Journal of Experimental Social Psychology 104 (2023) 104419

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Falling on deaf ears: The effects of sender identity and feedback dimension on how people process and respond to negative feedback – An ERP study☆

Inga K. Rösler [a,b,*], Félice van Nunspeet [b], Naomi Ellemers [b]

[a] University of Amsterdam, the Netherlands
[b] Utrecht University, the Netherlands

ARTICLE INFO

Keywords:
Feedback
ERP
Morality
Affective responses

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 1, $N = 194$), and additionally recorded an EEG in Study 2 ($N = 49$). In both studies, participants reported being more emotionally affected by negative feedback from ingroup compared to outgroup senders. Participants in Study 1 also reported to perceive feedback on their morality (vs. competence) as more negative. Complementing these findings, ERP results of Study 2 revealed greater preferential processing (i.e., increased P200) of feedback messages delivered by 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.

To regulate social behavior and to ensure norm adherence in groups, humans give feedback to each other (Ellemers, Pagliaro, & Barreto, 2013; Ellemers & van den Bos, 2012). How successful this feedback is in regulating behavior may depend on 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 can influence how people emotionally and behaviorally respond to the feedback (Hornsey, Oppes, & Svensson, 2002; Tajfel & Turner, 1979). Specifically, people perceive negative feedback from outgroup senders as less constructive than negative feedback from ingroup senders (Esposo, Hornsey, & Spoor, 2013; Hornsey et al., 2002) and are therefore more reluctant to accept such feedback (Hornsey, Trembath, & Gunthorpe, 2004; Rösler, Van Nunspeet, & Ellemers, 2021). 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 influence people's responses to feedback is the social dimension the feedback 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, Cuddy, & Glick, 2007; Judd, James-Hawkins, Yzerbyt, & Kashima, 2005). Past research has shown that people are less inclined to accept criticism of their morality than of their competence, because moral criticism is perceived as less constructive (Rösler et al., 2021), more harmful to the self-image (Pagliaro, Ellemers, Barreto, & Di Cesare, 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

☆ This paper has been recommended for acceptance by Dr Lasana Harris

* Corresponding author at: Department of Social Psychology, University of Amsterdam, Nieuwe Achtergracht129-B, 1018 WS Amsterdam, the Netherlands.
E-mail address: i.k.rosler@uva.nl (I.K. Rösler).

https://doi.org/10.1016/j.jesp.2022.104419
Received 25 November 2021; Received in revised form 9 September 2022; Accepted 12 September 2022
Available online 26 September 2022

revealing the receipt of such messages. 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). Specifically, using these Event-Related brain Potentials (ERPs), we examined participants' initial attentional deployment (i.e., P200, Kissler, Assadollahi, & Herbert, 2006; Trauer, Andersen, Kotz, & Müller, 2012), and higher-order cognitive functions such selective attention and working memory updating (i.e., P300, Polich, 2007), and sustained attention and decoding of affective meaning (i.e., LPP, Schupp et al., 2004; Schupp, Flaisch, Stockburger, & Junghöfer, 2006). 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).

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, Iffland, Schindler, Wabnitz, & Neuner, 2015; Schindler & Kissler, 2018; Schindler, Miller, & Kissler, 2020; 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, 2013; Lang, Bradley, & Cuthbert, 1997). 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. Attention to motivationally relevant stimuli may be indicated by two ERPs related to higher-order cognitive functions (Lang et al., 1997): the P300, associated with selective attention and working memory updating (Polich, 2007) and the LPP, associated with sustained attention and decoding of affective meaning (Schupp et al., 2004; Schupp et al., 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). Not only the P300 and LPP are modulated by source relevance, but also the P200, an ERP associated with initial attention deployment and lexical processing of feedback messages (i.e., P200, Kissler et al., 2006; Trauer et al., 2012). Schindler et al. (2020) demonstrated that when delivering feedback to participants coming from human lay persons and experts, this feedback induced deeper processing (i.e., greater P300 and LPP) of feedback messages compared to computer feedback. Moreover, expert feedback induced greater initial attention (e.g., greater P200) amplitudes than lay person feedback and computer feedback. These findings thus suggest that, besides contextual information, also the self-relevance of a stimulus can increase initial and sustained 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 as it may offer important cues on how to be a good group member and because ingroup sources are seen as having better motives than outgroup sources when delivering feedback (Ellemers et al., 2013; Hornsey et al., 2002). Secondly, feedback on one's morality is perceived as less relevant than feedback on one's competence because people perceive such feedback as harder to cope with and see less opportunity to improve their behavior (Rösler et al., 2021; Täuber & van Zomeren, 2013; Van der Lee et al., 2016). Research on social impression formation (e.g., Martijn, Spears, Van Der Pligt, & Jakobs, 1992) has demonstrated that negative information about a persons' morality is more diagnostic than negative information about one's competence (potentially because the latter behaviors are perceived as occurring more often in the real world, Mende-Siedlecki, Baron, & Todorov, 2013). Such heightened diagnosticity increases the threat that one's social reputation is harmed. This might happen, because, when people form impressions of others, more diagnostic information is more memorable than less diagnostic information (Gilron & Gutchess, 2012). Having other people witness an immoral action is therefore especially harmful for one's reputation because it may be remembered more by others. This is why people often get defensive when receiving feedback on their morality and perceive such feedback as less self-relevant (e.g., Sun & Goodwin, 2020; Täuber & van Zomeren, 2013).

