indicated by increased P300- and LPP-amplitudes). However, we also examined another cognitive process which became apparent during visual inspection of our EEG data. That is, we also investigated whether initial attentional processing was modulated by our intervention, as indicated by increased or decreased P200-amplitudes. Such modulations may give an indication on how attentionally vigilant participants were when processing negative feedback messages on their morality.

A potential drawback of Study 3.1 relates to how the control condition was implemented in our experimental design. Since we aimed to have a group of participants taking the task without being presented with any intentions, the control condition was part of a between-participants factor. However, this implies that our effects of the factor 'communicated intentions', which was manipulated in the other two between-participants conditions, could potentially be explained by a contrast effect. That is, because we cannot directly compare the two types of intentions to a control condition where no intentions were communicated. Higher scores of perceived fairness for feedback where other-enhancement intentions were communicated could potentially stem from the comparison group (i.e., self-enhancing intentions) being perceived as very negative. In Study 3.2 we addressed this issue by including the control condition in the within-participants factor of 'communicated intentions'. This way, we can directly compare trials where other-enhancement intentions were communicated to trials where no intentions were communicated.

4.6.3 | Method

4.6.3.1 | Design and participants

The experiment had a 2: Sender's group-membership (in- vs. outgroup, within-participant factor) x 3: Communicated intentions (other- vs. self-enhancing vs. control [no intention communicated], within-participant factor) factorial design. The trials were presented in a, prior to the experiment randomized, order. Hypotheses, exclusion criteria for participants and outliers, sample size, and preprocessing steps for the EEG analyses were preregistered before data collection at [aspredicted.org](https://osf.io/vhix7/) and can be found at <https://osf.io/vhix7/>.

We based our effect size estimate for sample size calculation on the same research as Study 3.1 now using ERP effect sizes as an indicator to estimate the minimum sample size. In this research (Rösler, Van Nunspeet, et al., 2021b, p. 223) another strategy was tested to increase the effectiveness of feedback: Contrasting feedback of one's morality with feedback on one's competence. For Study 3.2, we focused on the effect of this strategy on ERP amplitudes related to motivated attention (i.e., P300, LPP). The effect of dimension (i.e., morality vs. competence) addressed in the negative feedback message on the P300 was: $\eta_p^2 = .20$, and on the LPP: $\eta_p^2 = .12$. We used repeated measure ANOVA as a proxy for linear mixed models. Sample size calculation with G*power ($\alpha = .05$, $1-\beta = .80$) revealed a minimum of 22 participants for the factor communicated intentions (3 levels). We oversampled this number because we anticipated having to exclude participants with poor data quality (i.e., not having enough trials left to investigate ERPs).

A total of 38¹⁶ right-handed Dutch participants with no history of neurological or psychiatric problems and normal to corrected vision took part in the experiment. As in Study 3.1, it was necessary for our intention manipulation to be credible that participants found it important to give money to people in need. However, one participant indicated that they did not find it important (i.e., a score below 4 on the scale of our measure) and was therefore excluded from further analyses. Three participants had to be excluded due to technical problems. Therefore, the behavioral sample consisted of 34 participants (M_{age}

¹⁶ Initially, we had plan to test 50 participants, however, due constraints in lab time and since we already had made our minimum sample we stopped here

= 23.48, $SD = 4.17$, 27 female, one participant not indicating age). For the EEG analyses, we had to exclude 13 more participants as the data retained too few observations to achieve a good signal-to-noise ratio (SNR) after preprocessing. This number was also influenced by our unbalanced design, as trials included in the analyses depended on participants' choices in the paradigm (i.e., to be able to focus on participants' processing of negative feedback, we only included trials in which they made proself decisions). The final EEG sample consisted of 21 participants. The study was approved by the local Ethics Review Board. Participants were fully debriefed, compensated with course credit or €18, and thanked for their participation.

4.6.3.2 | Procedure, instruments, and stimuli

The procedure was very similar to Study 3.1. We made a few adjustments to the stimuli, dependent variables, and the length of the experiment to measure an EEG: We increased the trial number to 170 to be able to examine ERP responses. For that reason, we also extended the list of charities for which participants were asked to make an allocation decision. Additionally, we included an affirmation phase in each trial (e.g., Schindler & Kissler, 2018): We explained to participants that the feedback senders were presented with several traits and had to select which traits they thought were representative of the character of the participant—based on the participant's allocation decision. We used this affirmation phase to ensure that ERPs were related solely to receiving the actual feedback, rather than the sensory language processing of the trait. Finally, because the main focus of this study was on the effect of our intervention on cognitive processing of feedback messages and our manipulation proved to be successful in Study 3.1, we did not measure 'inferred good intentions' in this study. Including it would have increased the duration of the experiment (which in its current form was 90 minutes).

Each trial (see Figure 4.3) started with the presentation of the charity and participants were asked to allocate a hypothetical amount of money between themselves and the charity. This was followed by a fixation cross (jittered duration, 750-1250 ms); a picture of a sender's face with the background color indicating ingroup or outgroup membership (500 ms); an intention (or no intention, in the control condition, 5000 ms); followed by a fixation

cross (1000 ms). Then, the feedback message (i.e., in form of a character trait) was first presented to the participant (jittered duration, 750-1250 ms), and then either affirmed (i.e., trait turned orange) or not affirmed (i.e., trait stayed black, 1000 ms). Finally, the dependent measure perceived fairness of feedback was measured.

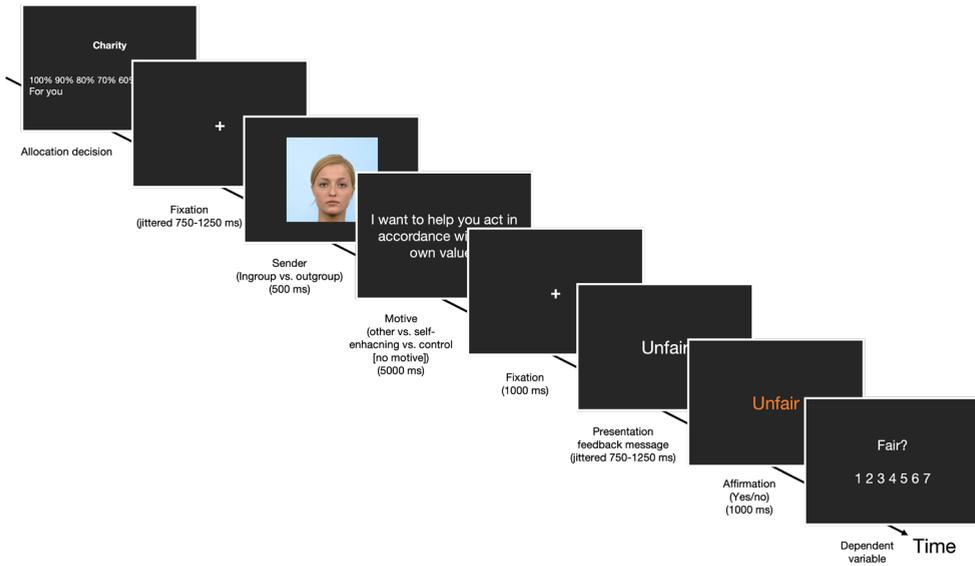


Figure 4.3. Example of trial

4.6.3.3 | EEG acquisition

We recorded EEG from 35 electrodes embedded in a stretch head cap and positioned according to the 10-10 system: F7, F3, F1, Fz, F2, F4, F8, FC3, FC1, FCz, FC2, FC4, T7, C3, C1, Cz, C2, C4, T8, CP3, CP1, CPz, CP2, CP4, P7, P5, P3, P1, Pz, P2, P4, P6, P8, POZ, and Oz. For the recording of the signal, we used the Biosemi active-electrode system with a sampling rate of 256 Hz. This system uses an analog hardware filter at 1/3 of the sampling frequency to prevent aliasing. For ground recording voltages and initial referencing, we used CMS and DLR. Electrode impedance was kept below 5 k Ω . We recorded horizontal and vertical eye movements to correct for eye movement. EEG activity was recorded with ActiView software.

Offline, we analyzed the EEG using Brain Vision Analyzer 2.1 (Brainproduct GmbH, Munich, Germany). The EEG was re-referenced to the average of the left and right mastoids. We used the regression approach (Gratton et al., 1983) to correct for ocular artifacts, filtered the signal (0.01-30Hz), and rejected trials with movement artifacts. We created separate stimulus-locked epochs for the feedback stimulus -200 ms prior to 800 ms after the event. Epochs were averaged and baseline corrected with the averaged signal 200 to 0 ms before the event. Lastly, we created separate epochs for each of the three conditions (i.e., other- vs. self-enhancing vs. control [i.e. no] intention communicated) of our experimental paradigm.

4.6.3.4 | EEG analyses

Our experimental design was unbalanced because the exact trial number per condition was dependent on participants' allocation decisions. As this led to a different number of trials between conditions, we used mean rather than peak amplitudes for the ERP analyses to avoid bias in noise levels that can occur with differing trial numbers (Luck & Kappenman, 2018).

Consistent with past literature (Herbert et al., 2006; León et al., 2010; Rösler et al., 2021b, p. 223), P200 amplitudes were largest at fronto-central and centro-parietal electrodes Fz, FCz, Cz, and CPz. Mean amplitudes for the P200 were averaged between 200–230 ms post-stimulus-onset. Consistent with past literature (Polich, 2007; Rösler et al., 2021b, p. 223; Schupp et al., 2004, 2006), P300 and LPP amplitudes were largest at centro-parietal electrodes Pz, CPz, and Cz. Mean amplitudes for the P300 were averaged between 300-400 ms and for the LPP between 400–700 ms post stimulus-onset.

4.6.4 | Results and Discussion: Self-report measures

In the following, we first report checks of our experimental manipulations, followed by the self-report and the EEG results.

4.6.4.1 | Checks

As intended, a paired sample t-test showed that participants identified more with ingroup senders ($M = 4.65$, $SD = 1.02$) than with outgroup senders ($M =$

2.52, $SD = .93$, $t(33) = 10.77$, $p < .001$, 95% CI = [1.73, 2.53].

4.6.4.2 | Comparing negative to positive feedback messages

Replicating the results of Study 3.1, one-sample t-tests against the midpoint of the scales (i.e., 4) showed that averaged scores of perceived fairness were significantly lower from the midpoint of the scale for negative feedback ($M = 1.88$, $SD = .63$), $t(33) = -19.81$, $p < .001$, 95% CI = [1.66, 2.09]. This means that negative feedback was generally perceived as not very fair. Positive feedback was perceived as fair, as indicated by averaged scores of perceived fairness being significantly higher from the midpoint of the scale, ($M = 4.66$, $SD = 1.47$), $t(31) = 2.53$, $p = .017$, 95% CI = [4.13, 5.18]. Both means were also significantly different from each other, $t(42) = -9.90$, $p < .001$, 95% CI = [-3.35, -2.12].

4.6.4.3 | Self-report results: Perceived fairness of negative moral feedback messages

To test whether communicating other-enhancing intentions increased participants' perceived fairness of negative moral feedback compared to self-enhancing intentions (replicating the results of Study 3.1) and compared to communicating no intentions, we fitted an LMM (ML) with perceived fairness as the dependent variable. We added participant number as a random effect and communicated intentions (other- vs. self-enhancing intention vs. control [no intention]) and the sender's group-membership (ingroup vs. outgroup) as fixed effects. In line with our predictions and the results of Study 3.1, there was an effect of communicated intentions on the perceived fairness of negative moral feedback messages. Compared to the trials where no intentions were communicated (i.e., control condition), when senders communicated other-enhancing intentions participants perceived feedback messages as more fair, $B = .09$, $t = 2.34$, $p = .019$, 95% CI = [.01, .16]. Similarly, communicating other-enhancing intentions also increased perceived fairness compared to when self-enhancing intentions were communicated, $B = .10$, $t = 2.85$, $p = .029$, 95% CI = [.01, 1.78]. The control condition did not differ from the self-enhancing condition, $B < .00$, $p = .911$. Also as in Study 3.1, adding the sender's group-membership to this model did not significantly increase the

model fit and was therefore not included in the final model, $p = .187$.

Thus, we replicated the finding that even though participants were reluctant to perceive negative feedback on their morality as fair, communicating other- (vs. self-)enhancing intentions enhanced perceived fairness. Moreover, we show that this effect is not only a contrast effect that occurs because we compare other-enhancement intentions to self-enhancement intentions which are likely perceived as very negative. Rather, we demonstrate that other-enhancement intentions also increase perceived fairness when compared to trials where no intentions were communicated. Furthermore, the sender's group-membership did not interact with this effect (as in Study 3.1). Thus, also in this study people did not respond more defensively to outgroup as compared to ingroup feedback senders when intentions were communicated.

4.6.5 | Results and Discussion: Event-related brain potentials (ERPs)

4.6.5.1 | P200¹⁷

Visual inspection of the data in the time window 200–230 ms post-stimulus-onset seemed to show potentially interesting changes in the P200 amplitudes for the factor 'communicated intentions'. We therefore decided to test to what extent P200-amplitudes were modulated by this factor using an LMM (ML) to predict mean amplitudes. We submitted participant number as a random effect and communicated intentions (other vs. self-enhancing intention vs. control [no intention communicated]) and electrode site (Fz vs. FCz vs. Cz vs. CPz) as fixed effects to this model. Interestingly, communicated intentions modulated P200-amplitudes elicited by the negative moral feedback messages (see Figure 4.4). Compared to the control condition, where no intentions were communicated, P200-amplitudes were decreased when other-enhancing intentions were communicated $B = -1.84$, $t = -5.89$, $p < .001$, 95% CI = [-2.45, -1.23]. This may indicate that negative moral feedback messages were semantically processed with less vigilance when other-enhancement (vs. no) intentions were communicated. Similarly, communicating other-enhancing intentions also decreased P200-amplitudes compared to the self-enhancing condition, $B = -1.38$, $t = -4.42$, $p < .001$, 95% CI = [-1.99, -.77]. The difference between the self-enhancing and the control

condition was non-significant, $B = .46$, $p = .145$. There were also main effects of electrode site. The electrodes FCz, Cz, and Fz elicited larger amplitudes ($B = 2.17$, $p < .001$, $B = 1.62$, $p < .001$, $B = 1.24$, $p < .001$, respectively) than CPz. Communicating intentions did not interact with electrode site, all B s $< .49$, p s $> .581$.

These findings seem to suggest that communicating other-enhancing intentions reduced vigilance towards negative moral feedback compared to when no intentions were communicated or when self-enhancing intentions were communicated. Strikingly, this pattern is in line with the pattern of our self-report findings: Communicating other-enhancing (vs. no and self-enhancing) intentions enhanced the perceived fairness of negative moral feedback. Taken together, these findings may suggest that providing context by embedding negative moral feedback in a helpful frame by communicating other-enhancing intentions might make participants perceive such feedback as less threatening. Interestingly, feedback with no frame or embedded in a frame where self-enhancing intentions are emphasized led to similar responses. It is possible that in the control condition, self-enhancing (or other hostile) intentions were inferred—as per default.

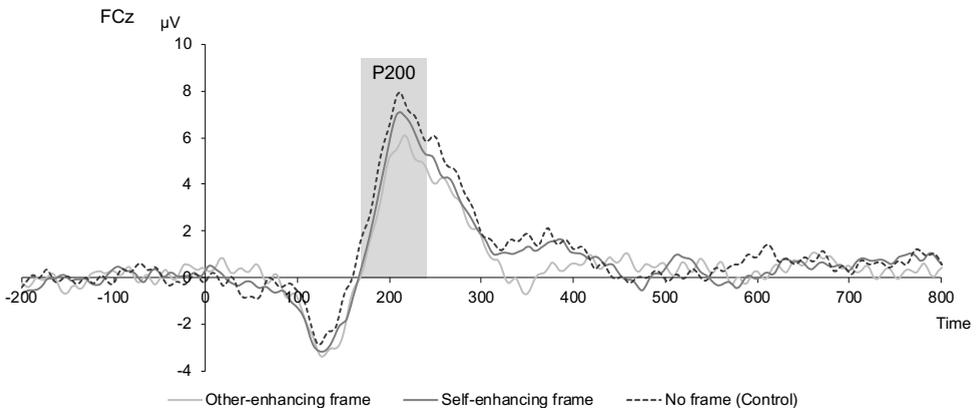


Figure 4.4. ERP-waveforms for negative moral feedback messages at frontal midline electrode FCz for other-enhancing, self-enhancing, and no intentions (i.e., control condition) communicated

4.6.5.2 | P300

We used an LMM (ML) predicting mean P300 amplitudes and submitted participant number as a random effect and communicating intentions (other- vs. self-enhancing intention vs. control [no intention communicated]) and electrode site (Pz vs. CPz vs. Cz) as fixed effects to this model. Contrary to our predictions, communicating other-enhancing (vs. no) intentions did not increase sustained attention towards negative moral feedback (as would have been indicated by increased P300 amplitudes), $B = -.28$, $p = .315$. Rather, communicating self-enhancing intentions increased P300 amplitudes compared to the control condition, $B = -.82$, $t = -2.93$, $p = .004$, 95% CI = [-1.37, -.28], and the other-enhancing condition, $B = -1.10$, $t = -3.94$, $p < .001$, 95% CI = [-1.65, -.56] (see Figure 4.5). There was also a main effect of electrode: Compared to CPz, Pz amplitudes were larger, $B = .75$, $p = .007$. There were no interaction effects, $B_s < .44$, $p_s > .505$.

These results suggest that participants were especially motivated to attend to negative feedback messages on their morality when these were embedded in a self-enhancing (vs. no or other-enhancing) frame. People were possibly angry or annoyed by receiving negative feedback on their morality when senders communicated that they wanted to demonstrate moral superiority (Minson & Monin, 2011; Monin et al., 2008).

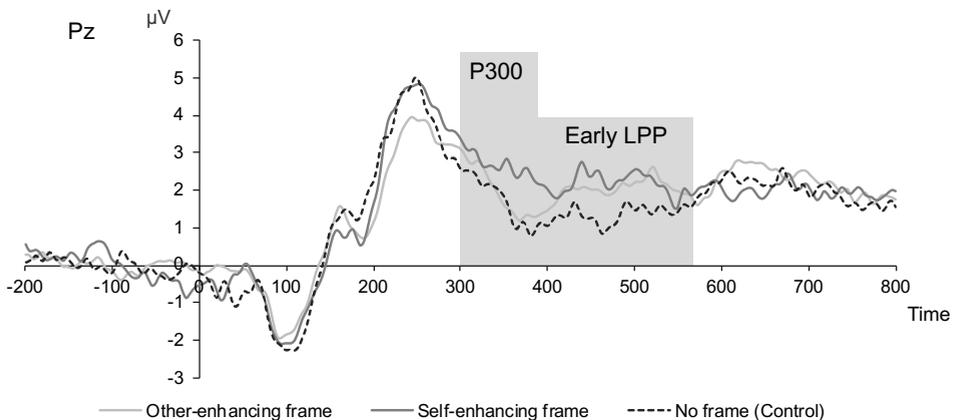


Figure 4.5. ERP-waveforms for negative moral feedback messages at parietal midline electrode Pz for other-enhancing, self-enhancing, and no intentions (i.e., control condition) communicated

4.6.5.3 | LPP

We used an LMM (ML) to predict mean LPP amplitudes. We submitted participant number as a random effect and communicating intentions (other- vs. self-enhancing intention vs. control [no intention]) and electrode site (Pz vs. CPz vs. Cz) as fixed effects to this model. There was a main effect of electrode: Compared to CPz, Pz amplitudes were larger, $B = .56$, $p = .013$, and Cz amplitudes were smaller, $B = -.70$, $p = .002$. Similar to the results for the P300, communicating self-enhancing intentions compared to the control condition (no intentions) seemed to slightly increase LPP amplitudes, $B = .39$, $t = 1.76$, $p = .080$, 95% CI = [-.04, .83]. Visually inspecting the grand averages (see Figure 4.5) seemed to suggest that this effect was more pronounced for the earlier LPP (400–600 m) rather than the later LPP (600–800ms). To investigate this further, we ran separate analyses for these two time-windows (as in previous research, e.g., Schindler & Kissler, 2018). This revealed that indeed, compared to the control (i.e., no intentions), communicating self-enhancing intentions increased early LPP amplitudes, $B = .69$, $t = 2.82$, $p = .005$, 95% CI = [.21, 1.61]. The differences between the other-enhancing and self-enhancing intentions condition and other-enhancing intentions and the control condition were non-significant, $B = .35$, $p = .154$, $B = -.34$, $p = .168$, respectively. For the late LPP, there were no effects of communicated intentions, $ps \geq .953$.

The ERP-results show two main findings which we had not anticipated, but which, as we will further explain in the General Discussion, are very consistent with our self-report findings. Compared to the control condition, a helpful frame (i.e., communicating other-enhancing intentions) seemed to decrease vigilance/ initial attention (i.e., lower P200) towards receiving negative moral feedback and a self-enhancing frame seemed to increase sustained attention (larger P300, LPP).

4.6.6 | General Discussion

People often try to motivate others to act morally by criticizing them. However, this type of negative feedback often backfires as it makes the person being criticized defensive (Rösler et al., 2021a; Täuber & van Zomeren, 2013).

Moreover, if the feedback is delivered by an outgroup, rather than by an ingroup, member people are even more likely to respond defensively to being criticized (Hornsey & Imani, 2004; Hornsey et al., 2002). This poses a problem because there are many situations where an outsider perspective is needed to call out the immoral behavior of ingroups (e.g., Cavazza et al., 2014; Ellemers & van Nunspeet, 2020; Thau et al., 2015; Van der Lee et al., 2017). Therefore, with the current research, we aimed to find new ways of making feedback receivers more open to such feedback. For this aim, we identified three problems that outgroup senders face when trying to motivate others to show moral behavior, building on social psychological research from three research lines (i.e., intergroup sensitivity effect [Hornsey & Imani, 2004; Hornsey et al., 2002], moral criticism and moral ‘do-gooders’ [Monin et al., 2008; Rösler et al., 2021a; Täuber & van Zomeren, 2013], and social comparison theory [Festinger, 1954; Wills, 1981]). The first problem is that people often respond defensively when their morality is criticized. The second problem is that people assume that outgroup critics have less constructive intentions than ingroup members. The third problem is that people may be inclined to interpret moral critique as a way for the feedback sender to show off that they are morally superior. Experiencing such behavior often annoys people, which gives them another reason to disregard moral critique by getting defensive. To alleviate these problems, the current research tested a new intervention: Having feedback senders communicate their helpful intentions when giving feedback. Even though feedback senders usually have a specific intention or a goal in mind with giving feedback, they do not always make this explicit. This may cause feedback receivers to incorrectly *infer* intentions and such assumptions may give rise to the three problems mentioned above. Our intervention, however, spoke to all these three problems as it a) eliminated greater defensiveness towards outgroup (vs. ingroup) criticism, b) increased perceived fairness of negative moral feedback, and c) decreased attentional vigilance when cognitively processing feedback messages (i.e., decreased P200-amplitudes). In the following we discuss these main findings in more detail.

4.6.6.1 | Communicating wanting to help to increase impact of negative moral feedback

Our results provide evidence that the three problems we identified are indeed relevant when outgroup members deliver feedback on other people's morality. Firstly, in line with growing literature on the ineffectiveness of moral criticism (Rösler et al., 2021a; Täuber & van Zomeren, 2013; Van der Lee et al., 2016), our results revealed that participants generally perceived negative feedback on their morality as unfair and as less fair than positive feedback on their morality. And secondly, in line with previous research that demonstrated the ineffectiveness of outgroup criticism (Hornsey & Imani, 2004; Hornsey et al., 2004; Rösler et al., 2021a) the group-membership of feedback senders affected participants' (self-reported) responses to receiving feedback. More specifically, we found that participants inferred more negative intentions from outgroup, compared to ingroup, feedback senders before receiving feedback from them. Moreover, in our control condition in Study 3.1, where feedback senders did not communicate their intentions (often the default in feedback situations), participants perceived negative feedback on their morality as less fair when delivered by an outgroup than by an ingroup member. However, we demonstrated across two studies that communicating intentions can extinguish these effects. Firstly, when intentions were made explicit, increased defensiveness towards outgroup (vs. ingroup) feedback, as indicated by perceiving negative moral feedback as less fair, was eliminated when intentions were communicated (Study 3.1). Secondly, we also demonstrated that communicating other-enhancing (vs. self-enhancing and no) intentions can decrease defensiveness towards negative moral feedback. Participants perceived negative moral as more fair when other-enhancing (vs. self-enhancing, Studies 3.1 and 3.2 and no, Study 3.2) intentions were communicated.

Past research on the intergroup sensitivity effect has demonstrated how difficult it is to change people's perception of outgroup feedback (Esposito et al., 2013). For example, participants rejected outgroup feedback regardless of the quality of this feedback (Esposito et al., 2013). Our results paint a more hopeful picture. As an outgroup critic, communicating your intentions and not 'giving people the space' to make assumptions about your intentions

is an easy and effective way to increase the impact of your feedback. That said, using our intervention most likely also increases the receipt of ingroup feedback, but possibly to a lesser extent. To explain, whereas ingroup members might also get defensive when their morality is criticized, people often infer that ingroup members have better intentions when giving feedback than outgroup members (Hornsey et al., 2004; Hornsey & Imani, 2004).

Whereas these are very promising findings, it is important to note that our manipulation of social group-membership in this current research was relevant to the task participants had to perform (i.e., donation task). That is, ingroup members were classified as having the same SVO personality type as the participant and outgroup members as having a different SVO personality type. Moreover, for our manipulation to be credible we needed to include only participants who found it important to give money to people in need and most of our participants were classified as a 'prosocial' SVO type (see Supplementary Materials). It would be interesting to test in future research whether our findings also extend to other types of SVO types and other social categorizations. For example, it could be possible that participants classified as a 'proself' SVO type are more responsive to feedback from outgroup (e.g., prosocial SVO types) compared to ingroup members (e.g., other proself SVO types). However, based on previous research showing robust effects and using different types of group-membership manipulations (e.g., real groups at the workplace, (Esposito et al., 2013; Rösler et al., 2021a; Thürmer et al., 2019), we would predict that they would also be more accepting of feedback coming from ingroup (vs. outgroup) members.

4.6.6.2 | Reduced vigilance towards negative moral feedback

Complementing these self-report findings, ERP results of Study 3.2 seemed to suggest that participants were less vigilant towards receiving negative moral feedback when feedback senders communicated other-enhancing (vs. no or self-enhancing) intentions. Embedded in such a helpful frame, P200-amplitudes elicited by the processing of negative moral feedback were reduced compared to the control condition (in which no intentions were communicated) and when self-enhancing intentions were communicated. The functionality of the P200 for word processing (i.e., in our case feedback

messages) is less well understood than other ERPs. But, modulations of P200-amplitudes have been associated with the initial lexical encoding of words (Kissler, Assadollahi, & Herbert, 2006; Trauer, Andersen, Kotz, & Müller, 2011). Moreover, previous research has shown that associated ‘emotionality’ of a word can affect its semantic encoding (Trauer et al., 2011). In research on the social categorization of faces, larger P200 amplitudes elicited by the viewing racial outgroup (vs. ingroup) faces, have been associated with ‘rapidly occurring vigilance’ because the P200 is indicating attention deployment (Ito & Bartholow, 2009). Thus, finding that communicating other-enhancing (vs. no and self-enhancing) intentions can decrease P200 amplitudes might suggest that participants paid less initial attention to and were less vigilant towards the processing of negative moral feedback messages. Therefore, when an (outgroup) feedback sender wants to criticize other people’s morality, both our self-report and ERP results suggest that they should embed the feedback in a helpful frame by letting the person being criticized know that they intend to help. This might make receivers perceive the negative moral feedback as less threatening and as more fair.

Another interesting finding is that the self-report measure of perceived fairness did not seem to differ between the condition where self-enhancing intentions were communicated and the control condition, where no intentions were communicated. Complementing these findings, ERP results for the P200 also showed no differences between these two conditions, but both showed increased amplitudes compared to the other-enhancing condition. Participants might have been equally vigilant towards messages in a self-enhancing frame and when no intentions were communicated. It could be possible that participants inferred self-enhancing intentions when no intentions were communicated.

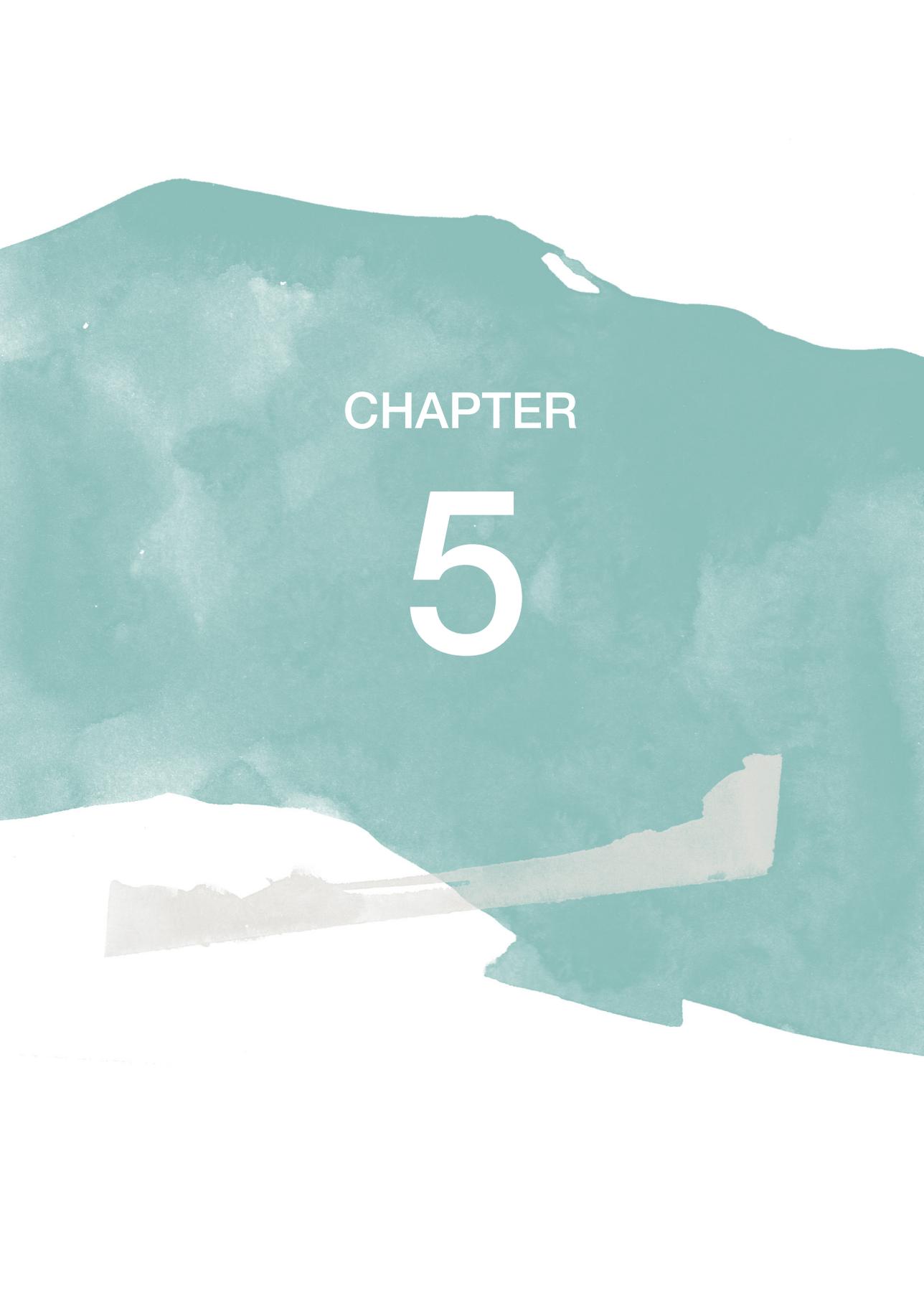
Analyses of higher-level processing ERPs (i.e., P300, LPP) showed that participants paid sustained attention to feedback when self-enhancing (vs. no and other-enhancing) intentions were communicated before the negative moral feedback was delivered, as indicated by larger P300 and Early LPP amplitudes. Thus, after showing heightened initial attention (i.e., larger P200 than the other-enhancing condition), participants seemed to also pay

sustained attention to messages in a self-enhancing frame. To understand this finding, we should consider that a self-enhancing intention entailed that senders communicated wanting to show that they are morally superior. Social psychological research on 'moral do-gooders' and social comparisons (e.g., Minson & Monin, 2011; Wills, 1981) has demonstrated that people often get annoyed if others show off their morality to make themselves feel better. Thus, increased motivated attention might stem from annoyance or increased negative impact of the feedback when it was delivered by senders with this type of intention. The finding is also in line with the theoretical reasoning of the motivated attention framework (Lang et al., 1997; Lang & Bradley, 2013). More extreme or emotional information (e.g., negative or positive vs. neutral) about the self might increase motivated attention paid to the stimulus to deal with the potential threat of it (Lang & Bradley, 2013). Participants might have been especially negatively impacted by feedback in a self-enhancing frame and therefore were more motivated to attend to these types of messages. Moreover, these findings might give new insights into why people often are more emotionally impacted by and dwell on negative (vs. positive) information about themselves (negativity bias, Baumeister et al., 2001; Rozin & Royzman, 2001). Information that is (especially) threatening to the self-image might increase initial and sustained attention and therefore also the memory of this event.

A limitation of our research is that our approach is somewhat new and different from previous studies and, therefore, our results have to be replicated by future research. In the current research, we combined social psychology with neuroscientific methods, using a new experimental paradigm that is suitable for both. This paradigm was unbalanced, meaning that the feedback participants received on their behavior depended on their choices. Even though this was needed (and intended to) increase the relevance and credibility of feedback for participants, it also comes with drawbacks. That is, participants were not always presented with the same number of trials per condition. In the current research, we, therefore, used linear mixed modeling for the analyses of self-report and EEG data. This method can account for 'missing' data. Moreover, we used mean, rather than peak amplitudes for the ERP analyses so a differing number of trials is less influential. However,

future research should replicate our findings using a more balanced design.

Nevertheless, the current research offers promising findings with important implications: Rather than letting others make assumptions about their intentions when giving feedback, we advise outgroup feedback senders to communicate their intentions and to emphasize their willingness to help. Doing so might prevent defensive reactions from feedback receivers and with that improve the chances of getting their voices heard.

The background is a teal watercolor wash with a torn paper effect at the bottom. The text is centered in white.

CHAPTER

5

General Discussion

When we try to encourage others to change their behavior, we often use feedback to communicate our views. When we give positive feedback to others, such as complimenting a friend about donating to a charity, this is often welcomed by the person who receives such positive feedback. People are motivated to engage in similar behaviors again. Giving negative feedback to someone, on the other hand, often elicits very different kinds of reactions. Rather than being motivated to change their behavior in line with the feedback, people often dismiss such feedback by getting defensive. Such a defensive response can take different forms. For instance, a person who was criticized may blame the feedback sender for wanting to intentionally hurt them or wanting to make themselves feel better. This may especially happen when the person giving the feedback does not communicate why they give such feedback, or in other words, what their intentions were when giving the feedback. This leaves much room for the person who is being criticized to make negative assumptions about the motives of critics such as wanting to make themselves feel better. In such situations, where people do not have much information about why they were criticized, they may then infer motives and base these assumptions on information they do have access to in the given moment. For example, they may use information such as whether the feedback sender seems similar to oneself or shares other similarities such as belonging to the same religion or having the same nationality (i.e., being an ingroup member, Tajfel & Turner, 1979). If perceived as similar, people may infer that the feedback sender has good intentions and wants to help them grow or become better at something (Hornsey et al., 2004). In contrast, negative feedback from people who do not share such similarities and who are perceived as very different from the self (e.g., practicing a different religion and therefore belonging to an outgroup) is often perceived as potentially harmful (Hornsey & Esposito, 2009). In daily life, we can witness such differing responses to feedback coming from ingroup compared to outgroup sources frequently. For instance, at the workplace, we might witness a colleague reacting defensively when their work is criticized by someone who holds a different position. However, when we see their work being criticized by a close colleague, they might be more open to listen to such criticism. Similarly,

a politician from the ruling party might defend a certain standpoint on an issue when being criticized for it by someone from the opposition, but not when being criticized for it from a close family member.

Such differing responses depending on whether a feedback sender shares (or does not share) a social group-membership with the self might be even amplified when the feedback is addressing something that is a core concept of how people see themselves and are seen by others, namely whether they are a moral person (Aquino & Reed, 2002; Brambilla & Leach, 2014; Pagliaro et al., 2016; Strohmingner & Nichols, 2014). Being criticized for past moral failures, such as not being a team player at work or lying about something makes people feel threatened because it threatens their self-view of being a 'good person' (Gausel & Leach, 2011). Moreover, people do not want others to notice their morally questionable behaviors because signaling immoral and rule-breaking behavior might lead to social exclusion from the communities they belong to (Ellemers & van den Bos, 2012; Van der Lee et al., 2017). Thus, people often defend themselves when they are criticized for immoral behavior, for example by hiding their moral failures (Carlson et al., 2020; Gausel & Leach, 2011; Täuber & van Zomeren, 2013). However, this represents a problem. In many of these social settings where people receive criticism from outsiders and where their morality is criticized, the positive side of negative feedback – being encouraged to improve oneself – is lost. That is, because people infer that those others want to harm them and therefore dismiss the negative feedback by getting defensive. Besides such obvious spoiled opportunities for personal growth, dismissing negative feedback might additionally create potential for disagreement between groups in our societies in a time that is already marked by conflict and polarization. To explain, during the global Covid-19 pandemic we as humans had to face ethical questions which we did not have to face to such an extent before. For instance, hospital workers and politicians had to make decisions on moral dilemma situations such as who should be prioritized for ICU beds. Moreover, the distribution of the vaccine created grounds for potential conflicts between countries. In such a tense climate, it is essential to deliver and receive potential negative feedback from others on ethical questions without further escalating the already critical situation. In addition to this, our societies have

become increasingly diverse and polarized in the last decades, creating more potential conflicts between group members and making it more difficult to communicate constructively with each other (e.g., Finkel et al., 2020). With the research in this dissertation, I aimed to contribute to addressing such pressing issues by finding new ways to encourage each other to be more open to receiving feedback from people who are different than us (i.e., outgroup members) and on sensitive topics, such as moral behavior.

5.1 | How has past research studied how people respond to moral criticism and criticism from outgroups?

Studying moral behavior has become extremely popular in social psychology over the last decades (Ellemers et al., 2019). However, not much research has specifically addressed morality in the context of intergroup relations and communication (Ellemers et al., 2019). Moreover, whereas understanding how people judge others in terms of their moral behavior has been extensively studied, little research has focused on how individuals who receive these judgments respond to them. Only in the last decade, the interest in understanding how people process being evaluated on their moral behavior has inspired new research questions (e.g., Täuber et al., 2015, 2018; Täuber & van Zomeren, 2013; Van der Lee et al., 2016; Van Nunspeet et al., 2015; Van Nunspeet et al., 2014). This previous research demonstrated that: 1) People are motivated to be perceived as moral by others, 2) that people find it harder to cope with being negatively evaluated on their morality than on their competence, and 3) that moral criticism can make people respond defensively because such criticism is perceived as threatening. However, most of this research faces a fundamental problem: People do not like to reflect on their moral failures, often hide these in front of themselves and others, or emphasize having changed since the moral failure to keep a positive self-image (Carlson et al., 2020; Gausel & Leach, 2011; Kouchaki & Gino, 2016; Mazar et al., 2008; Shalvi et al., 2015; Stanley et al., 2019). To deal with this problem, with the research in this dissertation I took a different approach. I combined traditional social psychological measures (e.g., self-reports) with measures from neuroscience (i.e., measuring participants' EEG) and economics (e.g., monetary contributions in an experiment). Combining these measures provides insights into both explicit and implicit responses

to receiving negative feedback. That is, besides insights into what people decide to report on (i.e., reflecting on receiving feedback) I also investigated brain responses to receiving feedback in the actual moment using event-related potentials (ERPs). These measures can give insight into how people cognitively process feedback. More specifically, the ERPs investigated in this dissertation (i.e., P200, P300, LPP) can measure how motivated participants are to pay initial and prolonged attention to feedback messages (Lang et al., 1997; Schindler et al., 2020; Schindler & Kissler, 2018). Thus, they may represent a way to measure what people want to hide or cannot consciously reflect on (e.g., paying more attention to feedback coming from certain sources over others). Together with the self-report and behavioral data they can therefore offer a more comprehensive picture of which psychological processes occur when people receive negative feedback from outgroup members and on their morality.

Besides investigating how people respond to being criticized for their immoral behavior, the second aim of the research in this dissertation was to investigate how people may respond differently to being criticized based on who is delivering the criticism. Previous research has shown that when outsiders criticize groups we belong to, for example the country we grew up in, people are less likely to accept such criticism compared to when the critic shares our nationality (Esposito et al., 2013; Hornsey et al., 2002). The research in this dissertation first investigated whether these effects also occur when not the own group, but the self, is criticized by an outgroup member. After showing in Study 1.1 that this is indeed the case, I then tested several strategies to make people more open to such criticism.

To summarize, I investigated two main research questions:

- 1) How do people deal with being criticized for their moral behavior and how can we make people more open to such criticism?
- 2) Do people react differently to criticism if it is delivered by an outgroup compared to an ingroup member – and if yes – how can we intervene to prevent this from happening?

5.2 | Successful strategies to make people more open to receiving criticism

The results are promising. They show that there are two ways to increase the effectiveness of outgroup feedback on people's morality which can easily be applied in practice: 1) Making a person's competence, rather than their morality, the focal point addressed in a feedback message and 2) communicating one's goodwill when delivering feedback. Both strategies refer to very simple adaptations of how feedback is (re)framed and therefore can be used by anyone who wants to prevent defensive responses when delivering constructive feedback. Moreover, the strategies can universally be applied to all types of intergroup situations and might help outgroup critics getting their voices heard in situations where they typically are not being listened to (i.e., when delivering feedback).

In the following, I will elaborate on how the research findings from each chapter contributed to these two main findings. I will discuss self-report, behavioral, and ERP findings for these two strategies separately. I will embed the findings in the social psychological and neuroscientific literature and show how the current dissertation extends previous research. Finally, I will discuss societal implications, limitations, and avenues for future research of the research in this dissertation.

5.3 | Criticism delivered by outgroup senders and on one's morality is often dismissed

When asked to reflect on situations where participants received criticism at their workplace, I got a first insight into who usually delivers criticism and what type of criticism people receive in work contexts. Content coding of participants' open-ended questions (Studies 1.1 and 1.2) revealed that when asked to reflect on criticism situations where participants were criticized by someone who is perceived as similar (i.e., an ingroup member) they often mentioned to have received criticism from colleagues in similar positions. When asked about criticism delivered by people who aren't similar or are outsiders at the workplace (i.e., an outgroup member), they most frequently specified such critics as colleagues from different work teams or departments. In terms of content of criticism, people seem to receive criticism of their

morality at the workplace for non-cooperative and unethical behaviors such as letting colleagues do more work or not following company guidelines. When criticized for their competence, such critique often seems to refer to performance related behavior such as not performing a task fast or careful enough. Most importantly, people also seem to react differently to these types of critiques. That is, results of Chapters 2, 3, and 4 (Studies 1.1., 1.2., 2.1, 2.2, and 3.1) demonstrated that people respond more negatively to criticism delivered by outgroup compared to ingroup members. It was perceived as less legitimate, less impactful, and less effective than criticism delivered by ingroup members. In terms of content of criticism, criticism seems to be less effective when it addressed unethical behaviors compared to when it referred to performance-related behaviors (Studies 1.1., 1.2). When criticism refers to unethical behaviors at the workplace, people seem to be less motivated to improve their behavior, made more negative assumptions about the motives of the critic and the credibility of the criticism, and less often changed their behavior. The findings demonstrate that there are many ways how people can respond to criticism at the workplace and that these responses may partly depend on who is delivering the criticism and what the criticism refers to. This is important, because often when we deliver feedback to others and receive a defensive response, we might not realize that part of this defensiveness stems from the person perceiving us as an outsider or from how we framed the critical message. But these findings not only offer new perspectives on why people may not respond well to criticism in our daily life, they also make important contributions to previous social psychological literature.

What do we know now that we did not know before?

The findings extend research which investigated how people respond to criticism from outsiders when this criticism is aimed at a social group they belong to (e.g., a non-Australian criticizing Australians for being racist, intergroup-sensitivity effect, Esposito et al., 2013; Hornsey et al., 2002, 2004). Whereas this previous research showed that people often do not accept such criticism from outgroups, the research in this dissertation demonstrates that people respond similarly negative to outgroup criticism when they themselves as individuals are criticized. In both real-life criticism

situations at the workplace and when delivering criticism to participants in an experimental setting, people were less responsive to outgroup criticism than to ingroup criticism. Thus, the findings of this dissertation make an important contribution to this previous research. Firstly, they make theoretical contributions by showing that similar psychological processes appear when individuals (as when groups) are criticized by outgroup members. Secondly, they demonstrate that these group processes may appear daily in our workplaces but also that they additionally occur under strict experimental settings in the lab. With this, they open exciting new possibilities for new research questions investigating how people may respond to being criticized from outsiders in different social settings and what type of interventions may increase the reception of such criticism.

The findings extend previous literature studying how people respond to moral criticism (Täuber & van Zomeren, 2013; Van der Lee et al., 2016) by offering new insights into how and why people respond defensively when they are criticized for their morality. That is, people seem to be less motivated to improve their behavior (Studies 1.1. and 1.2), make more negative assumptions about critics (Study 1.2), perceive feedback as more negative (Study 2.1), and remember feedback messages less when feedback addresses their morality instead of their competence (Study 2.2). Thus, the findings make important contributions to previous research by offering a more fine-grained understanding of which cognitive, motivational, and behavioral aspects may contribute to the ineffectiveness of moral criticism. A critic aiming to encourage others to act more prosocial and moral may not realize that emphasizing immorality in a feedback message elicits all these counterproductive mechanisms in the person being criticized. Even though these (coping) mechanism help the person criticized protect their positive self-view, they also make moral criticism practically ineffective. Future research could address each of these aspects with interventions, testing whether they might offer new ways to make people more open to accept and learn from criticism regarding their morality.

Interestingly, there was one aspect of moral criticism that was positively, rather than negatively, correlated with the acceptance and effectiveness of

criticism. That is, in Study 1.2, participants reported to have experienced more moral emotions (i.e., guilt and shame) when reflecting on being criticized for their morality than for their competence (in line with Van der Lee et al., 2016), and this, in turn, increased the effectiveness of criticism. In other words, the more guilty or ashamed people felt the more often they changed their behavior based on the criticism. One may take away from this finding that we should therefore aim to evoke these moral emotions in the people we criticize to encourage behavior change. However, besides ethical considerations of specifically aiming to make people feel ashamed of themselves or their behavior, doing so is also easier said than done. To explain this, it is important to consider how people deal with their moral failures. Gausel and colleagues (Gausel et al., 2012; Gausel & Leach, 2011) argue that people evaluate their moral failures in two possible ways: 1) As a defect of the self (i.e., a threat to the self-image), or 2) as possible condemnation from others (i.e., a threat to the social image). Only if they evaluate their moral failure as the former - a defect of the self - and additionally perceive the failure as addressing something specific (rather than addressing their global self), they are motivated to improve their behavior. If the moral failure is perceived as a global self-defect or a threat to their social image, they might self-defend or hide their moral failures. It is therefore inadvisable to try to make people change their behavior by emphasizing that they should be ashamed or that they are 'not a good' person when criticizing their moral behavior as most of the time this will lead to a self-defense response (also see Gausel, 2013; Täuber et al., 2015 for similar points).

Altogether, the (self-report) findings of Chapters 2 and 3 extend previous research investigating intergroup criticism (Esposito et al., 2013; Hornsey et al., 2002, 2004), the ineffectiveness of criticizing people's morality (Gausel & Leach, 2011; Täuber et al., 2015; Täuber & van Zomeren, 2013; Van der Lee et al., 2016), and by connecting these two literature streams to each other (which hasn't been done before to my current knowledge). However, to get a full picture of what type of psychological process may play a role in criticism situations which cannot be captured with traditional social psychological research methods, I additionally investigated how people cognitively processes receiving feedback on their morality and on their competence in

the actual moment (Chapter 3, Study 2.2). For this I made use of a method from neuroscience (i.e., EEG). Adding such measures to the self-report and behavioral results may help us understand better why people are less responsive to outgroup criticism and to criticism of their morality. That is, because they may perceive ingroup criticism as more relevant and may not be able to cope well when receiving painful evaluations of the moral self.

5.4 | Ingroup feedback is favorably processed

To get insights into psychological processes related to the cognitive processing of feedback that people may not be able to reflect on or are reluctant to share, I analyzed event-related potentials (ERPs). ERPs refer to stimulus- or response-locked fluctuations in the electrical activity that was measured with an EEG and are believed to represent cognitive processes such as attentional processing of stimuli. More specifically, I examined ERPs related to initial attention and perceptual processing of words (i.e., P200, e.g., Carretié, Martín-Loeches, et al., 2001; Kissler et al., 2006) and higher-level ERPs that have been related to whether people pay prolonged attention to stimuli (i.e., P300, LPP, Lang et al., 1997; Polich, 2007; Schupp et al., 2004, 2006). The results from using these neuroscientific markers for the attentional processing of feedback complement the self-report and behavioral findings of Chapters 2 and 3. In line with ingroup feedback being more effective (Chapter 2) and impactful (Chapter 3) than outgroup feedback, the ERP results showed that participants also pay more initial attention to feedback messages coming from ingroup members (i.e., larger P200-amplitudes, indicating that more attentional resources were allocated to the perceptual processing of messages). This is in line with previous ERP-research (Schindler et al., 2015, 2020) where feedback messages coming from more relevant senders (i.e., from a human or an expert) increased initial attention (i.e., larger P200-amplitudes) than feedback messages coming from less relevant senders (i.e., from a computer). The findings of Study 2.2 extend this research by demonstrating that also social group-memberships (in this case fellow national or international students) can modulate P200-amplitudes that are elicited by receiving feedback messages. Modulation of amplitudes refers to the changes in voltage of the brain response (ERP) that are elicited by the manipulation in the experiment (here receiving feedback

from an ingroup or outgroup sender). Finding such a modulation is noteworthy as most research in (social) neuroscience investigates rather basic stimuli (e.g., positive or negative word stimuli, happy or angry faces). Investigating more social phenomena such as cognitive processes related to social identity is still only a small part of social neuroscience (however, take a look at these reviews Amodio & Cikara, 2021; Cikara & Van Bavel, 2014; Ellemers & van Nunspeet, 2020). Thus, the research in this dissertation adds to this ‘more social stream’ of social neuroscience by demonstrating how the social identity of a feedback source can influence our perceptual and subsequent attentional processing of information before we become even aware of what we see (i.e., a few hundred milliseconds after receiving feedback).

To sum up, the social context feedback is embedded in can change how participants attentionally process feedback coming from ingroup and outgroup senders. This means that before even recognizing a person as such, people might allocate less attentional resources to feedback when it is delivered by outgroup (vs. ingroup) members. This is an important new piece of information when trying to understand how and why people respond worse to outgroup feedback than they do to ingroup feedback. It suggests that such feedback is already less effective before we consciously are aware of the meaning of it. Such early discrimination against outgroup criticism might then also influence later (and more conscious) responses such as making negative assumptions about the motives of critics.

The second factor that I investigated in the EEG study of Chapter 3 was related to how people process the actual content of the feedback. That is, whether the feedback addressed the morality or the competence of a person.

5.5 | People devote less sustained attention to morality (vs. competence) feedback

Relating to how people process the content of feedback, the ERP findings show that the social dimension addressed as the focal point of a feedback message can modulate participants’ brain responses elicited by the cognitive processing of feedback messages. More specifically, participants were more motivated to pay sustained attention (larger P300- and LPP-amplitudes)

to feedback messages that were referring to their competence instead of their morality. Thus, higher level cognitive processes such as motivated and sustained attention paid to stimuli were affected by the social dimension the feedback addressed. This finding is noteworthy because it demonstrates that people may show different cognitive mechanism depending on what a feedback message is addressing. They may potentially shift away attention from feedback when it relates to their morality, or they may perceive feedback on their competence as more relevant than feedback on their morality and therefore pay more sustained attention to it. The findings complemented the self-report and behavioral results of Chapter 3 which demonstrated that participants perceive moral (vs. competence) feedback messages as more negative (Study 2.1), remember these feedback messages about themselves less often, and remember negative feedback on their morality the least (Study 2.2). Together, these findings may suggest that people devote less sustained attention to the processing of feedback messages on their morality compared to their competence and consequently misremember (or misreport) these messages. The combined evidence offers an important new insight into why moral feedback may not always be effective and therefore has interesting implications. That is, people may not want to engage with such painful feedback and therefore perceive such feedback as less relevant. Consequently, they may also not remember such messages as well as when these referred to their competence. People may also not want to let others know that they received negative feedback on their morality to avoid signaling uncooperative behavior and therefore reported less of such messages (also see Van der Lee et al., 2016).

What do we know now that we did not know before?

Complementing the self-report and behavioral findings of Chapters 2 and 3 that showed that ingroup feedback is more effective and more impactful than outgroup feedback, the ERP findings relating to the group-membership of the feedback sender extend previous social psychological literature (Esposito et al., 2013; Hornsey et al., 2002, 2004). They demonstrate that ingroup feedback is preferentially processed milliseconds after people receive the feedback. Any interventions aimed at increasing the effectiveness of

outgroup feedback should therefore be targeted at changing the perception of the social context beforehand, rather than trying to do ‘damage control’ such as emphasizing goodwill after the criticism was delivered.

These ERP-findings extend previous neuroscientific research (e.g., Klein et al., 2015; Schindler et al., 2020; Schindler & Kissler, 2018; Schwarz et al., 2013; Wieser et al., 2014) by demonstrating that even subtle social cues such as the social group-membership of a feedback sender can influence how we attentionally process feedback messages. These findings open many new and exciting avenues for future research. One can imagine that many other kinds of social context information influence our perception in a similar fashion. For instance, the perceptual and attentional processing of feedback messages could also be influenced by how powerful or powerless a feedback receiver feels in the feedback situation (Fiske, 1993; Galinsky et al., 2006). Feeling powerful (vs. powerless) has been shown to decrease perspective taking (Galinsky et al., 2006) and according to Fiske’s attention theory may increase stereotyping by decreasing attention paid to outgroup members (Fiske, 1993). Thus, powerful (vs. powerless) feedback receivers might therefore even pay less attention to feedback coming from outgroup (vs. ingroup) members.

The ERP findings relating to how people process feedback on their morality (vs. on their competence) extend previous social psychological literature questioning the effectiveness of moral criticism (Gausel et al., 2012; Gausel & Leach, 2011; Täuber & van Zomeren, 2013; Tetlock, 2002; Van der Lee et al., 2016). The findings suggest that the ineffectiveness of moral criticism can already be captured in very early brain responses which are elicited by the cognitive processing of such messages. Interventions to increase the effectiveness of moral criticism should therefore also be aimed at changing the social context feedback is delivered in, similarly as when aiming to increase the effectiveness of outgroup criticism.

These ERP findings may inspire new research questions related to what other factors influence the attention paid to the processing of feedback messages on people’s morality. For instance, the processing could also be influenced by the language abstractness a feedback sender uses (e.g., Moscatelli et al.,

2019). To explain, previous research has shown that people are responding less negative to outgroup criticism when such criticism is framed in more concrete, rather than abstract, terms (Moscatelli et al., 2019). Thus, abstract feedback messages on people's morality (e.g., "you are a liar") might also decrease people's sustained attention paid to feedback messages compared to feedback messages framed in more concrete terms (e.g., "you shouldn't have lied in this situation", also see Gausel & Leach, 2011).

More broadly, the ERP findings of Chapter 3 add to the neuroscientific and social psychological literature on feedback processing in two additional ways. Firstly, they provide more evidence for the notion that social cues such as the social group membership of an interaction partner and the social dimension addressed with a feedback message can influence our perception before we become aware of the meaning of these cues (also see Amodio & Cikara, 2021; Ratner et al., 2014; Schindler et al., 2020; Simon et al., 2020; Van Nunspeet et al., 2014). Thus, they show how top-down information can shape our perception and how social context cues can direct attention towards more relevant stimuli (i.e., ingroup feedback, competence feedback). Secondly, they demonstrate that social context information influences cognitive and attentional processes often before what we traditionally can pick up on with social psychological measures (e.g., self-reports).

As previously discussed, these insights from Chapters 2 and 3 inspired the intervention in the final empirical chapter which was aimed at increasing effectiveness of outgroup feedback and feedback on one's morality. That is, I aimed to change the social context that feedback is embedded in by making feedback senders communicate their intention to help delivering negative moral feedback.

5.6 | Communicating good intentions can make people more open to criticism of their morality

Since previous research (Esposito et al., 2013; Hornsey et al., 2002, 2004; Thürmer et al., 2019) and the results of Chapter 2 have shown that people make more negative assumptions about the motives of outgroup compared ingroup critics, in Chapter 3 I aimed to intervene with this process, by having

feedback senders communicate their intention to help when giving feedback. Results of Studies 3.1 and 3.2 showed that communicating their intentions when delivering feedback can make people more open to feedback from outgroup senders and to feedback on their morality. In these two studies, the effectiveness of feedback was operationalized as less dismissal of the feedback by measuring how fair participants perceived the feedback to be. Results of Study 3.1 showed that this intervention can make people perceive negative feedback from outgroup members as more fair, that is, similarly fair as feedback from the ingroup is perceived. Moreover, negative feedback on participants' morality coming from senders who communicated the intention to help was received more positively. It was perceived as more fair than when, instead, these senders communicated the intention to enhance their self-view by demonstrating moral superiority (Minson & Monin, 2011; Monin et al., 2008, Study 3.1 and Study 3.2) and compared to when senders did not communicate their intentions (Study 3.2). Thus, making intentions explicit when giving feedback represents an effective and new way to increase both the effectiveness of (outgroup) feedback and feedback on people's morality. This finding is exciting because previous research has shown that both these types of critiques make people defensive and disregard criticism. Combining them in this research therefore increased the chance of people dismissing the negative feedback. Moreover, reframing a critical message in such a way that it refers to the competence rather than the morality of a person (as demonstrated being effective to increase openness in Chapters 2 and 3) might not always be appropriate. This new strategy, to communicate goodwill when providing feedback and to focus on the positive side of negative feedback (improving the self), therefore offers an easily to be applied and effective way to decrease the chance of people dismissing negative moral feedback.

These findings additionally rule out a possible concern of the research in Chapter 2. In Studies 1.1 and 1.2, the variables that mediated the effects of group-membership on the effectiveness of criticism (i.e., the cognitive attributions about critic, motivation to improve based on the criticism) were measured rather than manipulated. Therefore, it could be possible that the directionality of the effect is in fact not in the direction we hypothesized (i.e., making more negative attributions about outgroup vs. ingroup critics causes

the criticism to be less effective), but could possibly also be the other way around (i.e., lower effectiveness of criticism causes participants to make more negative attributions about outgroup vs. ingroup senders retrospectively, Fiedler et al., 2018). However, the findings of Studies 3.1 and 3.2 rule out the possibility of a reversed causal effect. The assumptions about the motives of feedback senders—why feedback was given—were experimentally manipulated. The results show a clear picture. When no intentions are communicated, participants infer that outgroup senders have worse intentions compared to ingroup senders and they perceive their negative feedback as less fair. When intentions are made explicit, these effects are eliminated, showing evidence for a causal relationship in the direction of the studies in Chapter 2.

What do we know now that we did not know before?

The results of Studies 3.1 and 3.2 extend previous research on intergroup criticism (Esposito et al., 2013; Hornsey et al., 2002, 2004; Thürmer et al., 2019), by demonstrating an easily to be applied method to reduce the ineffectiveness of outgroup criticism in real life situations. More broadly, the results of Chapter 4 add to social psychological research on moral evaluations (Gausel & Leach, 2011; Täuber & van Zomeren, 2013; Van der Lee et al., 2016). They introduce a new strategy to reduce the threat to the self-view when being criticized for one's morality (Gausel & Leach, 2011) which does not require to rephrase moral criticism as not actually addressing the morality of a person (e.g., addressing one's competence, see Chapter 2 and 3). Thus, this strategy offers a way to potentially make people more open to receiving moral criticism situations in situations where one needs to directly address a moral concern.

The implications of this research are straightforward. In situations where outsiders deliver feedback people often assume they do so with bad intent. When assuming so, even the best advice is perceived as potentially harmful. Making goodwill explicit, however, might be able to reduce these assumptions about potentially harmful intentions. Therefore, this method might help with the extreme negativity outgroup critics face when trying to encourage behavior change, such as defensiveness and hostile reactions of the person

being criticized (Esposito et al., 2013; Thürmer & McCrea, 2018). For instance, in situations where other people were affected by the immoral action of the person criticized (such as when others were (physically) harmed).

As in Chapter 3, I also measured participants' EEG in Study 3.2 to investigate how changing the social context, here in terms of a critic's good intentions made explicit, may affect cognitive processing of negative feedback messages on people's morality. In the following, I will discuss how the ERP findings extend previous neuroscientific and social psychological research.

5.7 | Communicating helpful intentions may reduce vigilance in response to negative moral feedback

The ERP findings of Chapter 4 shed more light on the processes that take place when people cognitively process receiving negative feedback on their morality and how these might be modulated by the social context feedback is embedded in (i.e., what type of intentions feedback senders communicate). The results demonstrate that initial attentional processes (i.e., P200) and later cognitive processes (i.e., motivated attention, P300 and LPP) can, in fact, be modulated by such social context information. These findings compliment the self-report findings of Studies 3.1 and 3.2 which showed that communicating goodwill can make people perceive negative feedback on their morality as more fair. More specifically, the modulations of the P200 amplitudes mirrored the self-report findings of these two studies. Participants showed reduced P200 amplitudes when viewing negative moral feedback after feedback senders had communicated the intention to help the participant (compared to when they communicated wanting to demonstrate moral superiority or did not communicate their intentions). To make inferences about the meaning of these modulations, it is important to note that these findings were unexpected and that the functionality of the P200 for word processing (i.e., in our case feedback messages) is less well understood than other ERPs. This is why we can only make assumptions about the underlying cognitive processes that might have led to such modulations. Previous research has associated P200 modulations with initial lexical encoding and sensory processing of words which can be modulated by stimuli features such as valence or emotionality (Kissler, Assadollahi, & Herbert, 2006; Trauer, Andersen, Kotz, & Müller,

2011, Carretje et al, 2001). This cognitive functionality in combination with the fact that the ERP modulations mirror the self-report findings might indicate that the decreased amplitudes refer to a lower vigilance associated with the processing of the negative feedback.

Such inferences would also be in line with previous research on social categorization where increased P200 amplitudes elicited by the viewing of racial outgroup (vs. ingroup) faces have been associated with a 'rapidly occurring vigilance' because the P200 ERP is indicating attention deployment (Ito & Bartholow, 2009). Participants might have perceived negative moral feedback as less threatening when feedback senders communicated their intention to help with the feedback.

Another interesting (and unexpected) ERP finding of Chapter 4 was that participants were especially motivated to pay prolonged attention to negative feedback messages when feedback senders communicated wanting to show that they are morally superior to the person receiving the feedback (as indicated by larger P300 and LPP responses compared to when wanting to help was communicated or no intentions were communicated). This might indicate that participants were especially negatively impacted by such negative feedback (or especially annoyed, Cramwinckel et al., 2013; Minson & Monin, 2011; Monin et al., 2008). Such an interpretation would also be in line with the theoretical reasoning of the motivated attention framework (Lang et al., 1997; Lang & Bradley, 2013) which states that more extreme information (e.g., negative or positive vs. neutral) about the self might increase motivated attention paid to a stimulus (Lang & Bradley, 2013).

What do we know now that we did not know before?

The ERP findings extend previous social and affective neuroscientific research on cognitive and attentional feedback processing which often focuses more on the neural basis or on clinical aspects of the processing of social feedback (e.g., Achterberg et al., 2016; Gu et al., 2020). For instance, researchers investigated how social anxiety or emotional regulation (e.g., anger in response to negative feedback) might influence how people perceptually and attentionally process social feedback (Achterberg et al., 2016). More related

to the research in this dissertation, recent research has demonstrated that social context factors such as characteristics of a feedback sender (i.e., being a human vs. being a computer) can influence how we attentionally process feedback messages (Schindler et al., 2014, 2015, 2020). The research of Chapter 4 adds to this by demonstrating the importance of taking into account other social context factors such as whether feedback senders communicate their intentions. We now have a better understanding not only of how embedding feedback in different types of contexts can modulate initial attention (i.e., the P200), but additionally might also affect later and more higher-level cognition (i.e., P300, LPP, related to sustained and motivated attention).

These ERP findings additionally extend previous research in social psychology which investigated how people react to others showing their 'moral superiority' (Cramwinckel et al., 2013; Minson & Monin, 2011; Monin et al., 2008). They do so by demonstrating that when feedback senders deliver negative feedback on someone's morality with the intention to show moral superiority, such feedback messages seem to especially hold feedback receivers' attention. Thus, our ERP measure might reflect a new way to measure annoyance associated with responses to people showing moral superiority.

The findings are important as they uncover potential underlying cognitive mechanisms that make negative feedback on one's morality more or less effective depending on what type of intentions are communicated alongside it. When the intention to help is communicated, people may perceive negative feedback messages as less threatening and therefore perceive them as more fair. This has interesting implications for using this as a strategy to encourage others to improve their moral behavior. That is, without changing the actual content of the feedback, communicating the intention to help makes people perceive the feedback they received as more fair and potentially more valuable. The results also show, though, that feedback senders should be aware not to come across as someone wanting to demonstrate moral superiority when delivering negative feedback as this might make people especially annoyed.

5.8 | Conclusion empirical evidence

The combined empirical evidence of the research in this dissertation shows three main findings: (1) People respond more negatively to criticism on their behavior when it was delivered by someone belonging to a different social group than themselves, (2) criticizing others for their morality is less effective than criticizing them for their competence, and (3) when critics want to increase the effectiveness of criticism on someone's morality, they should communicate their helpful intentions (also see Figure 5.1 for an overview). In the following I will shortly reflect on broader societal and practical implications of these main findings as well as discuss limitations and avenues for future research.

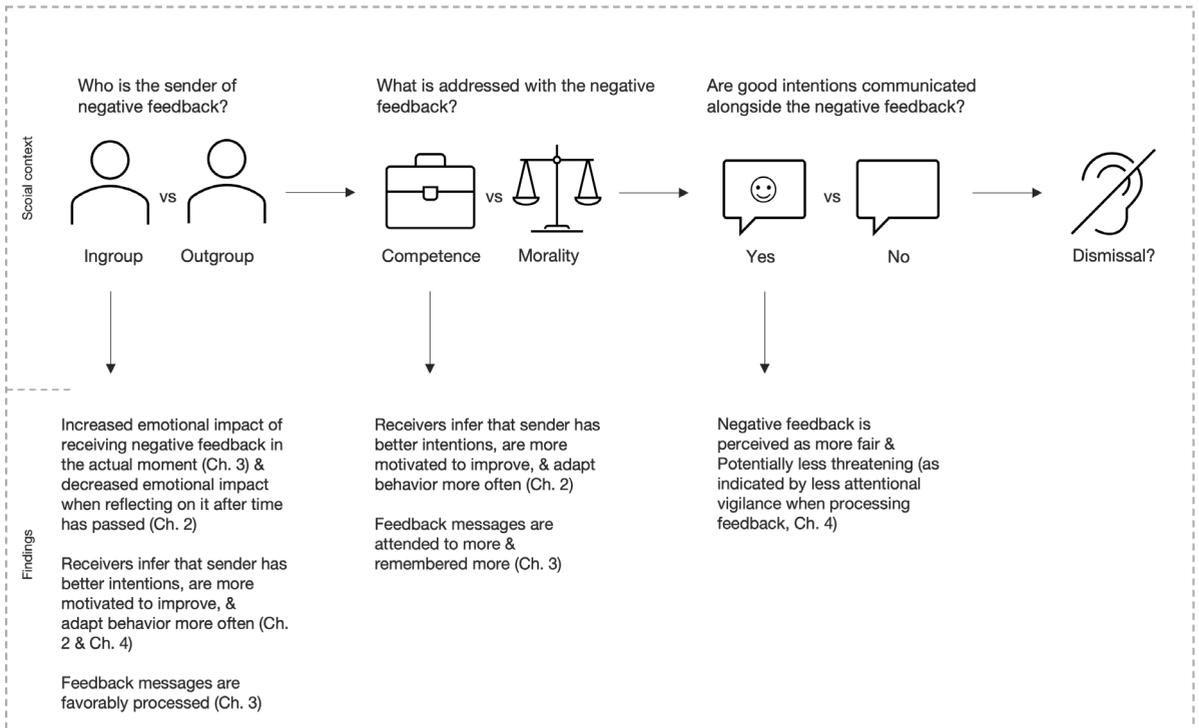


Figure 5.1. Schematic illustration of potential responses to receiving negative feedback from ingroup and outgroup members. The figure shows how the social group-membership of a feedback sender (ingroup vs. outgroup), which social dimension was addressed in the feedback (morality vs. competence), and whether good intentions were communicated alongside the feedback (yes vs. no) influences how people may respond to negative feedback. The lower part of the figure indicates the findings related to each of these three factors and is based on the empirical evidence from Chapters 2 – 4. If delivered by an outgroup (vs. ingroup) member, people may allocate less initial attentional resources to a feedback message and the feedback message is less effective. However, when competence, instead of morality, is made the focal point addressed in the feedback message, this effectiveness may increase. If morality is addressed and feedback senders do not communicate having good intentions, their feedback may be dismissed. However, communicating (vs. not communicating) having good intentions may decrease the respondent's perceived threat when receiving the negative feedback message which may consequently increase its effectiveness.

5.9 | Societal and practical Implications

The research in this dissertation has direct societal and practical implications for current politics, the media and journalism, our daily life living in diverse societies, and for the workplace.

5.9.1 | Politics

In current politics, we can see politicians criticize the morality of opponents belonging to different political parties often. For instance, a left-wing politician may criticize a member from a right-wing party for voting for a legislation preventing refugees from entering the country, calling such a standpoint immoral and disgusting. Whereas using such moral language may get the attention of the media and potential voters (e.g., on online platforms, Brady et al., 2020), it often does not have the intended effects of encouraging others to change their viewpoints or behavior (Gausel, 2013; Gausel & Leach, 2011; Täuber et al., 2015; Täuber & van Zomeren, 2013; Van der Lee et al., 2016). It might only make the opponent defensive and make them dismiss the criticism. Moreover, in voters who witness politicians using such moral language, it may lead to more 'othering' and consequently feeds polarization (Finkel et al., 2020). Since politicians need to make legislative decisions on moral topics such as abortion, climate change, and immigration daily, it is therefore essential to find new ways to increase such intergroup communication. The research in this dissertation offers two potential strategies politicians can use when discussing moral topics and delivering feedback to opponents that is referring to their morality. That is, where possible, they could reframe the feedback so that it addresses the opponent's competence, rather than their morality. Or, in situations where this is not possible, they could communicate their goodwill when delivering the feedback. Emphasizing goodwill in the above-mentioned situation could look as follows: The critic could emphasize that they delivered criticism with the ultimate goal of protecting people seeking refuge not to demonstrate that they themselves are a morally better person or want to shame the opponent. This may not only decrease defensive reactions of the person criticized, but it may additionally communicate to the public that political parties are working together on finding solutions on moral issues, rather than fighting each other.

5.9.2 | Media and journalism

Another context where people could benefit from using the strategies investigated in this dissertation is in the media and in journalism. Interviews on moral topics (e.g., racism, gender equality) where the interviewer and interviewee do not share the same moral viewpoints often end up in at least one of these communication parties getting defensive. As shown in this dissertation, this may be related to people assuming that the other party has bad intentions. For example, an interviewee that receives difficult questions about their moral standpoint might assume that the interviewer wants to publicly shame them. Even though this might actually be the case, if it was not, the strategies tested in this dissertation might help to prevent such assumptions from occurring. Shifting the focus away from the moral domain (for example by focusing on competence) or communicating having goodwill when questioning interviewee's moral standpoints might help to decrease defensiveness from the interviewee and stimulate a more open-minded discussion on the differences in moral viewpoints and how they could possibly be integrated with each other.

5.9.3 | Diverse societies

The strategies tested in this dissertation may also have recommendations for how we can increase intergroup communication in societies that become increasingly diverse and where polarization and 'othering' become normalized (Finkel et al, 2020). To explain, living in diverse societies, we often meet people belonging to different groups in daily life. And we might discuss moral issues with them. For instance, we may discuss vegetarian meat options with a meat-eater at the supermarket. Or, we may discuss whether our children should take part in controversial traditions such as 'Zwarte piet' in the Netherlands at a school parent meeting. Using the strategies of this dissertation may help to decrease defensive reactions when discussing such 'hot' topics with people who hold different moral values. And this, in turn, will help to not further feed polarization and 'othering'.

5.9.4 | At the workplace

Finally, the findings of this dissertation may also help outgroup members get their voices heard at the workplace. Here, it is especially important to try

to increase the effectiveness of criticism from outgroups members. Without feedback from outgroups, an ingroup culture can easily develop which may be defective when ingroup members only become to rely on (questionable) ingroup norms and do not listen to negative feedback from the outgroup in fear of being ostracized (e.g., Cavazza et al., 2014; Ellemers & van Nunspeet, 2020; Thau et al., 2015; Van der Lee et al., 2017). Group members may engage in morally questionable and illegal activities because in their group such behaviors have become normalized and are not questioned by the members of the group. A recent example comes from the accounting ‘Wirecard scandal’ that occurred in 2020. The German payment processing company Wirecard AG had engaged in fraudulent accounting activities to inflate its profit resulting in €1.9 billion “missing” in the balances of the company. Even though the fast rise of the company was questioned by the Financial Times already in 2015 (McCrum, 2015), these allegations were dismissed by the company. Even further, the company started deflecting the blame back to the financial paper. They sued the paper and accused the author of the article of market manipulation. In this example, outgroup criticism wasn’t effective. It took 5 more years for the fraudulent behavior to come to light. Reports and newspaper articles (but also employees or spouses) that started questioning the company could have also made use of the strategies tested in this paper. They could have emphasized their intention to help with getting rid of irregularities in the balances rather than making accusations about potential fraudulent behavior. However, since this example is rather extreme and the actors consciously engaged in fraudulent behavior over years, it is likely that no type of criticism would have prevented this scandal. The strategies could have maybe helped in the beginning phase, though, when the first questionable decisions were made by the actors involved.

After discussing how the empirical findings of this dissertation address and may help with societal issues such as intergroup conflict, in the final two sections of this dissertation I will discuss potential avenues for future research, limitations, and conclude with some final words.

5.10 | Avenues for future research and limitations

The first point I would like to discuss is related to the operationalization of

'defensiveness'. Defensiveness is a key concept when investigating criticism situations and therefore a focal point of the research in this dissertation as well. However, the operationalization of defensiveness and how it can be measured in an experiment is not so straightforward. Often 'defensiveness' refers to a variety of different psychological processes. For example, Oxford languages (i.e., Google's English dictionary which is based on Oxford's English dictionaries) defines it as "the quality of being anxious to challenge or avoid criticism" or "behavior intended to defend or protect" (Lexico, n.d.). These definitions include three psychological processes: 1) Motivational action tendencies such as avoidance, 2) emotional responses such as anxiety, and 3) behavioral and cognitive reactions of defending or protecting the self (e.g., making negative attributions about a feedback sender, challenging criticism). To measure participants' defensiveness, I made use of an approach from previous research on intergroup criticism (Hornsey et al., 2004). Participants were asked to evaluate the person who criticized them and the critical message on several items (e.g., "The criticism was intended to be constructive"). I used the differences between the scores of these items for the evaluation of ingroup and outgroup critics as an indicator of defensiveness. Thus, defensiveness was defined as making more negative attributions about the critic and the criticism. Such cognitive attributions, however, merely capture the cognitive responses of a potential defensive response and they do not capture motivational (e.g., avoidance) or behavioral responses (e.g., defending the self). The ERP and behavioral findings of Study 2.2 offer another way to measure defensiveness towards negative feedback that gets a bit more at these two additional parts of defensive responses. That is, the higher-level ERPs (i.e., the P300 and LPP) can give insight into whether people pay sustained attention to the processing of negative feedback messages. Decreased amplitudes may indicate that people see feedback on their morality as less self-relevant than on their competence. Since this was complemented by reporting fewer, and especially fewer negative, feedback messages participants received on their morality (vs. competence), we can speculate whether these two measures are an additional way to measure defensiveness. That is, because it can be seen as a way to (not) deal with negative feedback by disengaging or

evaluating the feedback as not relevant to the self. To sum up, measuring defensiveness remains difficult, however, combining different types of measures as in the current dissertation might offer a way to understand the different aspects of defensive reactions.

A second point I would like to discuss builds on this difficulty to measure defensiveness and is related to how participants responded to receiving criticism from ingroup and outgroup members across chapters. Interestingly, the findings on how people emotionally experienced, or how defensively they reacted to, criticism showed an opposite pattern in Chapter 3 as compared to Chapters 2 and 4. In Chapter 2, people reported a stronger negative emotional response and they made more negative assumptions about the critic and the critical message (i.e., they responded more defensively) when they recalled a situation in which they had been criticized by outgroup compared to ingroup members. In Chapter 3, where participants received real-time feedback (i.e., in the actual moment) on their character in an experimental setting, this finding was reversed. Here, participants reported more negative emotions and reacted more defensively when they were confronted with the overall negative criticism they had received from the ingroup compared to the outgroup. In Chapter 4, participants again received real-time feedback, but were now asked to reflect on each negative judgement they received from an individual in- or outgroup member. Here, participants reported to perceive negative evaluations of their morality as less fair when these were given by outgroup (vs. ingroup) members, indicating more defensiveness towards feedback from outgroup members. To understand these differences, it helps to take a closer look at the differences in the experimental designs of the studies and at the potential psychological experience of the participants in these studies.

In Chapter 2, participants reflected on past autobiographical situations in which they had been criticized and they were asked to recall how they emotionally responded to the situation in the past. In Chapter 3, participants received the criticism in the actual moment in an experimental paradigm in the form of one overall negative evaluation of their character (e.g., being an immoral person) that was allegedly the sum of judgments from all ingroup

(vs. outgroup) members they had received judgments from in the task. In Chapter 4, participants received many negative judgments regarding their character (i.e., on each trial one) after they had made a 'proself' donation decision in the task and were asked after each judgment to indicate how fair they perceived the judgment to be.

The first explanation is related to the difference in experimental and social settings across chapters. It is possible that people simply have a stronger negative emotional and defensive response to receiving negative feedback when they receive an overall negative evaluation of their character coming from the whole ingroup (vs. outgroup, Chapter 3) compared to when they reflect on and receive negative evaluations from individual ingroup (vs. outgroup) member (Chapter 2 and 4). It would be interesting to directly test this explanation with further research. One could, for example, combine the experimental designs of Chapters 3 and 4. This could look as follows: Participants make trial-by-trial donation decisions as in the studies of Chapter 4 and are asked to rate the perceived fairness of the feedback they are receiving. However, the feedback messages should have an equal amount of positive and negative feedback, as in the studies of Chapter 3. Then, also as in Chapter 3, participants should receive an overall negative evaluation from the ingroup and the outgroup. Such an experimental design could give insight into whether participants respond more defensively towards feedback messages coming from individual outgroup (vs. ingroup) members in the experimental task, but more defensively towards an overall negative evaluation by the ingroup (vs. outgroup) after the experimental task.

The second explanation is related to what type of psychological processes are actually picked up on by the experimental measures across chapters. Since we asked participants in Chapter 2 to reflect on previous criticism situations, it is, for example, possible that they have cognitively re-evaluated the criticism situation. They might have 'forgiven' ingroup (but not outgroup) members for delivering criticism or they may have realized that ingroup critics had good and helpful motives to deliver criticism (e.g., wanting to help). The findings of Chapter 3 may reflect that participants were more hurt and especially negatively impacted by negative judgments of their character when such

judgments were delivered by fellow ingroup (vs. outgroup) members. And finally, participants in Chapter 4 might have focused more on the cognitive part of defensiveness (i.e., making negative cognitive attributions) making them perceive negative feedback on their morality as less fair when outgroup members delivered such feedback (and no intentions were communicated). Importantly, all these different psychological processes hint at the overall conclusion that feedback from the outgroup is less effective and less impactful than feedback from the ingroup. In conclusion, even though the differing findings initially seem contradictory to previous research, they are in line with previous research showing that people care more and are more influenced by the behavior and feedback from ingroup (vs. outgroup) members (e.g., Ellemers, 2012; Ellemers & van den Bos, 2012; Esposito et al., 2013; Hornsey et al., 2002; Hornsey & Imani, 2004; Tajfel & Turner, 1979).

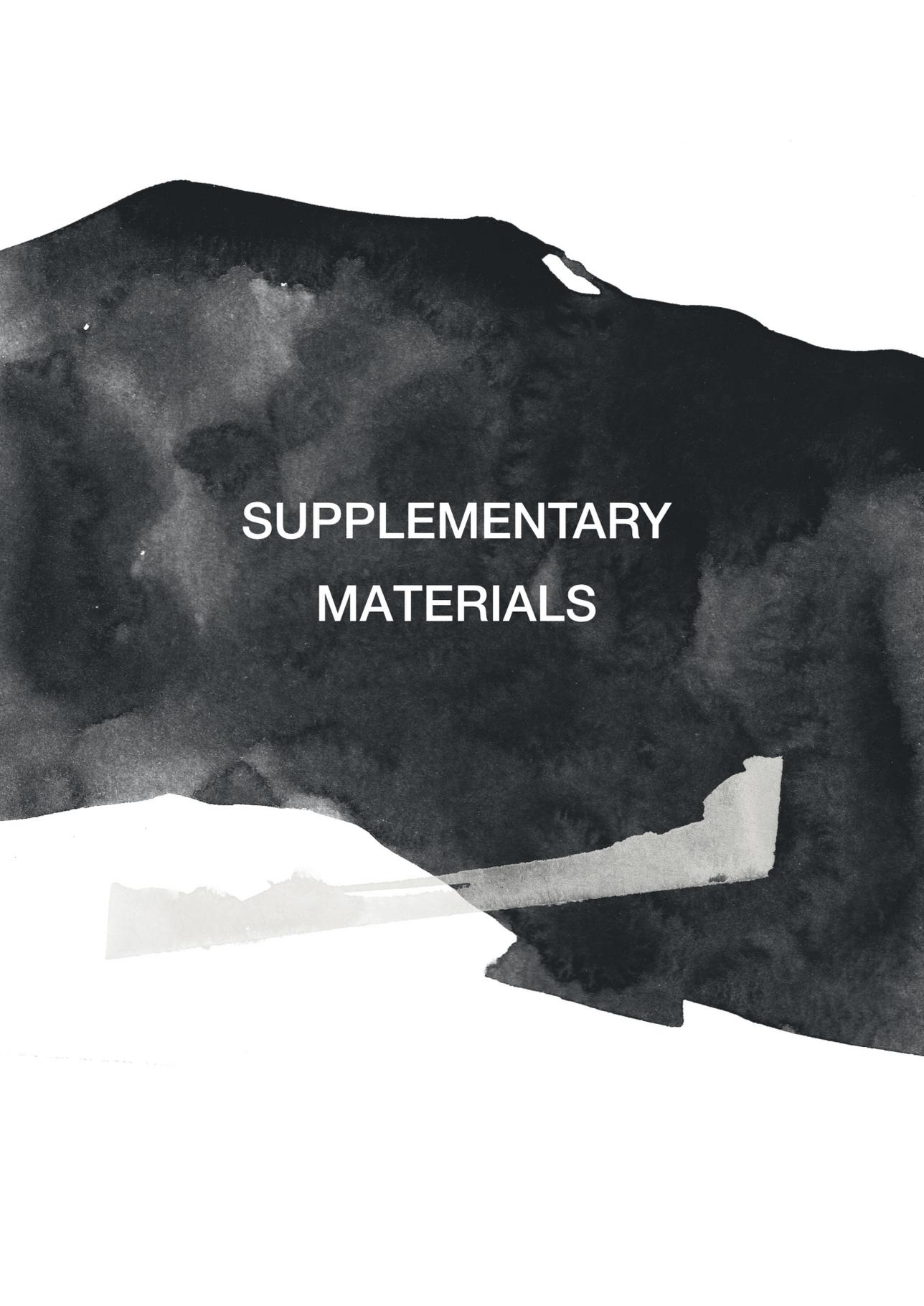
A final point worth noting is related to how the social group-membership of feedback sources was manipulated across chapters. When aiming to make general claims about group effects, it is important to demonstrate the consistency of (the direction of) an effect using different social groups. Therefore, the specific group-memberships that specified a feedback source as an ingroup or outgroup member in the given situation varied between chapters. In Chapter 2, participants reflected on criticism coming from real ingroup and outgroup members at the workplace such as a colleague from their own or from a different department. In Chapter 3, group-memberships were manipulated using groups that participants interacted with in real life (e.g., Dutch vs. international students) and in Chapter 4 by using a minimal group paradigm (i.e., other participants presented as being classified with a similar or different personality type as the participant). It is important to note that each of these manipulations brings their own benefits and drawbacks with it. To demonstrate, one could imagine that real groups at the workplace are more interdependent and that instrumental factors such as having a common goal to make money for the company might influence interpersonal experiences more strongly (e.g., Van Lange & Balliet, 2014) than when participants 'interact' with members from minimal groups in an experimental setting (Ellemers & Haslam, 2011; Tajfel & Turner, 1979). Moreover, participants share a superordinate identity with members from workplace

groups which is the company (Hornsey & Hogg, 2000). These two factors could influence participants' responses by making them more receptive to outgroup feedback at the workplace than for situations where groups do not share a superordinate identity such as in the lab. However, this does not seem to be the case. Participants were still less receptive of outgroup over ingroup criticism. In a similar fashion, one might argue that when using groups such as national (i.e., Dutch) and international students (the group-membership manipulation applied in Study 2.2) participants' responses could be influenced by prior stereotypical beliefs about these groups. In fact, the research in Study 2.2 was conducted in the lab at Utrecht University at a time where students of the international community reported to face discrimination and structural inequalities on the housing market (Fang & van Liempt, 2020). Using these real groups might be influenced by individual stereotypical beliefs, but the benefit of using real groups is that the results have greater external validity than when using, for instance, minimal groups. To sum up, each way to manipulate group-memberships of feedback sources has their unique potential drawbacks. But the benefit of using different types of manipulations outweigh these individual limitations. That is, with using different group-memberships we receive greater robustness of findings across settings and may make assumptions about a general group effect.

5.11 | Conclusion

The research in this dissertation addressed two questions: 1) How do people deal with being criticized for their moral behavior and how can we make people more open to such criticism? and 2) Do people react differently to criticism if it is delivered by an outgroup compared to an ingroup member – and if so – how can we intervene to prevent this from happening? By combining methods from different fields to investigate these questions, the research in this dissertation extends previous research that has examined such feedback situations solely from a social psychological perspective (Esposito et al., 2013; Gausel et al., 2012; Gausel & Leach, 2011; Hornsey & Imani, 2004; Täuber & van Zomeren, 2013; Van der Lee et al., 2016) or from a neuroscientific perspective (e.g., Schindler et al., 2014, 2015, 2020). Across three empirical chapters, the findings show that being criticized by an outgroup (vs. ingroup) member and being criticized on one's morality is often

ineffective. However, the findings also show that there are two strategies that can be used to increase this effectiveness. That is, outgroup critics should either make a person's incompetence (rather than their immorality) the focal point of a critical message, or they should communicate their helpful intentions. These two strategies offer two novel ways to create an open mind towards moral criticism.

The image features a dark, textured background that resembles torn paper or a watercolor wash. The color is a deep, mottled black or dark charcoal, with some lighter, greyish areas where the texture is more pronounced. The edges of the dark area are irregular and jagged, suggesting it was torn from a white surface. In the center of this dark area, the words "SUPPLEMENTARY" and "MATERIALS" are printed in a clean, white, sans-serif font. The text is centered horizontally and vertically, with "SUPPLEMENTARY" on the top line and "MATERIALS" on the bottom line. The overall composition is minimalist and modern.

**SUPPLEMENTARY
MATERIALS**

Supplementary Materials – Chapter 2

Study 1.1

Measures. The following items were excluded from our measures since they did not load on any of the factors (factor loadings $\leq .47$): “The criticism I received changed my feelings for this person”, “The person who criticized me did this in her/ his own best interest”, “I felt the criticism had to do with something outside of my control, something I am unable to influence”, “I perceived the criticism as threatening”, “I felt attacked by the criticism”, “The criticism made me feel quiescent”. These items are therefore also not included in the factor analyses tables which show the final solutions for each scale.

Sample. Participants were employed in a broad spectrum of career sectors (e.g., 10.2% information technology, 9.2% education & training, 6.1% business management administration, 4.6% finance, 3.1% arts), holding different types of industry roles (e.g., 14.3% researcher, 9.7% trained professional, 8.2% administrative staff, 7.7% middle management, 6.1% supporting staff, 4.1% upper management). No other career sector or industry role dominated in the sample. More than half of the participants indicated their nationality as being from the UK (58.3%). Other reported nationalities included the US (9.6%), Italy (6.4%), and Portugal (3.7%).

Checks. Participants also reported higher collective self-esteem for the ingroup ($M = 5.90$, $SD = 1.58$) compared to the outgroup ($M = 5.10$, $SD = 1.66$), $t(195) = 5.57$, $p < .001$.

Alternative explanations. To be able to exclude fear of exclusion from the ingroup of the participant as an alternative explanation we investigated the fear of exclusion with a RM ANOVA, specifying the critic’s group-membership as a within-participant factor and the dimension of the criticism (i.e., morality vs. competence) as a between-participant factor. We did not find any effects for fear of exclusion, all $F(1, 195) < 3.35$, $p > .069$, and, thus, fear of exclusion can be ruled out as an alternative explanation.

Additionally, we analyzed the proportions of chosen answer options from the guided questions for the critic’s group-membership and dimension. There were no significant differences between the proportions of whether

participants liked the critic and the dimension of the criticism, neither for the ingroup: $\chi^2(2, N = 196) = .59, p = .746$, nor for the outgroup condition: $\chi^2(2, N = 196) = 3.53, p = .171$. Thus, any effects of dimension could not be explained by the interpersonal liking of the critic. An exact McNemar's test revealed that there was a significant association between group-membership and liking of the critic, $\chi^2(2, N = 196) = 43.14, p < .001$. Participants liked the critic more in the ingroup condition (71.4%) compared to the outgroup condition (40.8%) as well as dislike the critic more in the outgroup condition (21.4%) compared to the ingroup condition (34.2%). Since we did not have enough cases to investigate this further with statistical tests, we inspected the mean patterns of critics liked and critics not liked. We found the effect of group-membership in our predicted direction on our main dependent variable effectiveness (i.e., ingroup criticism rated more effective, also see Table A). Thus, any effects of the critic's group-membership did not seem to be caused by differences in interpersonal liking.

There were no significant differences for the two dimensions of the criticism and the status of the critic, neither for the ingroup: $\chi^2(2, N = 196) = .63, p = .732$, nor for the outgroup condition: $\chi^2(2, N = 196) = .32, p = .851$. Thus, any dimension effects cannot be attributed to a difference in status of the critic. There was, however, a significant association between the status and the critic's group-membership, $\chi^2(3, N = 196) = 9.18, p = .027$. Outgroup critics seemed to have had a higher status (39.8% high status) than ingroup critics (26% high status). Again, inspecting the means (see Table A) showed that the group-membership effect was evident for all critics with the same and higher status. For critics with lower status, there was no difference (but also note that there are much fewer cases in this cell compared to the other two status options). Moreover, since people would probably be more likely to accept criticism from higher status critics, the fact that outgroup critics had higher status should make participants rate outgroup criticism as more positive, which goes against our predictions. Thus, the effect of group-membership in the directions of our predictions would even be stronger for ingroup vs. outgroup critics with a similar status.

To further assess whether there were differences between conditions on how participants perceived how the criticism was delivered by the critic,

we checked whether the form of delivery varied between our experimental conditions. There were no significant differences between the dimension of the criticism and how the criticism was delivered, neither for the ingroup: $\chi^2(4, N = 196) = 1.92, p = .751$, nor for the outgroup condition: $\chi^2(4, N = 196) = 5.00, p = .288$. There was, however, a significant association between the critic's group-membership and the delivery of the criticism, $\chi^2(9, N = 196) = 26.48, p = .002$. Participants more often received criticism that was "Quick, clear, and friendly" from an ingroup member (40.3%) compared to an outgroup member (19.4%), and more unfriendly criticism from an outgroup (55.6%) compared to an ingroup member (41.8%). To address this, we performed additional analyses. Delivery of criticism was measured using 5 answer options (i.e., 1 = Quick, clear, and friendly; 2 = Quick and clear – but unfriendly; 3 = Took too long – the person was talking on and on; 4 = Not long enough – I would have liked to hear about what I could improve according to this person; 5 = Other [please specify]). We used this variable as a covariate (time-varying = measured for each of the two experiences) in a linear mixed model, specifying group-membership of critic as a within-participant factor, effectiveness, negative attributions, and motivation to improve as dependent variables, and delivery of criticism as a (time-varying) covariate (i.e., in three different analyses).

Negative attributions. Participants made more negative attributions when receiving criticism from an outgroup ($EMM = 4.70, SE = .11$) than from an ingroup member ($EMM = 4.03, SE = .11$), $F = 17.74, p < .001$, even after controlling for delivery of criticism.

Motivation to improve. Participants were more motivated to improve when receiving criticism from an ingroup ($EMM = 3.38, SE = .13$) than from an outgroup member ($EMM = 2.92, SE = .13$), $F = 7.95, p = .005$, even after controlling for delivery of criticism.

Effectiveness. Criticism was more somewhat more effective when receiving criticism from an ingroup ($EMM = 3.70, SE = .14$) than from an outgroup member ($EMM = 3.34, SE = .14$), $F = 3.51, p = .062$, even after controlling for delivery of criticism.

To summarize, ingroup criticism is still more effective compared to outgroup criticism across our main dependent variables when controlling for delivery of criticism.

To check for comparability across situations we checked whether the context and time of the criticism situations varied between the dimension and critic's group-membership conditions. There was a significant difference in what context the criticism was given between the morality and competence condition for the ingroup condition, $\chi^2(3, N = 196) = 8.22, p = .042$. However, there were no differences between the two dimensions in the outgroup condition, $\chi^2(3, N = 196) = 6.72, p = .081$ and, as intended, most criticism was offered spontaneously (72.4% of all answers) making the situations comparable. There were no significant differences between the proportions of the context of the situations and the critic's group-membership, $\chi^2(5, N = 196) = 4.09, p = .536$. Thus, the context of the situations was comparable between our conditions.

There was no significant difference between the time that reported situations took place and the dimension of the criticism (i.e., morality, competence), neither for the ingroup: $\chi^2(7, N = 196) = 5.93, p = .548$, nor for outgroup condition: $\chi^2(7, N = 196) = 6.80, p = .450$ ¹⁷. Most participants reported on situations that were one to three months ago (ingroup: 38.8% of all answers, outgroup: 44.9%). Thus, the time that situations occurred did not differ systematically with our conditions, which makes them comparable.

Finally, we compared the means from both the ingroup and outgroup condition on our bipolar check on whether the criticism was given on behavior or identity (1 = behavior, 7 = identity) to the midpoint of the scale (4) with one-sample t-tests. As intended, participants reported that criticism was given on their behavior (rather than identity) in both the ingroup condition, ($M = 2.65, SD = 1.80$), $t(195) = -10.46, p < .001$, and the outgroup condition ($M = 2.70, SD = 1.72$), $t(195) = -10.61, p < .001$. Independent sample t-tests revealed that there was no significant difference between the morality and the competence (condition in whether the criticism was given on behavior vs. identity for neither the ingroup condition (competence: $M = 2.46, SD = 1.74$, morality: $M = 2.87, SD = 1.86$), $t(194) = -1.61, p = .108$, nor the outgroup condition, (competence: $M = 2.52, SD = 1.54$, morality: $M = 2.89, SD = 1.89$), $t(194) = -1.50, p = .135$).

¹⁷ Too few cases for McNemar-Bowker test to test whether there are significant differences for group-membership

Tables

Table S2.1

Alternative explanations: Cases, Means, and Standard Deviation for Effectiveness by Liking (excl. 'other' option), Status, and Delivery of Criticism (Study 1)

		<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
		Ingroup			Outgroup		
Liking	Liked	140	4.45	1.90	80	3.88	2.07
	Disliked	42	2.67	1.66	67	2.33	1.67
Status	Lower	29	2.47	1.80	25	2.47	1.75
	Equal	93	4.21	2.02	93	3.00	1.96
	Higher	78	4.12	2.02	78	3.77	2.08
Delivery	Quick, clear, and friendly	79	4.89	1.78	38	5.08 ¹⁸	1.81
	Quick and clear – but unfriendly	82	3.12	3.12	109	2.85	1.79
	Took too long – the person was talking on and on	20	3.42	3.42	33	2.70	1.97
	Not long enough – I would have liked to hear about what I could improve according to this person	13	4.38	4.38	7	3.17	2.13

¹⁸ It might be the case that outgroup criticism is more effective than ingroup criticism when it is given in a friendly manner, however the mean difference is very small.

Table S2.2

Means, Standard Deviation, and Reliability for Study 1 ($N = 196$)

Variables	<i>M</i>	<i>SD</i>	Rel
Negative attributions	3.90/	1.78/	(.96/
	4.82	1.62	.94)
Negative emotions	3.82/	1.60/	(.91/
	4.27	1.52	.89)
Positive emotions	2.37/	1.39/	(.83/
	2.10	1.3	.85)
Moral emotions	3.47/	1.97/	(.70/
	3.51	2.00	.73***)
Motivation to improve	3.46/	1.95/	(.92/
	2.83	1.72	.91)
Effectiveness	3.93/	2.07/	(.97/
	3.24	2.03	.97)

Note: ** $p < .01$, *** $p < .001$. Ingroup/ Outgroup condition, lowest significance of both conditions presented. Cronbach's alphas are shown in the diagonal (correlation for moral emotions, 2 items).

Table S2.3

Pearson Correlation Matrix for Study 1 ($N = 196$)

	1	2	3	4	5	6	7	8	9	10	11
Negative attributions _i											
Negative attributions _o	.02										
Motivation to improve _i	-.69**	-.08									
Motivation to improve _o	.02	-.66**	.19**								
Effectiveness _i	-.72**	-.06	.83**	.18*							
Effectiveness _o	-.00	-.73**	.06	.77**	.10						
Positive emotions _i	-.52**	-.17*	.67**	.22**	.53**	.15*					
Positive emotions _o	-.07	-.59**	.25**	.70**	.18**	.56**	.38**				
Negative emotions _i	.30**	.03	-.26**	.10	-.19**	.07	-.44**	-.07			
Negative emotions _o	-.10	.32**	.03	-.20**	.09	-.16*	-.07	-.47**	.40**		
Moral emotions _i	-.32**	.02	.32**	.05	.44**	.03	-.06	-.09	.49**	.35**	
Moral emotions _o	-.17*	-.21**	.13	.29**	.17*	.35**	.05	-.05	.28**	.48**	.43**

Note: ** $p < .01$, *** $p < .001$. i = ingroup, o = outgroup

Table S2.4

Correlation for three parts included in our main dependent variable effectiveness: Means, Standard Deviation and Pearson Correlation Matrix for Acceptance, Intention to Change, and Actual Behavioral Change, Study 1 ($N = 196$)

Variables	1	2	3
Acceptance	.83/		
Intention to Change	.87***		
	.94/	.81/	
Actual Behavioral Change	.93***	.80***	

Note: *** $p < .001$. Ingroup/ Outgroup condition. Lowest significance of both conditions presented.

Table S2.5

Summary of Exploratory Factor Analysis Results for Negative Attributions (Study 1) Using Maximum Likelihood Estimation and Varimax Rotation ($N = 196$)

Item	Factor Loadings
	Negative Attributions
I trusted the person who criticized me before I received the criticism	.76/ .69
I now trust the person	.90/ .87
I think the person who criticized me did this in my best interest	.84/ .89
I think the person who criticized me cares about me	.89/ .88
I think the person is credible	.92/ .88
I think the person who criticized me has the expertise to do this	.82/ .77
The criticism was intended to be constructive	.82/ .73
The criticism was intended to be destructive	.81/ .71
The criticism was credible	.78/ .78

Note: Ingroup/ Outgroup condition. Positive items of negative attributions were reverse coded.

Table S2.6

Summary of Exploratory Factor Analysis Results for Emotional Response (Study 1) Using Maximum Likelihood Estimation and Varimax Rotation ($N = 196$)

Item	Factor Loadings		
	Negative Emotions	Positive Emotions	Moral Emotions
The criticism made me feel dejected	.60/ .49	-.28/ -	.29/ .41
The criticism upset me	.77/ .65	-.35/ -.39	.22/ -
The criticism made me feel discouraged	.60/ .57	-.30/ -	.32/ .32
The criticism made me feel fearful	.54/ .53	-.09/ -	.31/ .39
The criticism made me feel disappointed	.64/ .65	-.23/ -	.33/ -
The criticism made me feel tense	.74/ .77	-.10/ -	.16/ -
The criticism made me feel agitated	.76/ .61	-.12/ -	-.01/ -
I felt distressed about the criticism	.79/ .78	-.12/ -	.23/ -
The criticism made me feel excited	-.05/ -	.78/ .72	.06/ -
The criticism made me feel relaxed	-.28/ -	.67/ .80	-.06/ -
I was happy to have received the criticism	-.36/ -.30	.61/ .76	.09/ -
The criticism made me feel confident	-.11/ -	.86/ .66	-.13/ -
I felt guilty when I was criticized in this way	.22/ -	.11/ -	.73/ .97
I felt ashamed about the criticism I received	.31/ -	-.07/ -	.85/ .70

Note: Ingroup/ Outgroup condition. No loadings: -. Loadings in bold and are included in the final scale.

Table S2.7

Summary of Exploratory Factor Analysis Results for Effectiveness (Study 1) Using Maximum Likelihood Estimation and Varimax Rotation ($N = 196$)

	Factor Loadings
	Effectiveness
I accepted the criticism	.74/ .75
I took the criticism into consideration	.80/ .80
I agreed with the criticism	.76/ .81
I intended to act upon the criticism by changing my behavior or to do or say things differently in the future	.92/ .93
I expect that I will use the criticism to change my behavior	.95/ .96
I planned to use the criticism to improve	.94/ .88
I actually acted upon the criticism by changing my behavior or to do or say things differently in the future	.94/ .95
I used the criticism to change something about my behavior	.95/ .96
I changed my behavior based on what was criticized	.95/ .95

Note: Ingroup/ Outgroup condition.

Table S2.7

Summary of Exploratory Factor Analysis Results for Effectiveness and Motivation to Improve (Study 1) Using Maximum Likelihood Estimation and Varimax Rotation ($N = 196$)

Item	Factor Loadings	
	Effectiveness	Motivation to improve
I accepted the criticism	.51/ .50	.59/ .63
I took the criticism into consideration	.59/ .58	.57/ .58
I agreed with the criticism	.56/ .64	.54/ .53
I intended to act upon the criticism by changing my behavior or to do or say things differently in the future	.76/ .78	.53/ .50
I expect that I will use the criticism to change my behavior	.79/ .81	.52/ .51
I planned to use the criticism to improve	.78/ .68	.52/ .59
I actually acted upon the criticism by changing my behavior or to do or say things differently in the future	.85/ .86	.43/ .41
I used the criticism to change something about my behavior	.87/ .92	.41/ .35
I changed my behavior based on what was criticized	.83/ .89	.46/ .36
The criticism made me feel energized about improving myself	.38/ .31	.75/ .80
I was interested to find out more about how to improve	.43/ .37	.79/ .80
I felt I could do something with the criticism	.39/ .32	.75/ .72
The criticism inspired me to do something about it	.44/ .40	.72/ .67

Note: Ingroup/ Outgroup condition. Loadings in bold and are included in the final scale.

Table S2.8

Summary of Exploratory Factor Analysis Results for Effectiveness and Negative Attributions (Study 1) Using Maximum Likelihood Estimation and Varimax Rotation ($N = 196$)

Item	Factor Loadings	
	Effectiveness	Negative attributions
I accepted the criticism	.62/ .59	-.49/ -.55
I took the criticism into consideration	.69/ .66	-.47/ -.50
I agreed with the criticism	.65/ .68	-.48/ -.52
I intended to act upon the criticism by changing my behavior or to do or say things differently in the future	.87/ .86	-/ -.35
I expect that I will use the criticism to change my behavior	.90/ .89	-.31/ -.35
I planned to use the criticism to improve	.89/ .79	-/ -.38
I actually acted upon the criticism by changing my behavior or to do or say things differently in the future	.90/ .91	-/ -
I used the criticism to change something about my behavior	.91/ .93	-/ -
I changed my behavior based on what was criticized	.90/ .92	-.31/ -
I trusted the person who criticized me before I received the criticism	-/ -	.74/ .65
I now trust the person	-/ -	.88/ .83
I think the person who criticized me did this in my best interest	-.37/ -.35	.75/ .82
I think the person who criticized me cares about me	-/ -	.86/ .85
I think the person is credible	-.30/ -.35	.87/ .80
I think the person who criticized me has the expertise to do this	-.44/ -.39	.71/ .69
The criticism was intended to be constructive	-.50/ -.39	.69/ .64
The criticism was intended to be destructive	-.35/ -	.74/ .67
The criticism was credible	-.52/ -.55	.65/ .64

Note: Ingroup/ Outgroup condition. Positive items of negative attributions were reverse coded No loadings: -. Loadings in bold and are included in the final scale.

Figures

Figures. There was some overlap between the acceptance part (three items) of our main dependent variable effectiveness and motivation to improve (see Table S2.7). We therefore ran the mediation model without the acceptance items. Taking out these items improves the distinction of these two factors but does not change mediation (see Figure S2.1).

Further, we include (non-significant) mediation effects for the between-participants factors (i.e., dimension of criticism, morality vs. competence) on effectiveness to this file (see Figure S2.2 and Figure S2.3). The results nicely demonstrate that, even though not always significant, all paths have the same direction as in the complementing within-factors models, providing more evidence for the consistency of our results across both studies (note: As we asked participants to reflect on two criticism situations per study, we included the mediation analyses separately for the two within factors (i.e., Study 1 = ingroup vs. outgroup criticism).

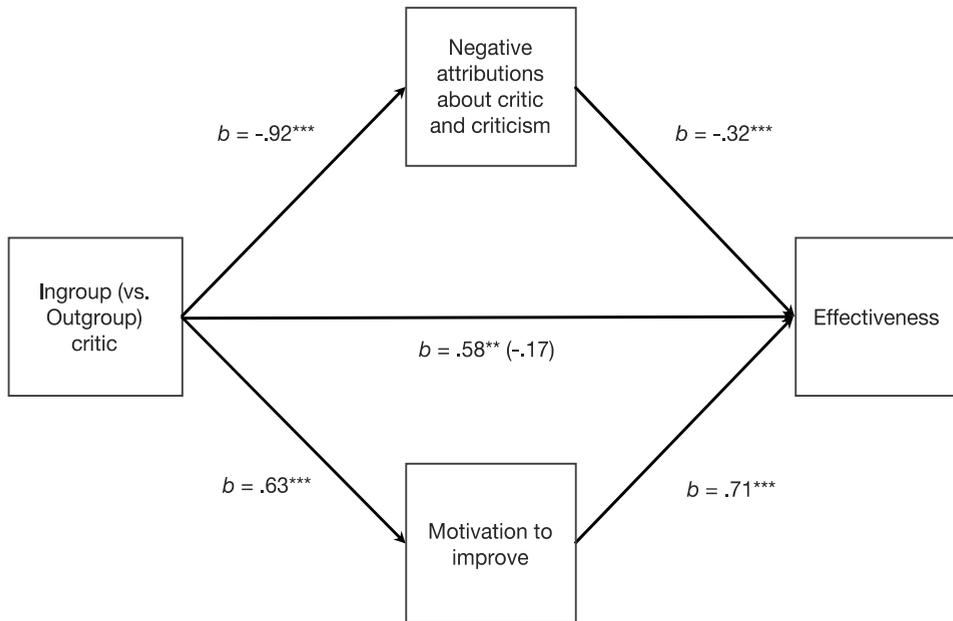


Figure S2.1. Mediation model Study 1 with unstandardized regression coefficients (b) depicting the relationship between the within-participants factor ‘the critic’s group-membership’ and ‘effectiveness of the criticism’ (without acceptance), as mediated by ‘negative attributions the about critic and ‘motivation to improve’. The unstandardized regression coefficient for the direct effect of ‘group-membership’ on ‘effectiveness of criticism’ is in parentheses.

Note: $*p < .05$, $***p < .001$

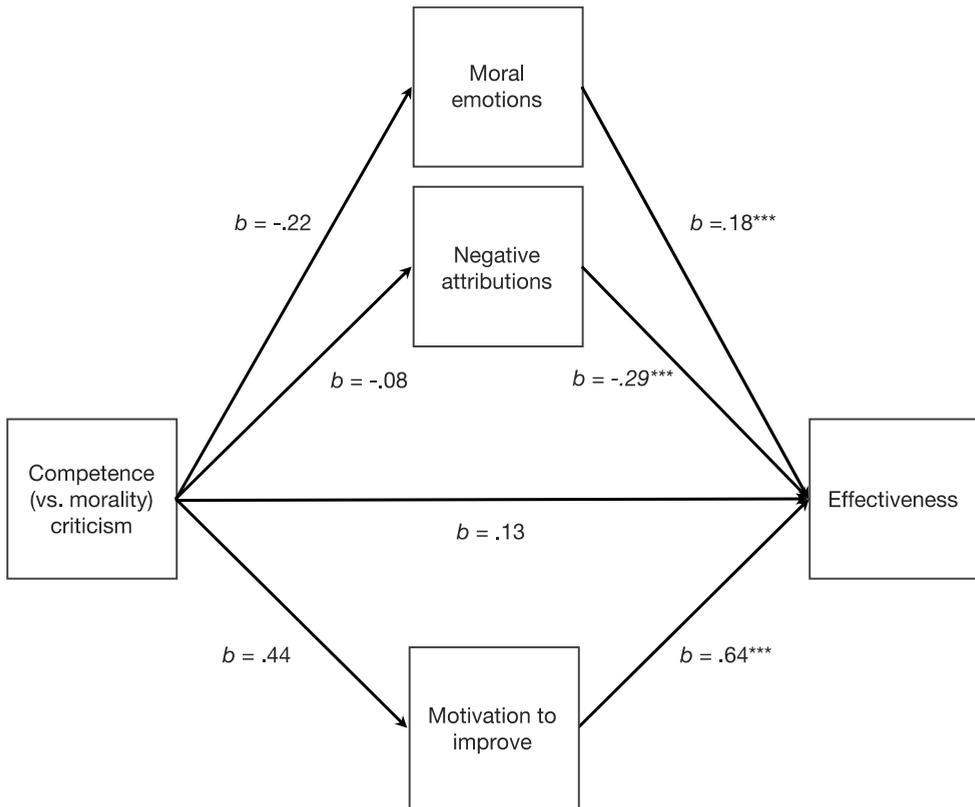


Figure S2.2. (Non-significant) mediation model Study 1 for the ingroup condition with unstandardized regression coefficients (b) depicting the relationship between the within-participants factor 'dimension of criticism (i.e., competence vs. morality)' and 'effectiveness of the criticism', with (non-significant) mediators 'negative attributions the about critic', 'moral emotions', and 'motivation to improve'.

Note: * $p < .05$, *** $p < .001$

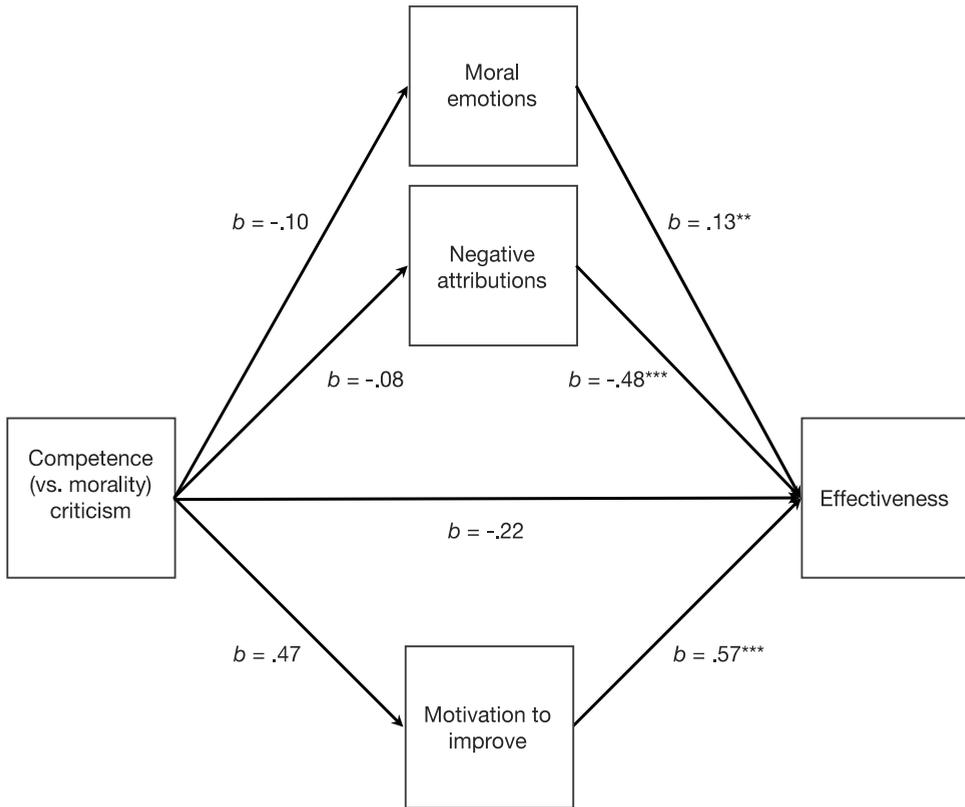


Figure S2.3. (Non-significant) mediation model Study 1 for the outgroup condition with unstandardized regression coefficients (b) depicting the relationship between the within-participants factor ‘dimension of criticism (i.e., competence vs. morality)’ and ‘effectiveness of the criticism’, with (non-significant) mediators ‘negative attributions the about critic’, ‘moral emotions’, and ‘motivation to improve’.

Note: $*p < .05$, $***p < .001$

Study 1.2

Measures. The following items were excluded from our measures since they loaded on several factors: “I perceived the criticism as threatening”, “I felt attacked by the criticism”, “I feel disgraced when I think about being criticized this way”, “I feel humiliated when I think about being criticized this way”. These items are therefore also not included in the factor analyses tables which show the final solutions for each scale.

Additional measures. The second study also included items related to group attachment (adapted from Prentice et al., 1994, e.g. “How important is belonging to this group to you?”) and perception of the criticism (adapted from Esposito et al., 2013, e.g., “The criticism was disappointing”). However, these were part of a student research project (i.e., master thesis) and were not part of the current research.

Sample and groups of critics. As in the previous study, participants indicated to work in different career sectors (e.g., 9.2% education & training, 6.8% retail, 5.6% government & public administration, 5% medicine), holding different positions in their workplace (e.g., 15% trained professional, 10.2 % middle management, 8.9% administrative staff, 6% self-employed/partner, 5.8% junior management). No other career sector or industry role dominated in the sample. Most of our participants indicated as their nationality being from the UK (78.7%), other reported nationalities included the US (13.6%), Canada (2.9%), and Ireland (1%).

Checks. Participants reported higher collective self-esteem for the ingroup ($M = 5.08$, $SD = 1.37$) compared to the outgroup ($M = 4.78$, $SD = 1.45$), $t(383) = 2.10$, $p = .036$, for the morality condition. For the competence condition however, participants reported only somewhat higher collective self-esteem in the ingroup condition ($M = 5.07$, $SD = 1.38$) compared to the outgroup condition ($M = 4.82$, $SD = 1.32$), $t(383) = 1.80$, $p = .073$.

Alternative explanations. To check for comparability between the situations in which criticism was provided, we investigated whether the time and context of the situations systematically varied between our conditions. There was no significant difference between the proportions of the time that the criticism situations took place and the critic’s group-membership

neither for the moral, $\chi^2(4, N = 385) = 5.62, p = .229$, nor for the competence condition, $\chi^2(4, N = 385) = 4.66, p = .324$. An exact McNemar's test yielded that there were no differences between the dimension and the time that the situations took place, $\chi^2(10, N = 385) = 11.16, p = .345$.

There were no significant differences between the context of criticism situations and the critic's group-membership, neither for the morality: $\chi^2(3, N = 385) = 4.24, p = .237$, nor for the competence condition: $\chi^2(3, N = 385) = 1.14, p = .767$. An exact McNemar's test determined that there was a significant association between the dimension and the situation of the criticism, $\chi^2(6, N = 385) = 20.11, p = .003$. However, in both conditions, feedback was most often offered spontaneously (morality: 81.3 %, competence: 72.2 %). Thus, the criticism situations were comparable.

To check whether our results of dimension (i.e., morality vs. competence) are confounded by concreteness/ abstractness of criticism we performed additional analyses. To check for this possibility, we compared means of our bipolar check on whether criticism was given on behavior or identity (1 = behavior, 7 = identity) to the midpoint of the scale (4) with one-sample t-tests. As intended, participants reported that criticism was given on their behavior (vs. identity) in both the morality, ($M = 3.30, SD = 2.01$), $t(384) = -6.85, p < .001$, and the competence condition, ($M = 2.79, SD = 1.81$), $t(384) = -13.09, p < .001$. Interestingly, there was a significant difference between the morality and the competence condition in whether the criticism was given on behavior vs. identity, $t(384) = 3.80, p < .001$. Even though overall most criticism was related to behavior (as intended), moral criticism seems to be a bit more about participant's identity compared to competence criticism. To check whether this might confound our reported effects, we performed additional analyses. We used the bipolar measure of behavior/identity as a covariate for the most important analyses from Study 2. As this was a time-varying covariate, meaning that we have one measure for each of the two levels of the within-participants measure, we performed these analyses with linear mixed models, specifying dimension as a within-participant factor and negative attributions about the critic, motivation to improve, and effectiveness of the criticism as dependent variables, and a behavior/identity as a (time-varying) covariate (i.e., in three different

analyses). Importantly, for all three main DVs, we find that all effects hold.

Negative attributions. Participants made more negative attributions when being criticized for their morality ($EMM = 4.70$, $SE = .08$) as compared to for their competence ($EMM = 4.41$, $SE = .08$), $F = 6.49$, $p = .011$, even after controlling for behavior/identity.

Motivation to improve. Participants were more motivated to improve when being criticized for their competence ($EMM = 3.11$, $SE = .09$) as compared for their morality ($EMM = 2.63$, $SE = .09$), $F = 19.47$, $p < .001$, even after controlling for a behavior/identity.

Effectiveness. Criticism was more effective when it was about competence ($EMM = 3.35$, $SE = .10$) as compared to about morality ($EMM = 2.93$, $SE = .10$), $F = 13.58$, $p < .001$, even after controlling for behavior/identity.

Taken together, we see that most criticism situations are about specific behavior and that the effect of dimension remains after controlling for behavior/identity. We can therefore rule this out as a potential confound for our results.

Tables

Table S2.9

Means, Standard deviation, and Reliability of Study 2 ($N = 385$)

Variables	<i>M</i>	<i>SD</i>	Cronbach's alpha
Negative attributions	4.73/	1.60/	(.94/
	4.34	1.71	.89)
Negative emotions	4.53/	1.57/	(.92/
	4.44	1.55	.92)
Positive emotions	1.76/	1.01/	(.76/
	1.93	1.02	.66)
Moral emotions	3.09/	1.75/	(.89/
	2.88	1.63	.89)
Motivation to improve	2.59/	1.69/	(.90/
	3.15	1.93	.92)
Effectiveness	2.86/	1.85/	(.97/
	3.42	1.97	.97)

Note. Morality/ Competence condition

Table S2.10

Pearson Correlation Matrix for Study 2 ($N = 385$)

	1	2	3	4	5	6	7	8	9	10	11
Negative Attributions _m											
Negative Attributions _c	.20**										
Motivation to improve _m	-.57**	-.04									
Motivation to improve _c	-.16**	-.64**	.28**								
Positive Emotions _m	-.44**	-.00	.55**	.19**							
Positive Emotions _c	-.08	-.44**	.19**	.58**	.33**						
Negative Emotions _m	.26**	.02	-.06	.06	-.39**	.06					
Negative Emotions _c	-.00	.34**	.08	-.26**	.03	-.35**	.40**				
Moral Emotions _m	-.16**	.01	.21**	.08	.06	.09	.47**	.33**			
Moral Emotions _c	-.08	-.06	.14**	.09	.15**	.01	.25**	.44**	.52**		
Effectiveness _m	-.69**	-.09	.78**	.20**	.50**	.11*	-.09	.13*	.32**	.17**	
Effectiveness _c	-.19**	-.76**	.20**	.82**	.15**	.51**	.03	-.25**	.09	.18**	.23**

Note: * $p < .05$, ** $p < .01$, *** $p < .001$. m = morality, c = competence

Table S2.11

Summary of Exploratory Factor Analysis Results for Negative Attributions (Study 2) Using Maximum Likelihood Estimation and Varimax Rotation ($N = 385$)

Item	Factor Loadings
	Negative Attributions
I trusted the person who criticized me before I received the criticism	.61/ .65
I now trust the person	.87/ .88
I think the person who criticized me did this in my best interest	.88/ .90
I think the person who criticized me cares about me	.93/ .89
I think the person is credible	.90/ .89
I think the person who criticized me has the expertise to do this	.80/ .85
The criticism was intended to be constructive	.73/ .80
The criticism was intended to be destructive	.65/ .70
The criticism was credible	.73/ .81

Note: Morality/ Competence condition. Positive items of negative attributions were reverse coded.

Table S2.12

Summary of Exploratory Factor Analysis Results for Emotional Response (Study 2) Using Maximum Likelihood Estimation and Varimax Rotation ($N = 385$)

Item	Factor Loadings		
	Negative Emotions	Moral Emotions	Positive Emotions
The criticism made me feel dejected	.74/ .67	.20/ .15	-.11/ -.16
The criticism upset me	.83/ .83	.14/ .15	-.24/ -.19
The criticism made me feel discouraged	.78/ .78	.20/ .17	-.13/ -.20
The criticism made me feel fearful	.60/ .60	.30/ .33	.08/ .13
The criticism made me feel disappointed	.73/ .74	.12/ .13	-.13/ -.14
The criticism made me feel tense	.79/ .78	.07/ .14	-.16/ -.08
The criticism made me feel agitated	.73/ .77	.02/ .08	-.23/ -.11
I felt distressed about the criticism	.81/ .84	.18/ .19	-.16/ -.11
The criticism made me feel excited	-.02/ -	-/ .05	.74/ .72
The criticism made me feel relaxed	-.19/ -.14	-.01/ .08	.90/ .77
The criticism made me feel confident	-.18/ -.18	-.05/ -.05	.62/ .54
I was happy to have received the criticism	-.20/ -.40	.24/ .13	.47/ .37
I feel ashamed when I think about being criticized this way	.52/ .47	.52/ .57	-.08/ -.07
I feel guilty when I think about being criticized this way	.24/ .17	.83/ .87	-.02/ -
Right now, thinking of being criticized this way, I feel guilt	.17/ .16	.94/ .93	.05/ .04
As a member of the group, right now I feel guilty	.17/ .12	.85/ .83	.10/ .13

Note: Morality/ Competence condition. No loadings: -. Loadings in bold and are included in the final scale.

Table S2.13

Summary of Exploratory Factor Analysis Results for Effectiveness (Study 2) Using Maximum Likelihood Estimation and Varimax Rotation ($N = 385$)

	Factor Loadings
	Effectiveness
I accepted the criticism	.74/ .77
I took the criticism into consideration	.79/ .81
I agreed with the criticism	.76/ .84
I intended to act upon the criticism by changing my behavior or to do or say things differently in the future	.95/ .92
I expect that I will use the criticism to change my behavior	.94/ .94
I planned to use the criticism to improve	.91/ .92
I actually acted upon the criticism by changing my behavior or to do or say things differently in the future	.94/ .93
I used the criticism to change something about my behavior	.92/ .94
I changed my behavior based on what was criticized	.93/ .95

Note: Morality/ Competence condition.

Table S2.14

Summary of Exploratory Factor Analysis Results for Effectiveness and Motivation to Improve (Study 2) Using Maximum Likelihood Estimation and Varimax Rotation ($N = 385$)

Item	Factor Loadings	
	Effectiveness	Motivation to improve
I accepted the criticism	.58/ .54	.47/ .58
I took the criticism into consideration	.63/ .58	.50/ .60
I agreed with the criticism	.61/ .65	.46/ .52
I intended to act upon the criticism by changing my behavior or to do or say things differently in the future	.86/ .75	.41/ .53
I expect that I will use the criticism to change my behavior	.82/ .79	.47/ .50
I planned to use the criticism to improve	.72/ .71	.60/ .59
I actually acted upon the criticism by changing my behavior or to do or say things differently in the future	.87/ .81	.37/ .46
I used the criticism to change something about my behavior	.84/ .87	.39/ .39
I changed my behavior based on what was criticized	.88/ .89	.33/ .39
The criticism made me feel energized about improving myself	.30/ .33	.78/ .80
I was interested to find out more about how to improve	.37/ .38	.80/ .78
I felt I could do something with the criticism	.49/ .53	.69/ .69
The criticism inspired me to do something about it	.35/ .44	.70/ .71

Note: Morality/ Competence condition. Loadings in bold and are included in the final scale.

Table S2.15

Summary of Exploratory Factor Analysis Results for Effectiveness and Negative Attributions (Study 2) Using Maximum Likelihood Estimation and Varimax Rotation ($N = 385$)

Item	Factor Loadings	
	Effectiveness	Negative attributions
I accepted the criticism	.59/ .62	-.58/ -.54
I took the criticism into consideration	.68/ .68	-.44/ -.50
I agreed with the criticism	.62/ .70	-.54/ -.53
I intended to act upon the criticism by changing my behavior or to do or say things differently in the future	.91/ .84	-.29/ -.37
I expect that I will use the criticism to change my behavior	.89/ .87	-.31/ -.35
I planned to use the criticism to improve	.84/ .83	-.35/ -.41
I actually acted upon the criticism by changing my behavior or to do or say things differently in the future	.91/ .89	-.25/ -.28
I used the criticism to change something about my behavior	.90/ .92	-.24/ -.25
I changed my behavior based on what was criticized	.89/ .91	-.27/ -.28
I trusted the person who criticized me before I received the criticism	-.21/ -.15	.56/ .65
I now trust the person	-.17/ -.25	.86/ .86
I think the person who criticized me did this in my best interest	-.35/ -.40	.82/ .80
I think the person who criticized me cares about me	-.22/ -.28	.91/ .84
I think the person is credible	-.27/ -.35	.85/ .81
I think the person who criticized me has the expertise to do this	-.32/ -.42	.73/ .74
The criticism was intended to be constructive	-.39/ -.50	.65/ .66
The criticism was intended to be destructive	-.21/ -.34	.62/ .61
The criticism was credible	-.49/ -.56	.62/ .68

Note: Morality/ Competence condition. Positive items of negative attributions were reverse coded. Loadings in bold and are included in the final scale.

Table S2.16

Summary of Exploratory Factor Analysis Results for Effectiveness and Moral Emotions (Study 2) Using Maximum Likelihood Estimation and Varimax Rotation ($N = 385$)

Item	Factor Loadings	
	Effectiveness	Moral Emotions
I accepted the criticism	.73/ .77	.15/ .06
I took the criticism into consideration	.78/ .81	.13/ .06
I agreed with the criticism	.74/ .83	.16/ .12
I intended to act upon the criticism by changing my behavior or to do or say things differently in the future	.93/ .92	.19/ .07
I expect that I will use the criticism to change my behavior	.93/ .93	.15/ .11
I planned to use the criticism to improve	.90/ .92	.15/ .10
I actually acted upon the criticism by changing my behavior or to do or say things differently in the future	.93/ .93	.14/ .06
I used the criticism to change something about my behavior	.91/ .94	.16/ .07
I changed my behavior based on what was criticized	.91/ .94	.19/ .09
I feel ashamed when I think about being criticized this way	.05/ -.05	.59/ .64
I feel guilty when I think about being criticized this way	.18/ .14	.84/ .87
Right now, thinking of being criticized this way, I feel guilty	.16/ .10	.95/ .95
As a member of the group, right now I feel guilty	.24/ .17	.83/ .82

Note: Morality/ Competence condition. Loadings in bold and are included in the final scale.

Figures

Figures. There was some overlap between the acceptance part (three items) of our main dependent variable effectiveness and motivation to improve (see this document Table S2.16). We therefore ran the mediation model without the acceptance items. Taking out these items improves the distinction of these two factors but does not change mediation (see Figure S2.4).

Further, we include (non-significant) mediation effects for the between-participants factors (i.e., group-membership of critic) on effectiveness to this file (see Figure S2.5 and Figure S2.6). The results nicely demonstrate that, even though not always significant, all paths have the same direction as in the complementing within-factors models, providing more evidence for the consistency of our results across both studies. Please note, as we asked participants to reflect on two criticism situations per study, we included the mediation analyses separately for the two within factors (i.e., competence and morality).

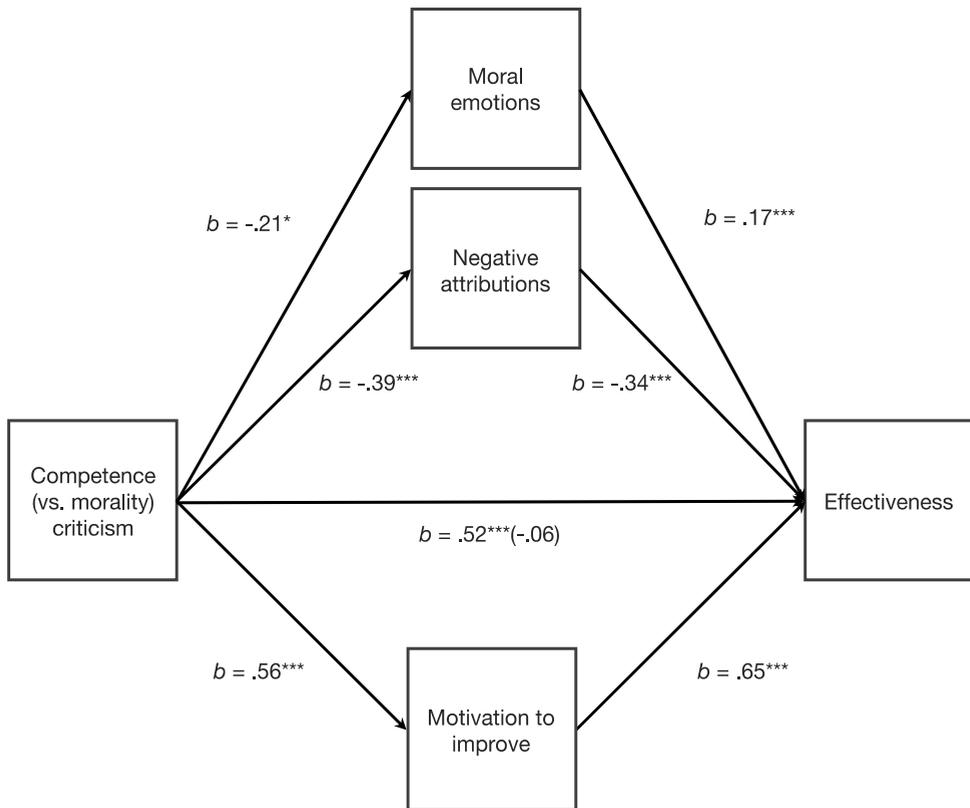


Figure S2.4. Mediation model Study 2 with unstandardized regression coefficients (b) depicting the relationship the within-participants factor ‘dimension’ and ‘effectiveness of the criticism’ (without acceptance), as mediated by ‘moral emotions’ (higher numbers = more moral emotions), ‘negative attributions about critic’ (higher numbers = more attributions), and the ‘motivation to improve’ (higher numbers = more motivation). The unstandardized coefficient for the direct effect of ‘dimension’ on ‘effectiveness of criticism’ is in parentheses.

Note: $*p < .05$, $***p < .001$

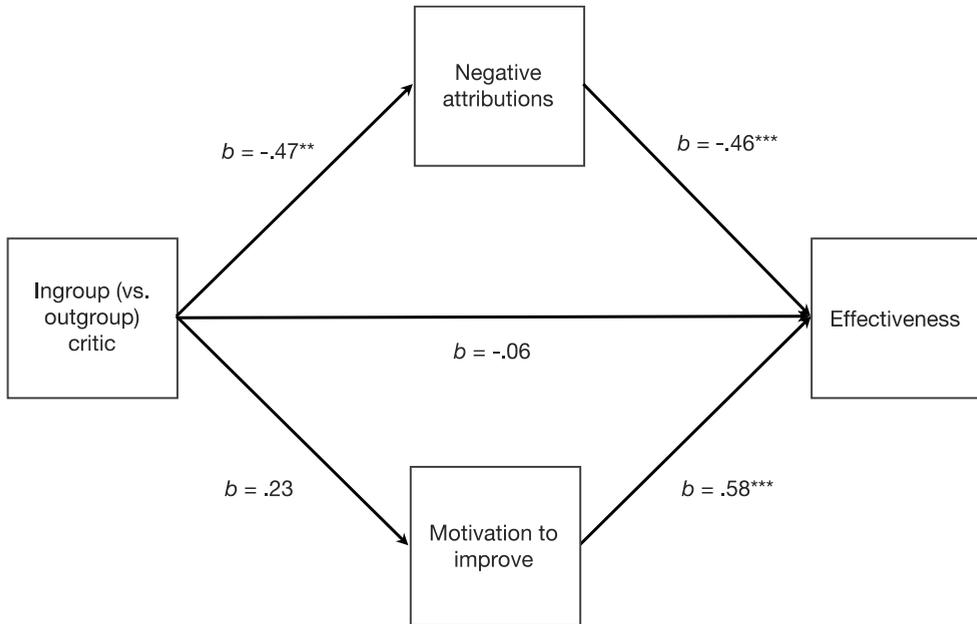


Figure S2.5. (Non-significant) mediation model Study 2 for the competence condition with unstandardized regression coefficients (b) depicting the relationship between the within-participants factor 'group-membership of critic' and 'effectiveness of the criticism', with (partly non-significant) mediators 'negative attributions the about critic' and 'motivation to improve'. Note: $*p < .05$, $***p < .001$

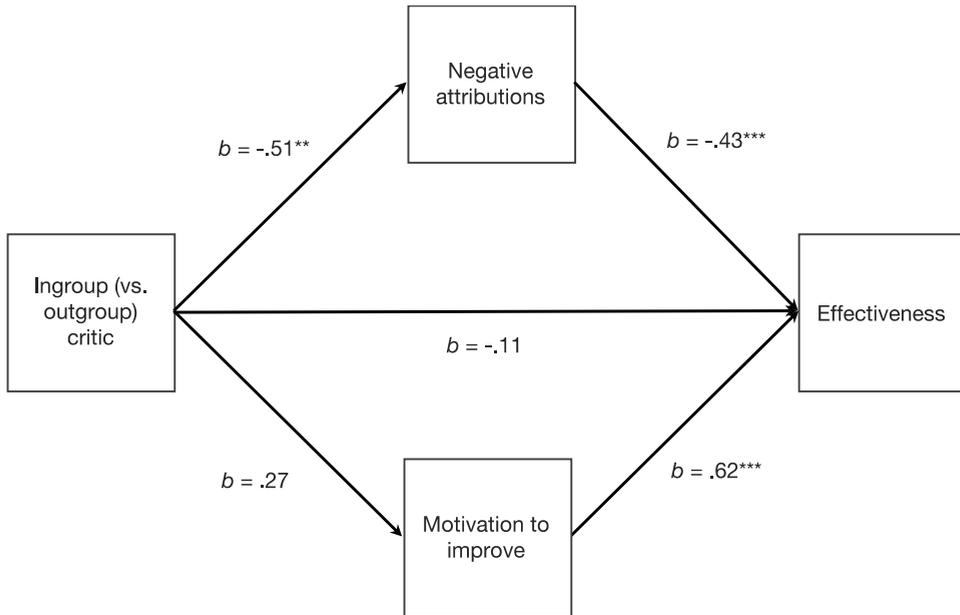


Figure S2.6. (Non-significant) mediation model Study 2 for the morality condition with unstandardized regression coefficients (b) depicting the relationship between the within-participants factor 'group-membership of critic' and 'effectiveness of the criticism', with (partly non-significant) mediators 'negative attributions the about critic' and 'motivation to improve'.

Note: $*p < .05$, $***p < .001$

Supplementary Materials – Chapter 3

Study 2.1

Tables

Table S3.1

English character traits presented as feedback messages per feedback dimension and valence

Competence		Morality	
Positive	Negative	Positive	Negative
Competent	Inexpert	Cooperative	Uncooperative
Knowledgeable	Unknowing	Trustworthy	Hurtful
Intelligent	Idiotic	Idealistic	Corrupt
Realistic	Inexperienced	Virtuous	Unfair
Smart	Silly	Responsible	Irresponsible
Clever	Incompetent	Worthy	Selfish
Bright	Stupid	Principled	Immoral
Skillful	Irrational	Moral	Shameless

Study 2.2

Instruments. We made a small adaptation to our behavioral measure (i.e., monetary contributions to common good) based on participants' answers in open-ended questions at the end of our experiment. That is, participants reported finding it difficult to make a contribution decisions twelve times (for each month of one hypothetical year). In Study 2, we therefore reduced the contributions to the overall common good (i.e., crowdfunding new housing) to one decision, but added a couple more contribution decisions so the ostensible other participants would have enough information to judge the participant's character. Specifically, we asked participants to decide what amount they would be willing to contribute to a gym room for the new housing complex, and what amount they would be willing to contribute to a study room. We used only the original contribution decision to the common good for our analysis to keep it consistent with Study 1 (for analyses with all 3 options see below).

Measures. We used the same measures as in Study 1. Reliability measures for our manipulation check identification with the group-membership of the senders were sufficient, ingroup, $\alpha = .89$, outgroup, $\alpha = .85$. Some of the other self-report measures showed lower reliability compared to Study 1 (probably due to the lower sample size). However, we decided to keep the scales as in Study 1 for consistency across studies. Reliability for *Negative emotions after group-judgements* was, ingroup: $\alpha = .81$, outgroup: $\alpha = .82$, for *negative attributions made about group evaluating*, ingroup: $\alpha = .56$, outgroup: $\alpha = .71$, for *positive emotions after group-judgements*, ingroup: $\alpha = .68$, outgroup: $\alpha = .51$, and for *moral emotions after group-judgements*, ingroup: $r[47] = .18$, $p = .223$, outgroup: $r[47] = .54$, $p < .001$.

Results. A binary logistic mixed model with accuracy of categorization (i.e., correct vs. incorrect) as dependent variable and sender's group-membership as predictor revealed no effect of group-membership on accuracy of categorization, $p = .232$. Thus, participants did not categorize one group better than the other.

Event-related brain potentials (ERPs): Social categorization and face perception

Method. Consistent with past literature it (Ito & Urland, 2003, 2005; Nunspeet, Ellemers, Derks, & Nieuwenhuis, 2014), ERPs related to perceptual attention and social categorization of faces (i.e., N100, N200, P200) were largest at fronto-central electrodes Fz, FCz, and Cz. Peaks for the N100 were scored within the time window of 90–120 ms post-stimulus-onset, for the P200 within 120–200 ms post-stimulus-onset, and for the N200 within 200–250 ms post-stimulus-onset. Peak-amplitude values were submitted to a 3 (Electrode site) x 2 (Sender's group-membership) RM ANOVA.

Consistent with past literature (Ito & Bartholow, 2009; Ito & Urland, 2005), N170 peaks were most evident at the left and right posterior sites (P7 and P8, respectively). Peak-amplitude values were submitted to a 2 (Electrode site) x 2 (Sender's group-membership) RM ANOVA.

Stimulus locked epochs for the face stimulus ranged from -200 prior to 500 ms after the event.

Results.

N100. There was no effect of sender's group-membership on N100-amplitudes, $F < 1$.

P200. There was no effect of sender's group-membership on P200-amplitudes, $F < 1$.

N200. There was a main effect of sender's group-membership on N200-amplitudes. The N200 showed higher amplitudes for outgroup ($M = -5.93$, $SE = .42$) compared to ingroup faces ($M = -5.34$, $SE = .48$), $F(1, 40) = 6.45$, $p = .015$, $\eta_p^2 = .14$ (see Figure 3.1). There was also a main effect of electrode, $F(2, 80) = 10.32$, $p = .001$, $\eta_p^2 = .21$, with largest amplitudes at Fz ($M = -5.92$, $SE = .44$). There were no other effects, $F_s \leq 3.22$, $p_s \geq .062$.

N170. There was no effect of sender's group-membership on N170-amplitudes, $F < 1$.

Tables

Table S3.2

Dutch character traits presented as feedback messages per feedback dimension and valence

Competence		Morality	
Positive	Negative	Positive	Negative
Ambitueus	Amateuristisch	Betrouwbaar	Bedrieglijk
Bedreven	Chaotisch	Eerlijk	Gemeen
Bekwaam	Dom	Ethisch	Gluiperig
Capabel	Dwaas	Gehoorzaam	Immoreel
Competent	Gebrekkig	Idealistisch	Misleidend
Deskundig	Hulpeloos	Integer	Onbetrouwbaar
Efficiënt	Incapabel	Loyaal	Oneerlijk
IJverig	Incompetent	Moreel	Onethisch
Intellectueel	Inefficiënt	Oprecht	Ongehoorzaam
Intelligent	Irrationeel	Principieel	Onoprecht
Professioneel	Onbekwaam	Rechtvaardig	Onrechtvaardig
Realistisch	Ondeskundig	Respectvol	Onverantwoordelijk
Succesvol	Ongeschikt	Trouw	Roekeloos
Talentvol	Onnozel	Voorbeeldig	Schaamteloos
Vastberaden	Onwetend	Waardig	Slecht
Wijs	Stom	Welgemeend	Verraderlijk

Table S3.3

	Study 1 (English)	Study 2 (Dutch)
Negative attributions	I think the [students/trainees] who judged me did this in my best interest	Ik denk dat de [Nederlandse/internationale] studenten die mij beoordeelden, dat in mijn belang deden
	I think the [students/trainees] were able to make these judgments	Ik denk dat de [Nederlandse/internationale] studenten in staat waren deze beoordelingen te maken
	The judgments from the [students/trainees] were credible	De beoordelingen van de [Nederlandse/internationale] studenten waren geloofwaardig
	The judgments from the [students/trainees] were meant to be constructive	De beoordelingen van de [Nederlandse/internationale] studenten waren constructief bedoeld
	The judgments from the [students/trainees] were meant to be destructive	De beoordelingen van de [Nederlandse/internationale] studenten waren destructief bedoeld
	<i>I liked the [students/trainees] who judged me</i>	
	<i>I think the [students/trainees] who judged me are stupid</i>	
Negative emotions	The judgments of the [students/trainees] upset me	Ik raak van streek door de beoordelingen van de [Nederlandse/internationale] studenten
	The judgments of the [students/trainees] made me feel discouraged	Ik voel me ontmoedigd door de beoordelingen van de [Nederlandse/internationale] studenten
	The judgments of the [students/trainees] made me feel disappointed	Ik voel me teleurgesteld door de beoordelingen van de [Nederlandse/internationale] studenten
	The judgments of the [students/trainees] made me feel tense	Ik voel me gespannen door de beoordelingen van de [Nederlandse/internationale] studenten
Positive emotions	The judgments of the [students/trainees] made me feel relaxed	Ik voel me ontspannen door de beoordelingen van de [Nederlandse/internationale] studenten
	The judgments of the [students/trainees] made me feel excited	Ik voel me enthousiast door de beoordelingen van de [Nederlandse/internationale] studenten
	The judgments of the [students/trainees] made me feel confident	Ik voel me zelfverzekerd door de beoordelingen van de [Nederlandse/internationale] studenten
Moral emotions	The judgments from the [students/trainees] made me feel ashamed	Ik schaam me voor de beoordelingen van de [Nederlandse/internationale] studenten
	The judgments of the [students/trainees] made me feel guilty	Ik voel me schuldig door de beoordelingen van de [Nederlandse/internationale] studenten

Note. Excluded items are in cursive. The two excluded items were added to items from previous research (Hornsey et al., 2008) for the specific context of this study. However, they did not seem to measure a similar construct.

Figures

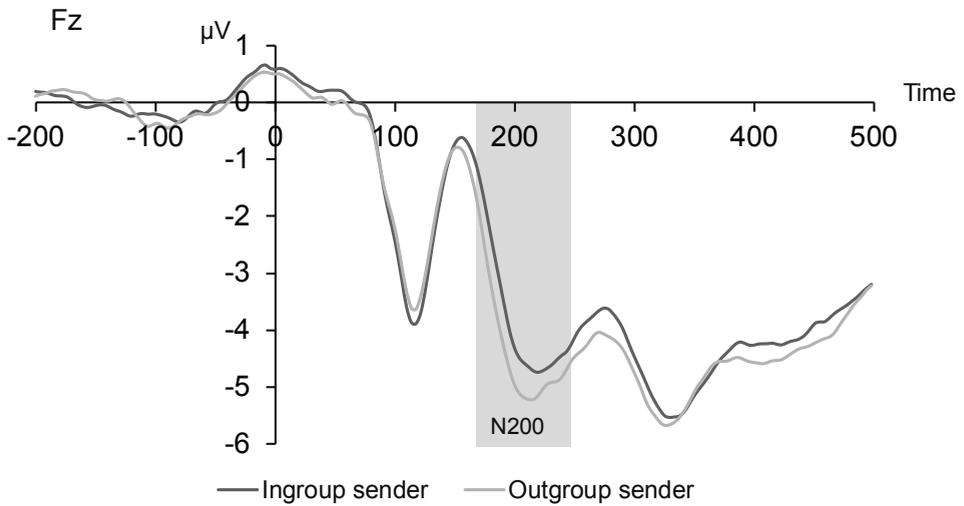


Figure S3.1. ERP-waveforms at frontal electrode Fz for ingroup (i.e., Dutch students) and outgroup (i.e., international students) target faces.

Supplementary Materials – Chapter 4

Study 3.1

Power calculations. Since we used an unbalanced design, meaning that the actual trial number per condition was dependent on participants choices, we made an estimation of the repeated measures. As input for the number of repeated measures we used data from a pilot test to get an estimate of how many times participants would make proself decisions. This was needed because in the current research we were only interested in proself decisions, which would lead to negative feedback, as compared to prosocial decisions, which would lead to positive feedback (see Instruments). In the pilot test, participants made proself decisions and received negative feedback around 30% of the trials. As we changed the definition of proself decisions to increase the amount of negative feedback participants would get in the current experiment compared to the pilot, we approximated that participants would receive negative feedback 50% of all trials. This would be equivalent to 10 trials of negative feedback per condition.

Recall task. The current study was used as a behavioral pilot for the second study in which we investigated cognitive processing of feedback messages. To explore whether the factor time of communicating intentions (i.e., before the sender was presented, before the message was presented, and no intentions communicated) would modulate cognitive processing of feedback messages, we added a recall task of feedback messages at the end of the experiment. Checking whether participants recalled feedback messages more often in one of the conditions, for example, if intentions were communicated right before the feedback message was shown, could indicate whether participants also allocated more attentional resources to process feedback. We did not find any differences for this task, $F < 1$.

SVO types. Analyses of the SVO task showed that 39 participants were classified as prosocial and 5 as individualist.

Study 3.2

As in past research (Rösler et al., 2021b, p. 223) we additionally investigated social categorization processes of sender faces to check whether participants initially distinguish between ingroup and outgroup faces by preferentially processing faces of ingroup or outgroup members. We predicted that participants would show greater amplitudes on ERPs related to social categorization (i.e., N200) for outgroup as compared to ingroup faces.

Event-related brain potentials (ERPs): Social categorization and face perception

Method

We matched the feedback sender's gender to the participant's gender. We counterbalanced the background color indicating the sender's group-membership. That is, the same set of faces was shown to half of the participants as belonging to the ingroup, whereas for the other half of participants it was shown as belonging to the outgroup. This way, we can be certain that ERPs related to the perceptual processing of faces (e.g., N170) are only due to the group-membership of the senders, rather than their different facial features.

Perceptual attention and social categorization of faces. N100, N170, P150. We selected the electrodes to include in the analyses based on the literature and the visual inspection of the grand-averages. Consistent with past literature (Ito & Urland, 2003; Rösler et al., 2021b; Van Nunspeet et al., 2014), ERPs related to perceptual attention and social categorization of faces (i.e., N100, P150) were evident at centro-frontal and centro-parietal electrodes Fz, FCz, Cz, CPz, and Pz. Mean amplitudes for the N100 were averaged between 100–125ms, and for the P150, between 140–170ms post stimulus-onset. Compared to previous research using a very similar design (Rösler et al., 2021b, p. 223), we did not visually identify a clear N200 in the current dataset.

For the N170, consistent with past literature (Ito & Bartholow, 2009; Rösler et al., 2021b, p. 223), there was a peak at the left posterior P7

and right posterior site P8. Mean amplitudes for the N170 were averaged between 120–180 ms post stimulus-onset.

Results

SVO types. Analyses of the SVO task showed that 29 participants were classified as prosocial and 5 as individualist.

Perceptual attention and social categorization of faces. We tested whether the social group-membership (ingroup vs. outgroup) modulated ERPs related to the perceptual attention and social categorization of faces of feedback senders. We used LMMs (ML) to predict N100 and P150 amplitudes and submitting the participant number as a random effect and the sender's group-membership (ingroup vs. outgroup) and electrode site (Fz vs. FCz vs. Cz vs. CPz vs. Pz) as fixed effects. For the N170, we used the same model but submitted relevant electrodes for the N170 (i.e., P7 vs. P8). As anticipated, there were no effects of sender's group-membership on these ERPs, $bs \leq -.12$, $p \geq .813$.

Since inspecting the grand-averages seemed to suggest that there was an effect of the sender's group-membership on the N170 for the left posterior electrode P7, we ran a separate model only including P7. This analysis showed that, interestingly, there was an effect of sender's group-membership on N170-amplitudes, $B = -.80$, $t = -2.33$, $p = .030$, 95% CI = [-1.51, -.10]. The N170 showed higher amplitudes for outgroup compared to ingroup faces. We had predicted this effect (and the direction of it) for the N200 rather than the N170. However, the N200 was not identifiable in the current dataset.

Tables

Table S4.1

List of sentences used for motives manipulation

Acknowledgment	Helping	Moral superiority
Ik zie dat je het belangrijk vindt om geld te geven aan goede doelen	Ik wil je helpen om overeenkomstig te handelen met jouw eigen waarden	Ik geef altijd meer geld aan goede doelen vergeleken met jou
Het is geweldig dat je geeft om doneren aan goede doelen	Ik wil je helpen jouw doel te bereiken als het gaat om doneren aan goede doelen	Ik ben een beter persoon dan jij, omdat ik meer geld geef aan goede doelen
Je lijkt het belangrijk te vinden om te geven aan mensen in nood	Ik wil je herinneren dat je hebt gezegd dat we geld zouden moeten geven aan goede doelen zodat je overeenkomstig met jouw waarden kunt handelen	Ik wil je met dit oordeel laten zien dat ik een beter persoon lijk te zijn dan jij
Fijn om te zien dat je erom geeft om anderen te helpen	Je zei dat het belangrijk was om geld te geven aan goede doelen en ik wil je helpen overeenkomstig te handelen met jouw waarden	Ik wil duidelijk maken dat ik een beter persoon ben dan jij
Het lijkt alsof het belangrijk is voor jou om te doneren aan goede doelen	Ik wil je helpen door je te herinneren dat je hebt gezegd dat het belangrijk is om te doneren aan goede doelen	Ik lijk een beter persoon te zijn dan jij en daarom geef ik je dit oordeel
Je laat zien dat je het op waarde stelt om mensen in nood te helpen		
Ik zie dat je geld hebt gegeven aan dit goede doel, dat is fijn		
Geven aan mensen lijkt iets te zijn wat jij op waarde stelt		
Het is fijn dat je geeft aan mensen in nood		
Wat fijn dat je geld hebt gegeven aan dit goede doel		



References

- Abele-Brehm, A., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2020). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. *Psychological Review*, 28(2), 290-314, doi: 10.1037/rev0000262
- Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness-communion with warmth and morality. *Frontiers in Psychology*, 7, 1–17. <https://doi.org/10.3389/fpsyg.2016.01810>
- Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, 93(5), 751–763. <https://doi.org/10.1037/0022-3514.93.5.751>
- Achterberg, M., van Duijvenvoorde, A. C. K., Bakermans-Kranenburg, M. J., & Crone, E. A. (2016). Control your anger! The neural basis of aggression regulation in response to negative social feedback. *Social Cognitive and Affective Neuroscience*, 11(5), 712–720. <https://doi.org/10.1093/scan/nsv154>
- Amodio, D. M., & Cikara, M. (2021). The social neuroscience of prejudice. *Annual Review of Psychology*, 72, 439–469. <https://doi.org/10.1146/annurev-psych-010419-050928>
- Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423–1440. <https://doi.org/10.1037/0022-3514.83.6.1423>
- Aristotle, ., & Roberts, W. R. (2004). *Rhetoric*. Dover Publications.
- Aristotle, ., Ross, W. D., & Brown, L. (2009). *The Nicomachean ethics*. Oxford University Press.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects

- modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37(2), 122–147. <https://doi.org/10.1037/0003-066X.37.2.122>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/doi:10.18637/jss.v067.i01>
- Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, 2(3), 412–414. <https://doi.org/10.1098/rsbl.2006.0509>
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad Is Stronger Than Good. *Review of General Psychology*, 5(4), 323–370. <https://doi.org/10.1037/1089-2680.5.4.323>
- Berlo, D. K. (1960). *The process of communication*. Holt, Rinehart, Wins.
- Bradley, M., & Lang, P. J. (1994). Measuring Emotion: The Self-assessment semantic differential manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49–59. [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
- Brady, W. J., Crockett, M. J., & Van Bavel, J. J. (2020). The MAD model of moral contagion: The role of motivation, attention, and design in the spread of moralized content online. *Perspectives on Psychological Science*, 15(4), 978–1010. <https://doi.org/10.1177/1745691620917336>
- Brambilla, M., & Leach, C. W. (2014). On the importance of being moral: The distinctive role of morality in social judgment. *Social Cognition*, 32(4), 397–408. <https://doi.org/10.1521/soco.2014.32.4.397>
- Brambilla, M., University, S. S., Rusconi, P., & Goodwin, G. P. (2021). The primacy of morality in impression development: Theory, research, and

- future directions. *Advances in Experimental Social Psychology*, 64.
- Brockner, J., & Higgins, E. T. (2001). Regulatory Focus Theory: Implications for the study of emotions at work. *Organizational Behavior and Human Decision Processes*, 86(1), 35–66. <https://doi.org/10.1006/obhd.2001.2972>
- Burke, K. (1969). *A rhetoric of motives*. University of California Press.
- Carlson, R. W., Maréchal, M. A., Oud, B., Fehr, E., & Crockett, M. J. (2020). Motivated misremembering of selfish decisions. *Nature Communications*, 11(1), 2100. <https://doi.org/10.1038/s41467-020-15602-4>
- Carretié, L., Martín-Loeches, M., Hinojosa, J. A., & Mercado, F. (2001). Emotion and attention interaction studied through event-related potentials. *Journal of Cognitive Neuroscience*, 13(8), 1109–1128. <https://doi.org/10.1162/089892901753294400>
- Carretié, L., Mercado, F., Tapia, M., & Hinojosa, J. A. (2001). Emotion, attention, and the ‘negativity bias’, studied through event-related potentials. *International Journal of Psychophysiology*, 41, 75–85.
- Cavazza, N., Pagliaro, S., & Guidetti, M. (2014). Antecedents of concern for personal reputation: The role of group entitativity and fear of social exclusion. *Basic and Applied Social Psychology*, 36(4), 365–376. <https://doi.org/10.1080/01973533.2014.925453>
- Cikara, M., & Van Bavel, J. J. (2014). The Neuroscience of intergroup relations: An integrative review. *Perspectives on Psychological Science*, 9(3), 245–274. <https://doi.org/10.1177/1745691614527464>
- Cramwinckel, F. M., van Dijk, E., Scheepers, D., & van den Bos, K. (2013). The threat of moral refusers for one’s self-concept and the protective function of physical cleansing. *Journal of Experimental Social Psychology*, 49(6), 1049–1058. <https://doi.org/10.1016/j.jesp.2013.07.009>

- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, 40(07), 61–149. [https://doi.org/10.1016/S0065-2601\(07\)00002-0](https://doi.org/10.1016/S0065-2601(07)00002-0)
- Eisenberger, N. I., & Lieberman, M. D. (2004). Why rejection hurts: A common neural alarm system for physical and social pain. *Trends in Cognitive Sciences*, 8(7), 294–300. <https://doi.org/10.1016/j.tics.2004.05.010>
- Ellemers, N. (2012). The group self. *Science*, 336(6083), 848–852. <https://doi.org/10.1126/science.1220987>
- Ellemers, N. (2017). *Morality and the regulation of social behavior: Groups as moral anchors*. Milton Park, UK.
- Ellemers, N., & Haslam, S. A. (2011). Social identity theory. In P. van Lange, A. Kruglanski, & T. Higgins (Eds.), *Handbook of Theories of Social Psychology* (pp. 379–398). Sage.
- Ellemers, N., Kingma, L., Van De Burgt, J., Barreto, M., Burgt, J. van de, & Barreto, M. (2011). Corporate social responsibility as a source of organizational morality, employee commitment and satisfaction. *Journal of Organizational Moral Psychology*, 1(2), 97–124.
- Ellemers, N., Kortekaas, P., & Ouwerkerk, J. W. (1999). Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. *European Journal of Social Psychology*, 29(2–3), 371–389. [https://doi.org/10.1002/\(SICI\)1099-0992\(199903/05\)29:2/3<371::AID-EJSP932>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1099-0992(199903/05)29:2/3<371::AID-EJSP932>3.0.CO;2-U)
- Ellemers, N., Pagliaro, S., & Barreto, M. (2013). Morality and behavioural regulation in groups: A social identity approach. *European Review of Social Psychology*, 24(1), 160–193. <https://doi.org/10.1080/10463283.2013.841490>
- Ellemers, N., Pagliaro, S., Barreto, M., & Leach, C. W. (2008). Is it better

- to be moral than smart? The effects of morality and competence norms on the decision to work at group status improvement. *Journal of Personality and Social Psychology*, 95(6), 1397–1410. <https://doi.org/10.1037/a0012628>
- Ellemers, N., & van den Bos, K. (2012). Morality in groups: On the social-regulatory functions of right and wrong. *Social and Personality Psychology Compass*, 6(12), 878–889. <https://doi.org/10.1111/spc3.12001>
- Ellemers, N., & Van der Toorn, J. (2015). Groups as moral anchors. *Current Opinion in Psychology*, 6, 189–194. <https://doi.org/10.1016/j.copsy.2015.08.018>
- Ellemers, N., van der Toorn, J., Paunov, Y., & van Leeuwen, T. (2019). The psychology of morality: A review and analysis of empirical studies published from 1940 through 2017. *Personality and Social Psychology Review*. 1-35. <https://doi.org/10.1177/1088868318811759>
- Ellemers, N., & van Nunspeet, F. (2020). Neuroscience and the social origins of moral behavior: How neural underpinnings of social categorization and conformity affect everyday moral and immoral behavior. *Current Directions in Psychological Science*. 1-8. <https://doi.org/10.1177/0963721420951584>
- Esposito, S. R., Hornsey, M. J., & Spoor, J. R. (2013). Shooting the messenger: Outsiders critical of your group are rejected regardless of argument quality. *British Journal of Social Psychology*, 52(2), 386–395. <https://doi.org/10.1111/bjso.12024>
- Fang, C., & van Liempt, I. (2020). ‘We prefer our Dutch’: International students’ housing experiences in the Netherlands. *Housing Studies*, 1–21. <https://doi.org/10.1080/02673037.2020.1720618>
- Festinger, L. (1954). A theory of social comparison processes. In *Human Relations*, 7(20), 1–31.
- Fiedler, K., Harris, C., & Schott, M. (2018). Unwarranted inferences from

- statistical mediation tests – An analysis of articles published in 2015. *Journal of Experimental Social Psychology*, 75, 95–102. <https://doi.org/10.1016/j.jesp.2017.11.008>
- Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., Mason, L., McGrath, M. C., Nyhan, B., Rand, D. G., Skitka, L. J., Tucker, J. A., Van Bavel, J. J., Wang, C. S., & Druckman, J. N. (2020). Political sectarianism in America. *Science*, 370(6516), 533–536. <https://doi.org/10.1126/science.abe1715>
- Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist*, 48(6), 621–628. <https://doi.org/10.1037/0003-066X.48.6.621>
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77–83. <https://doi.org/10.1016/j.tics.2006.11.005>
- Galinsky, A. D., Magee, J. C., Ena Inesi, M., & Gruenfeld, D. H. (2006). Power and perspectives not taken. *Psychological Science*, 17(12), 1068–1074. <https://doi.org/10.1111/j.1467-9280.2006.01824.x>
- Galizzi, M. M., & Navarro-Martinez, D. (2019). On the external validity of social preference games: A systematic lab-field study. *Management Science*, 65(3), 976–1002. <https://doi.org/10.1287/mnsc.2017.2908>
- Gausel, N. (2013). Self-reform or self-defense? Understanding how people cope with their moral failures by understanding how they appraise and feel about their moral failures. In M. Moshe & N. Corbu (Eds.), *Walk of shame* (Issue July, pp. 191–208). Nova Publishers.
- Gausel, N., & Brown, R. (2012). Shame and guilt—do they really differ in their focus of evaluation? Wanting to change the self and behavior in response to ingroup immorality. *The Journal of Social Psychology*, 152(5), 547–567. <https://doi.org/10.1080/00224545.2012.657265>
- Gausel, N., & Leach, C. W. (2011). Concern for self-image and social image in the management of moral failure: Rethinking shame. *European*

Journal of Social Psychology, 41(4), 468–478. <https://doi.org/10.1002/ejsp.803>

Gausel, N., Leach, C. W., Vignoles, V. L., & Brown, R. (2012). Defend or repair? Explaining responses to in-group moral failure by disentangling feelings of shame, rejection, and inferiority. *Journal of Personality and Social Psychology*, 102(5), 941–960. <https://doi.org/10.1037/a0027233>

Giner-Sorolla, R. (2012). *European monographs in social psychology. Judging passions: Moral emotions in persons and groups*. (1st ed.). Psychology Press.

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106(1), 148–168. <https://doi.org/10.1037/a0034726>

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, 55(4), 468–484. [https://doi.org/10.1016/0013-4694\(83\)90135-9](https://doi.org/10.1016/0013-4694(83)90135-9)

Gu, R., Ao, X., Mo, L., & Zhang, D. (2020). Neural correlates of negative expectancy and impaired social feedback processing in social anxiety. *Social Cognitive and Affective Neuroscience*, 15(3), 285–291. <https://doi.org/10.1093/scan/nsaa038>

Herbert, C., Kissler, J., Junghöfer, M., Peyk, P., & Rockstroh, B. (2006). Processing of emotional adjectives: Evidence from startle EMG and ERPs. *Psychophysiology*, 43(2), 197–206. <https://doi.org/10.1111/j.1469-8986.2006.00385.x>

Hornsey, M. J., & Esposito, S. (2009). Resistance to group criticism and recommendations for change: Lessons from the intergroup sensitivity effect. *Social and Personality Psychology Compass*, 3(3), 275–291. <https://doi.org/10.1111/j.1751-9004.2009.00178.x>

- Hornsey, M. J., & Imani, A. (2004). Criticizing groups from the inside and the outside: an identity perspective on the intergroup sensitivity effect. *Personality & Social Psychology Bulletin*, *30*(3), 365–383. <https://doi.org/10.1177/0146167203261295>
- Hornsey, M. J., Oppes, T., & Svensson, A. (2002). “It’s OK if we say it, but you can’t”: Responses to intergroup and intragroup criticism. *European Journal of Social Psychology*, *32*(3), 293–307. <https://doi.org/10.1002/ejsp.90>
- Hornsey, M. J., Robson, E., Smith, J., Esposito, S., & Sutton, R. M. (2008). Sugaring the pill: Assessing rhetorical strategies designed to minimize defensive reactions to group criticism. *Human Communication Research*, *34*(1), 70–98. <https://doi.org/10.1111/j.1468-2958.2007.00314.x>
- Hornsey, M. J., Trembath, M., & Gunthorpe, S. (2004). “You can criticize because you care”: Identity attachment, constructiveness, and the intergroup sensitivity effect. *European Journal of Social Psychology*, *34*(5), 499–518. <https://doi.org/10.1002/ejsp.212>
- Hornsey, M. J., & Hogg, M. A. (2000). Assimilation and diversity: An integrative model of subgroup relations. *Personality and Social Psychology Review*, *4*(2), 143–156. https://doi.org/10.1207/S15327957PSPR0402_03
- Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, *64*(4), 349–371. <https://doi.org/10.1037/0021-9010.64.4.349>
- Ito, T. A., & Bartholow, B. D. (2009). The neural correlates of race. *Trends in Cognitive Sciences*, *13*(12), 524–531. <https://doi.org/10.1016/j.tics.2009.10.002>
- Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity

- bias in evaluative categorizations. *Journal of Personality and Social Psychology*, 75(4), 887–900. <https://doi.org/10.1037/0022-3514.75.4.887>
- Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, 85(4), 616–626. <https://doi.org/10.1037/0022-3514.85.4.616>
- Ito, T. A., & Urland, G. R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, & Behavioral Neuroscience*, 5(1), 21–36.
- Judd, C. M., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology*, 89(6), 899–913. <https://doi.org/10.1037/0022-3514.89.6.899>
- Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, 42(3), 643–650. <https://doi.org/10.3758/BRM.42.3.643>
- Kissler, J., Assadollahi, R., & Herbert, C. (2006). Emotional and semantic networks in visual word processing: insights from ERP studies. *Progress in Brain Research*, 156, 147–183. [https://doi.org/10.1016/S0079-6123\(06\)56008-X](https://doi.org/10.1016/S0079-6123(06)56008-X)
- Klein, F., Iffland, B., Schindler, S., Wabnitz, P., & Neuner, F. (2015). This person is saying bad things about you: The influence of physically and socially threatening context information on the processing of inherently neutral faces. *Cognitive, Affective and Behavioral Neuroscience*, 15(4), 736–748. <https://doi.org/10.3758/s13415-015-0361-8>
- Koch, A., Yzerbyt, V., Abele, A., Ellemers, N., & Fiske, S. T. (2020). Social

- evaluation: Comparing models across interpersonal, intragroup, intergroup, several-group, and many-group contexts. *Advances in Experimental Social Psychology*, 63.
- Komorita, S., & Parks, C. (1994). *Brown & Benchmark's social psychology series. Social dilemmas*. Brown & Benchmark. <http://psycnet.apa.org/record/1993-99062-000>
- Kouchaki, M., & Gino, F. (2016). Memories of unethical actions become obfuscated over time. *Proceedings of the National Academy of Sciences of the United States of America*, 113(22), 6166–6171. <https://doi.org/10.1073/pnas.1523586113>
- Krippendorff, K. (2004). Content analysis: An introduction to its methodology. In *Education* (Vol. 79). <https://doi.org/10.2307/2288384>
- Krueger, J. I., Heck, P. R., Evans, A. M., & Theresa, E. (2020). European review of social psychology social game theory: preferences, perceptions, and choices. *European Review of Social Psychology*, 31(1), 222–253. <https://doi.org/10.1080/10463283.2020.1778249>
- Lang, P. J., Bradley, M. M., & Cuthbert, M. M. (1997). Motivated attention: Affect, activation and action. In P. J. Lang, R. F. Simons, & M. T. Balaban (Eds.), *Attention and orienting: Sensory and motivational processes* (pp. 97–136). Hillsdale.
- Lang, Peter J, & Bradley, M. M. (2013). Appetitive and Defensive Motivation: Goal-Directed or Goal-Determined? *Emotion Review*, 5(3), 230–234. <https://doi.org/10.1177/1754073913477511>
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the radboud faces database. *Cognition and Emotion*, 24(8), 1377–1388. <https://doi.org/10.1080/02699930903485076>
- Lasswell, H. D. (1948). The structure and function of communication in society. In L. Bryson (Ed.), *The Communication of Ideas*. Harper.

- Leach, C. W., Ellemers, N., Barreto, M., Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: the importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, *93*(2), 234–249. <https://doi.org/10.1037/0022-3514.93.2.234>
- León, I., Díaz, J. M., de Vega, M., & Hernández, J. A. (2010). Discourse-based emotional consistency modulates early and middle components of event-related potentials. *Emotion*, *10*(6), 863–873. <https://doi.org/10.1037/a0019983>
- Levitt, S. D., & List, J. A. (2007). About the real world? *Journal of Economic Perspectives*, *21*(2), 153–174.
- Lexico. (n.d.). *Lexico.com dictionary*. Retrieved August 4, 2021, from <https://www.lexico.com/definition/defensiveness>
- Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. MIT Press.
- Luck, S. J., & Kappenman, E. S. (2018). *ERP Boot Camp Tip: Why mean amplitude is usually superior to peak amplitude*. <https://erpinform.org/blog/2018/7/5/mean-versus-peak-amplitude>
- Lüdecke, D., Makowski, D., Waggoner, P., & Patil, I. (2020). *Performance: Assessment of regression models performance*. R package version 0.4.7. <https://cran.r-project.org/package=performance>
- Martijn, C., Spears, R., Van Der Pligt, J., & Jakobs, E. (1992). Negativity and positivity effects in person perception and inference: Ability versus morality. *European Journal of Social Psychology*, *22*(5), 453–463. <https://doi.org/10.1002/ejsp.2420220504>
- Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, *45*(6), 633–644. <https://doi.org/10.1509/jmkr.45.6.633>
- McCrum, D. (2015). The house of wirecard. *Financial Times*. <https://www.>

ft.com/content/534e7c4d-3101-3f6a-abc8-dc70beab35b7

- Messick, D. M., & McClintock, C. G. (1968). Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology*, 4(1), 1–25. [https://doi.org/10.1016/0022-1031\(68\)90046-2](https://doi.org/10.1016/0022-1031(68)90046-2)
- Minson, J. A., & Monin, B. (2011). Do-gooder derogation. *Social Psychological and Personality Science*, 3(2), 200–207. <https://doi.org/10.1177/1948550611415695>
- Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology*, 95(1), 76–93. <https://doi.org/10.1037/0022-3514.95.1.76>
- Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods*, 22(1), 6–27. <https://doi.org/10.1037/met0000086>
- Moscатели, S., Prati, F., & Rubini, M. (2019). If you criticize us, do it in concrete terms: Linguistic abstraction as a moderator of the intergroup sensitivity effect. *Journal of Language and Social Psychology*, 38(5–6), 680–705. <https://doi.org/10.1177/0261927X19864686>
- Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. J. (2011). Measuring social value orientation. *Judgment and Decision Making*, 6(8), 771–781 Measuring. <https://doi.org/10.1002/fut.2001>
- Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, 15(3), 263–280. <https://doi.org/10.1002/ejsp.2420150303>
- Otten, K., Buskens, V., Przepiorka, W., & Ellemers, N. (2020). Heterogeneous groups cooperate in public good problems despite normative disagreements about individual contribution levels. *Scientific Reports*, 10(1), 1–12. <https://doi.org/10.1038/s41598-020-73314-7>

- Pagliaro, S., Ellemers, N., & Barreto, M. (2011). Sharing moral values: Anticipated ingroup respect as a determinant of adherence to morality-based (but not competence-based) group norms. *Personality and Social Psychology Bulletin*, *37*(8), 1117–1129. <https://doi.org/10.1177/0146167211406906>
- Pagliaro, S., Ellemers, N., Barreto, M., & Di Cesare, C. (2016). Once dishonest, always dishonest? The impact of perceived pervasiveness of moral evaluations of the self on motivation to restore a moral reputation. *Frontiers in Psychology*, *7*(APR), 1–12. <https://doi.org/10.3389/fpsyg.2016.00586>
- Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*(5), 519–526. <https://doi.org/10.3758/BF03197524>
- Pfabigan, D. M., Holzner, M. T., & Lamm, C. (2016). Performance monitoring during a minimal group manipulation. *Social Cognitive and Affective Neuroscience*, *11*(10), 1560–1568. <https://doi.org/10.1093/scan/nsw071>
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & Team, R. C. (2020). *Nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-148.
- Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clin Neurophysiol.*, *118*(10). <https://doi.org/10.1016/j.clinph.2007.04.019>. Updating
- Prentice, D. A., Miller, D. T., & Lightdale, J. R. (1994). Asymmetries in Attachments to Groups and to their Members: Distinguishing between Common-Identity and Common-Bond Groups. *Personality and Social Psychology Bulletin*, *20*(5), 484–493.
- Ratner, K. G., Dotsch, R., Wigboldus, D. H. J., Knippenberg, A. van, & Amodio, D. M. (2014). Visualizing minimal ingroup and outgroup faces: Implications for impressions, attitudes, and behavior. *Journal*

- of Personality and Social Psychology*, 106(6). <https://doi.org/10.1037/a0036498>
- Rösler, I. K., Van Nunspeet, F., & Ellemers, N. (2021a). Don't tell me about my moral failures but motivate me to improve: Increasing effectiveness of outgroup criticism by criticizing one's competence. *European Journal of Social Psychology*, 51(3), 597–609. <https://doi.org/10.1002/ejsp.2764>
- Rösler, I. K., Van Nunspeet, F., & Ellemers, N. (2021b). Sender identity and social dimension moderate the attentional processing of, and behavioral responses towards, negative feedback – an ERP study. *In Preparation*.
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 5(4), 296–320. https://doi.org/10.1207/S15327957PSPR0504_2
- Schindler, S., & Kissler, J. (2016). People matter: Perceived sender identity modulates cerebral processing of socio-emotional language feedback. *NeuroImage*, 134, 160–169. <https://doi.org/10.1016/j.neuroimage.2016.03.052>
- Schindler, S., & Kissler, J. (2018). Language-based social feedback processing with randomized “senders”: An ERP study. *Social Neuroscience*, 13(2), 202–213. <https://doi.org/10.1080/17470919.2017.1285249>
- Schindler, S., Miller, G. A., & Kissler, J. (2020). Attending to Eliza: Rapid brain responses reflect competence attribution in virtual social feedback processing. *Social Cognitive and Affective Neuroscience*, 14(10), 1073–1086. <https://doi.org/10.1093/scan/nsz075>
- Schindler, S., Wegrzyn, M., Steppacher, I., & Kissler, J. (2014). It's all in your head - how anticipating evaluation affects the processing of emotional trait adjectives. *Frontiers in Psychology*, 5(NOV), 1–10. <https://doi.org/10.3389/fpsyg.2014.01292>

- Schindler, S., Wegrzyn, M., Steppacher, I., & Kissler, J. (2015). Perceived communicative context and emotional content amplify visual word processing in the fusiform gyrus. *The Journal of Neuroscience*, *35*(15), 6010–6019. <https://doi.org/10.1523/JNEUROSCI.3346-14.2015>
- Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Hillman, C. H., Hamm, A. O., & Lang, P. J. (2004). Brain processes in emotional perception: Motivated attention. *Cognition and Emotion*, *18*(5), 593–611. <https://doi.org/10.1080/02699930341000239>
- Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: event-related brain potential studies. *Progress in Brain Research*, *156*, 31–51. [https://doi.org/10.1016/S0079-6123\(06\)56002-9](https://doi.org/10.1016/S0079-6123(06)56002-9)
- Schwarz, K. A., Wieser, M. J., Gerdes, A. B. M., Mühlberger, A., & Pauli, P. (2013). Why are you looking like that? How the context influences evaluation and processing of human faces. *Social Cognitive and Affective Neuroscience*, *8*(4), 438–445. <https://doi.org/10.1093/scan/nss013>
- Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-serving justifications: Doing wrong and feeling moral. *Current Directions in Psychological Science*, *24*(2), 125–130. <https://doi.org/10.1177/0963721414553264>
- Shiffman, S., Hufford, M., Hickcox, M., Paty, J. A., Gnys, M., & Kassel, J. D. (1997). Remember that? A comparison of real-time versus retrospective recall of smoking lapses. *Journal of Consulting and Clinical Psychology*, *65*(2), 292–300. <https://doi.org/10.1037/0022-006X.65.2.292.a>
- Simon, J. C., Styczynski, N., & Gutsell, J. N. (2020). Social perceptions of warmth and competence influence behavioral intentions and neural processing. *Cognitive, Affective and Behavioral Neuroscience*, *20*(2), 265–275. <https://doi.org/10.3758/s13415-019-00767-3>
- Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social

- memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology*, 52(4), 689–699. <https://doi.org/10.1037/0022-3514.52.4.689>
- Stanley, M. L., Henne, P., & De Brigard, F. (2019). Remembering moral and immoral actions in constructing the self. *Memory and Cognition*, 47(3), 441–454. <https://doi.org/10.3758/s13421-018-0880-y>
- Strohming, N., & Nichols, S. (2014). The essential moral self. *Cognition*, 131(1), 159–171. <https://doi.org/10.1016/j.cognition.2013.12.005>
- Sun, J., & Goodwin, G. P. (2020). Do people want to be more moral? *Psychological Science*. <https://doi.org/10.1177/0956797619893078>
- Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. *The Social Psychology of Intergroup Relations, January 1979*, 33–47. [https://doi.org/10.1016/S0065-2601\(05\)37005-5](https://doi.org/10.1016/S0065-2601(05)37005-5)
- Täuber, S., Gausel, N., & Flint, S. W. (2018). Weight bias internalization: The maladaptive effects of moral condemnation on intrinsic motivation. *Frontiers in Psychology*, 9(SEP), 1–15. <https://doi.org/10.3389/fpsyg.2018.01836>
- Täuber, S., & van Zomeren, M. (2013). Outrage towards whom? Threats to moral group status impede striving to improve via out-group-directed outrage. *European Journal of Social Psychology*, 43(2), 149–159. <https://doi.org/10.1002/ejsp.1930>
- Täuber, S., van Zomeren, M., & Kutlaca, M. (2015). Should the moral core of climate issues be emphasized or downplayed in public discourse? Three ways to successfully manage the double-edged sword of moral communication. *Climatic Change*, 130(3), 453–464. <https://doi.org/10.1007/s10584-014-1200-6>
- Tetlock, P. E. (2002). Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. *Psychological Review*, 109(3), 451–471. <https://doi.org/10.1037/0033-295X.109.3.451>

- Thau, S., Derfler-Rozin, R., Pitesa, M., Mitchell, M. S., & Pillutla, M. M. (2015). Unethical for the sake of the group: Risk of social exclusion and pro-group. *Journal of Applied Psychology, 100*(1), 98–113.
- Thürmer, J. L., & McCrea, S. M. (2018). Beyond motivated reasoning: Hostile reactions to critical comments from the outgroup. *Motivation Science, 4*(4), 333–346. <https://doi.org/10.1037/mot0000097>
- Thürmer, J. L., McCrea, S. M., & McIntyre, B. M. (2019). Motivated collective defensiveness: Group members prioritize counterarguing out-group criticism over getting their work done. *Social Psychological and Personality Science, 10*(3), 382–392. <https://doi.org/10.1177/1948550618762298>
- Tibon, R., & Levy, D. A. (2015). Striking a balance: Analyzing unbalanced event-related potential data. *Frontiers in Psychology, 6*, 1–4. <https://doi.org/10.3389/fpsyg.2015.00555>
- Trauer, S. M., Andersen, S. K., Kotz, S. A., & Müller, M. M. (2011). Capture of lexical but not visual resources by task-irrelevant emotional words: A combined ERP and steady-state visual evoked potential study. *Neuroimage, 60*(1), 130–138, <https://doi.org/10.1016/j.neuroimage.2011.12.016>
- Van der Lee, R., Ellemers, N., & Scheepers, D. (2016). Mastering moral misery: Emotional and coping responses to intragroup morality (vs. competence) evaluations. *Cognition and Emotion, 30*(1), 51–65. <https://doi.org/10.1080/02699931.2015.1050357>
- Van der Lee, R., Ellemers, N., Scheepers, D., & Rutjens, B. T. (2017). In or out? How the perceived morality (vs. competence) of prospective group members affects acceptance and rejection. *European Journal of Social Psychology, 47*(6), 748–762. <https://doi.org/10.1002/ejsp.2269>
- Van der Toorn, J., Ellemers, N., & Doosje, B. (2015). The threat of moral transgression: The impact of group membership and moral opportunity. *European Journal of Social Psychology, 45*(5), 609–622.

<https://doi.org/10.1002/ejsp.2119>

- Van Lange, P. A., & Balliet, D. (2014). *Interdependence Theory*. American Psychological Association. <https://doi.org/10.4135/9781446201022.n39>
- Van Lange, P. A. M. (1999). The pursuit of joint outcomes and equality in outcomes. *Journal of Personality, 77*(2), 337–349.
- Van Nunspeet, F., Derks, B., Ellemers, N., & Nieuwenhuis, S. (2015). Moral impression management: Evaluation by an in-group member during a moral IAT affects perceptual attention and conflict and response monitoring. *Social Psychological and Personality Science, 6*(2), 183–192. <https://doi.org/10.1177/1948550614548076>
- Van Nunspeet, F., Ellemers, N., Derks, B., & Nieuwenhuis, S. (2014). Moral concerns increase attention and response monitoring during IAT performance: ERP evidence. *Social Cognitive and Affective Neuroscience, 9*(2), 141–149. <https://doi.org/10.1093/scan/nss118>
- van Prooijen, A. M., & Ellemers, N. (2015). Does it pay to be moral? How indicators of morality and competence enhance organizational and work team attractiveness. *British Journal of Management, 26*(2), 225–236. <https://doi.org/10.1111/1467-8551.12055>
- Wanous, J. P., Reichers, A. E., & Hudy, M. (1997). Overall job satisfaction: How good are single-item measures? *Journal of Applied Psychology, 82*(247–252).
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Wieser, M. J., Gerdes, A. B. M., Büngel, I., Schwarz, K. A., Mühlberger, A., & Pauli, P. (2014). Not so harmless anymore: How context impacts the perception and electrocortical processing of neutral

faces. *NeuroImage*, 92, 74–82. <https://doi.org/10.1016/j.neuroimage.2014.01.022>

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, 90(2), 245–271. <https://doi.org/10.1037/0033-2909.90.2.245>

Wilson, T., & Gilbert, D. (2003). Affective Forecasting. Knowing what to want (part 1). *Advances in Experimental Social Psychology*, 35, 345–411.

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials (ERPs) in studies of perception and attention. *Attention and Perceptual Psychophysiology*, 72(8), 1–29. <https://doi.org/10.3758/APP.72.8.2031.A>



Dutch Summary

Wanneer we anderen proberen aan te moedigen om hun gedrag te veranderen, geven we vaak feedback om onze visies mee over te brengen. Wanneer we positieve feedback aan anderen geven, bijvoorbeeld het complimenteren van een vriend die aan een goed doel doneert, wordt dit vaak graag ontvangen door de persoon die zulke positieve feedback krijgt. We worden gemotiveerd om hetzelfde gedrag te blijven vertonen. Het geven van negatieve feedback roept daarentegen vaak verschillende reacties op. In plaats van gemotiveerd te worden om op basis van de feedback ons gedrag te veranderen, wijzen we zulke feedback vaak af door een defensieve respons. Een dergelijke defensieve respons kan op verschillende wijzen tot uiting komen. Iemand die bekritiseerd is, kan bijvoorbeeld de feedbackzender verwijten dat diegene bewust wil kwetsen of dat diegene zichzelf goed wil laten voelen. Dit kan vooral gebeuren wanneer de zender van de feedback niet communiceert waarom zulke feedback wordt gegeven of – in andere woorden – wat de intenties waren om dergelijke feedback te geven.

Wanneer de zender de reden van de feedback niet communiceert biedt dit de bekritiseerde persoon veel gelegenheid om de feedback te negeren en ervan uit te gaan dat de zender negatieve intenties heeft, bijvoorbeeld om zichzelf in plaats van de ander beter te laten voelen. Bij het verklaren van het gedrag van de feedbackzender of het bepalen van diens intenties, kan de bekritiseerde persoon gebruik maken van verschillende soorten informatie. Bijvoorbeeld of de feedbackzender op de bekritiseerde lijkt of overeenkomsten vertoont zoals het aanhangen van dezelfde religie of het hebben van dezelfde nationaliteit (oftewel lid zijn van de in-groep, Tajfel & Turner, 1979). Als iemand als gelijke beschouwd wordt, kan men aannemen dat de feedbackzender goede bedoelingen heeft en de persoon in kwestie wil helpen te groeien of ergens beter in te worden (Hornsey et al, 2004). Negatieve feedback daarentegen, van mensen die geen overeenkomsten vertonen en dus als anders worden gezien door de bekritiseerde persoon (bijvoorbeeld omdat ze een andere religie aanhangen en daarom tot een uit-groep behoren), wordt vaak ontvangen als potentieel gevaar (Hornsey & Esposito, 2009).

Bovendien willen mensen niet dat anderen hun moreel twijfelachtige gedrag opmerken, omdat het signaleren van immoreel en grensoverschrijdend gedrag kan leiden tot sociale exclusie van de gemeenschappen waar zij toe behoren (Ellemers & van den Bos, 2012; Van der Lee et al., 2017). Daarom verdedigen mensen zich vaak wanneer ze bekritiseerd worden voor immoreel gedrag, bijvoorbeeld door hun morele falen te verbergen (Carlson et al., 2020; Gausel & Leach, 2011; Täuber & van Zomeren, 2013). Dit levert echter een probleem op. In veel van deze sociale gelegenheden waarin mensen kritiek van buitenstaanders ontvangen en waarin hun moraliteit wordt bekritiseerd, gaat de positieve kant van negatieve kritiek – aangemoedigd worden zichzelf te verbeteren – verloren. Bovendien zijn er veel situaties waarin buitenstaanders moeten wijzen op immoreel gedrag van groepen. Bijvoorbeeld wanneer in een groepssituatie onethisch gedrag genormaliseerd is, maar leden van de in-groep te bang zijn om zich hierover uit te spreken, omdat ze bang zijn te worden buitengesloten (e.g., Cavazza et al., 2014; Ellemers & van Nunspeet, 2020; Thau et al., 2015; Van der Lee et al., 2017).

Met het onderzoek in deze dissertatie was het mijn doel om een bijdrage te leveren aan deze kwesties door nieuwe manieren te vinden om anderen aan te moedigen om opener te staan voor feedback van mensen die anders zijn dan onszelf (dus van uit-groepsleden) en over gevoelige onderwerpen, zoals moreel gedrag. Ik heb twee hoofdvragen onderzocht: 1) Hoe gaan mensen om met het bekritiseerd worden om hun morele gedrag en hoe kunnen mensen zich meer openstellen voor zulke kritiek? 2) Reageren mensen anders op kritiek als het afkomstig is van een uit-groep in vergelijking met een lid van de in-groep – en als dit zo is – welke strategieën kunnen voorkomen dat dit gebeurt?

Om deze vragen te onderzoeken heb ik theorie, methodes en inzichten gebruikt uit de sociale psychologie en de neurowetenschap. Bestaand onderzoek naar kritiek tussen groepen en kritiek op iemands moraliteit is naar mijn weten alleen verricht op het gebied van de sociale psychologie (Esposito et al., 2013; Täuber et al., 2018; Täuber & van Zomeren, 2013). Specifieker gesteld betekent dit dat we enige kennis hebben over hoe mensen

bewust reageren op het bekritiseerd worden door uit-groepsleden en op het bekritiseerd worden op hun moraliteit. Echter, we weten niet of factoren van de sociale context (bijvoorbeeld het zijn van een lid van de uit-groep) ook van invloed zijn op meer *impliciete* processen die plaatsvinden voordat mensen zelfs maar bewust kunnen reflecteren op wat het met ze doet wanneer ze kritiek ontvangen. Op dit gebied bieden methoden uit de neurowetenschap uitkomst, omdat ze nieuwe inzichten kunnen bieden in de cognitieve processen die plaatsvinden voordat we in staat zijn bewust te reageren op een situatie (Ellemers & van Nunspeet, 2020). Neurowetenschappelijke meetmethoden kunnen bijvoorbeeld aantonen welk type kritiek onze aandacht krijgt en vasthoudt en welke kritiek niet. De combinatie van beide disciplines is daarom veelbelovend en biedt een uitbreiding van de kennis die al is gevonden in de sociaalpsychologische literatuur.

Samenvatting van empirische bevindingen

De resultaten laten zien dat er twee manieren zijn om de effectiviteit te vergroten van feedback uit de uit-groep over iemands moraliteit die eenvoudig in de praktijk kunnen worden toegepast. 1) Stel iemands competentie, in plaats iemands moraliteit, centraal wanneer feedback gegeven wordt en 2) communiceer van goede wil te zijn wanneer je feedback geeft. Beide strategieën verwijzen naar eenvoudige aanpassingen in hoe feedback wordt ge(re)framed en kan daarom door iedereen gebruikt worden die negatieve reacties wil voorkomen wanneer constructieve feedback gegeven wordt. Bovendien kunnen de strategieën algemeen worden toegepast in allerlei situaties tussen groepen en kan dit critici uit uit-groepen helpen hun stem te laten horen in situaties waarin doorgaans niet naar ze geluisterd wordt (bijv. wanneer ze feedback leveren).

Strategie 1: Focussen van kritiek op competentie (in plaats van moraliteit)

In hoofdstuk 2 laat ik zien dat door leden van de uit-groep (vs. in-groep) op het werk geuite kritiek minder effectief was. Deelnemers accepteerden minder vaak dergelijke kritiek en pasten hun gedrag minder vaak aan op basis van deze kritiek. Deze reactie wordt verklaard doordat kritiek ontvangen door

de uit-groep ervoor zorgde dat deelnemers meer defensief reageerden en minder gemotiveerd waren hun gedrag te verbeteren. Echter, ik ontdekte ook dat deelnemers minder negatief reageerden op kritiek die zich richtte op iemands competentie in plaats van diens moraliteit, ongeacht of deze kritiek afkomstig was van iemand van de in-groep of uit-groep. Deelnemers die deze vorm van kritiek ondergingen, waren minder defensief, gemotiveerder om hun gedrag te verbeteren en gaven aan vaker hun gedrag veranderd te hebben in overeenstemming met de gegeven kritiek.

In hoofdstuk 3 constateerde ik dat feedback van de in-groep (versus de uit-groep) van grote invloed was op het proces van informatieverwerking van de deelnemers, hun affectieve reacties op het ontvangen van de feedback, en op hun reflecties op de feedback. Vergeleken met feedback van de uit-groep waren de deelnemers op het moment van het ontvangen van feedback vaker emotioneel geraakt door feedback van de in-groep. Dit valt ook terug te zien in het verwerken van de informatie van boodschappen uit de feedback: aanvankelijk besteedden deelnemers meer aandacht aan de boodschap die afkomstig was van feedbackzenders uit de in-groep (versus de uit-groep, zoals aangegeven door met EEG gemeten verhoogde P200-amplitudes).

In relatie tot de inhoud van de feedback vond ik meer bewijs voor de ineffectiviteit van morele feedback. Deelnemers besteedden minder aanhoudende aandacht aan feedbackberichten over hun moraliteit (vs. competentie, zoals aangegeven door verminderde P300- en LPP-amplitudes). Dit werd aangevuld met zelfrapportagebevindingen waaruit bleek dat deelnemers zich slechter voelden na het ontvangen van negatieve feedback over hun moraliteit (versus competentie). Ten slotte herinnerden deelnemers ook minder negatieve feedbackboodschappen wanneer deze gerelateerd waren aan hun moraliteit (vs. competentie).

Strategie 2: Het communiceren van de intentie om te helpen

In hoofdstuk 4 toonde ik aan dat het overbrengen van de intentie om te helpen bij het leveren van morele kritiek kan voorkomen dat deelnemers verdere negatieve aannames doen over feedbackzenders uit de uit-groep (versus de in-groep), en dat het de ervaren mate van eerlijkheid van zulke

feedback vergroot. Bovendien leek het communiceren van behulpzaamheid de dreiging te verkleinen die mensen mogelijk ervaren bij het krijgen van negatieve feedback over hun moraliteit. Dit werd aangegeven door een verminderde oplettendheid wanneer deelnemers cognitief negatieve feedbackboodschappen over hun moraliteit verwerkten (zoals aangegeven door met EEG gemeten verminderde P200-amplitudes).

Het gecombineerde empirische bewijs van het onderzoek in deze dissertatie toont in hoofdzaken het volgende aan: 1) Mensen reageren negatiever op kritiek op hun gedrag als het afkomstig is van iemand die deel uitmaakt van een uit-groep dan van iemand die deel uitmaakt van de in-groep. 2) Iemand bekritisieren vanwege zijn moraliteit is minder effectief dan iemand bekritisieren op zijn competenties, en 3) wanneer critici de effectiviteit willen vergroten van hun kritiek op iemands moraliteit, dan zouden ze hun behulpzame intenties moeten overbrengen. De onderzoeksresultaten hebben directe sociale en praktische implicaties voor de hedendaagse politiek, media en journalistiek, ons dagelijks leven in diverse samenlevingen en op de werkvloer. Het toepassen van de twee in deze dissertatie geteste strategieën in alle hierboven genoemde gebieden van het dagelijks leven kan critici helpen om hun stem te laten horen wanneer ze de moraliteit van anderen willen bekritisieren.

Conclusie

Door methoden uit verschillende vakgebieden te combineren om te onderzoeken hoe mensen meer open kunnen staan voor morele kritiek die wordt geleverd door leden van de uit-groep, breidt het onderzoek in deze dissertatie eerder onderzoek uit dat deze vragen uitsluitend had onderzocht vanuit sociaalpsychologisch perspectief (Esposito et al., 2013; Gausel et al., 2012; Gausel & Leach, 2011; Hornsey & Imani, 2004; Täuber & van Zomeren, 2013; Van der Lee et al., 2016) of vanuit neurowetenschappelijk perspectief (e.g., Schindler et al., 2014, 2015, 2020). De bevindingen in de drie empirische hoofdstukken tonen aan dat bekritiseerd worden door een lid van een uit-groep (versus de in-groep) en bekritiseerd worden op iemands moraliteit vaak ineffectief is. Echter, de bevindingen tonen ook aan dat er twee strategieën zijn die kunnen worden ingezet om de effectiviteit te vergroten.

Critici uit een uit-groep zouden zich bij het overbrengen van een kritische boodschap kunnen richten op iemands incompetenties (in plaats van hun immoraliteit) of ze zouden hun behulpzame intenties duidelijk kunnen maken wanneer ze anderen bekritisieren. Beide strategieën kunnen een meer open houding bewerkstelligen tegenover morele kritiek bij de persoon die de kritiek ontvangt.

About the author

Inga Katharina Rösler was born on the 2nd of November in Hamburg, Germany. In 2013 she obtained her bachelor's degree from Leuphana University Lüneburg with a major in Business Psychology and a minor in Economics. After working for a year as a research assistant at Leuphana University, she started a research master's in Psychology at Leiden University, the Netherlands, and graduated in 2016 cum laude. During this time, she also completed an internship at the Max Planck Institute for Human Development in Berlin, Germany, and worked as a research assistant at Leiden University. She then started her PhD-project under the supervision of prof. dr. Naomi Ellemers and dr. Félice van Nunspeet on the topics of morality and social identity, financed by a Spinoza grant from NWO that was awarded to Naomi Ellemers. As of November 2020, Inga works as a postdoctoral researcher at the University of Amsterdam investigating topics such as moralized stereotypes and bias in media communication.

KLI dissertation series

The “Kurt Lewin Institute Dissertation Series” started in 1997. The following dissertations have been published during the last two years. The complete list can be found on our website: <https://kurtlewininstituut.nl>

- 2020-01: Florian Wanders: Rebels, Renegades, and Robin Hoods: The Social-Hierarchical Dynamics Surrounding Norm Violators
- 2020-02: Marko Milovanović: Intrinsically Motivating Social Influence
- 2020-03: Simon Columbus: Subjective Interdependence and Prosocial Behaviour
- 2020-04: Annemijn Peters: When well begun is half done: How the adoption of sustainable energy technologies can lead to sustainable use of the technologies and other pro-environmental behaviours
- 2020-05: Josefine Geiger: Context matters: Three ways of how the context influences recycling behavior
- 2020-06: Lianne Aarntzen: Work-family guilt: A straitjacket keeping parents in traditional gender roles
- 2020-07: Mandy Tjew-A-Sin: Contact Comfort: Psychological Effects of Actual and Simulated Affectionate Touch
- 2020-08: Melissa Vink: Who brings home the bacon? How gender stereotypes straitjacket men and women into traditional relationships
- 2020-09: Jesús Manuel Mascareño Apodaca: Orchestrating innovation: How leaders affect creativity and innovation
- 2020-10: Tatiana Chopova: Doing good in business: Examining the importance of morality in business contexts
- 2020-11: Margarita Leib: (Dis)honesty in Individual and Collaborative Settings: A behavioral ethics approach
- 2020-12: Samantha Antusch: On how we experience ourselves as intentional agents. An examination of the role of intentional action in the sense of agency
- 2021-01: Mengchen Dong: Understanding Moral Hypocrisy: Behavioral

Antecedents and Social Consequences

- 2021-02: Daniel Sloot: Bringing Community and Environment Together: The role of community environmental initiatives in sustainability transitions
- 2021-03: Burkhard Wörtler: Enhancing Blended Working Arrangements and Individual Work Performance
- 2021-04: Frank Doolaard: Social exclusion - put into context
- 2021-05: Laurens van Gestel: The psychology of nudging - An investigation of effectiveness and acceptability
- 2021-06: Ruddy Faure: Implicit partner evaluations: How they form and affect close relationships
- 2021-07: Rosabelle Illes: Between a rock and a hard place: Challenges, strategies and resolution of value conflict mediation
- 2021-08: Nadja Zeiske: The Intrinsic Route to Pro-Environmental Behaviour
- 2021-09: Katherina Tatiana Alvarez Durnov: The Psychological Impact of Receiving Aid
- 2021-10: Peikai Li: Looking on the Bright and Dark Sides of Working Life: Appraisals of Work Characteristics and Employee Outcomes
- 2022-01: Iris van Sintemaartensdijk: Burglary in virtual reality
- 2022-02: Lu Liu: Public participation in decision making on sustainable energy transitions
- 2022-03: Rabia Kodapanakkal: The role of tradeoffs and moralization in the adoption of big data technologies
- 2022-04: Elissa El Khawli: Why, When, and How Workers Regulate: A lifespan perspective on work design and emotion regulation at work
- 2022-05: Chantal van Anandel: Clinical Grade Differences Between Ethnic Minority and Majority Students: Institutional-, assessor and student-related factors
- 2022-06: Inga Rösler: Hear me out: How to create an open mind

towards moral criticism