Greater social relevance of ingroup sources and greater self-relevance of feedback on one's competence compared to one's morality may influence already the cognitive processing of such feedback messages by facilitating the attentional processing of ingroup and competence feedback. This is what we investigated with the current research. 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 initial and sustained attention to feedback messages from ingroup (vs. outgroup) members and to feedback on their competence (vs. morality). We extend prior research by combining insights from different theoretical research areas and literatures from social psychology and neuroscience.

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, Ellemers, & Barreto, 2011). That is, because following these recommendations can potentially secure respect and inclusion from fellow group members. Moreover, negative feedback from outgroups often makes receivers defensive because the feedback is perceived as less constructive than ingroup feedback and potentially harmful to the self or ingroup (Esposo et al., 2013; Hornsey et al., 2002, 2004; Rösler et al., 2021). 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 earlier mentioned ERPs related to attentional processing of motivationally relevant stimuli, the P200 (initial lexical encoding of emotional words (Kissler et al., 2006; Trauer et al., 2012), the P300 (selective attention and working memory updating, Polich, 2007), and the LPP (sustained attention and decoding of affective meaning, Schupp et al., 2004; Schupp et al., 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. Participants 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. 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, Bratslavsky, Finkenauer, & Vohs, 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; Carretié, 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., selective attention: P300, and sustained attention: 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, 2021; 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, Fiske, & Glick, 2008) and is central to how we see ourselves and are seen by others (Goodwin, Piazza, & Rozin, 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., 2021; 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, Amir, & Ariely, 2008; Shalvi, Gino, Barkan, & Ayal, 2015). This goes as far as misremembering past moral failures (Carlson, Maréchal, Oud, Fehr, & Crockett, 2020; Kouchaki & Gino, 2016), distancing the self when reflecting on past moral failures by emphasizing having changed since the occurrence (Stanley, Henne, & De Brigard, 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 & Esposo, 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., 2021; 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 perceive moral (vs. competence) feedback as less motivationally relevant, as reflected in decreased P300-, and LPP-amplitudes (Rösler et al., 2021; Täuber & van Zomeren, 2013.[1]

Moreover, we expected that the affective and behavioral measures complement these findings and based our predictions on previous social

[1] Based 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.

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., 2021; 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; Esposo et al., 2013; Hornsey et al., 2002, 2004; Pagliaro et al., 2011; Rösler et al., 2021).

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 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, conducted in the lab, we replicated the findings of Study 1. Moreover, complementing the self-report findings of Study 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.

5. Study 1

5.1. Methods

5.1.1. Design and participants

Participants performed a task in which they made monetary contribution decisions (see Instruments). After this, they received feedback on their behavior in form of character traits. The feedback came from ostensible ingroup members (i.e., fellow students) or outgroup members (i.e., professional trainees) and was supposedly based on their behavioral decisions. The feedback messages were presented in random order, were equated for valence (i.e., 50% positive, 50% negative), and were either referring to participants' morality or to 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. After this main task, participants additionally received an overall group-based negative feedback message that supposedly summarized the individual feedback from both the ingroup and the outgroup (see Fig. 1).

Sample size was determined before any data analyses. We based sample size calculations on previous research (Rösler et al., 2021) 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 $\eta_p^2 = 0.05$. 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 ($\alpha = 0.05$, $1\text{-}\beta = 0.80$) revealed a minimum sample of 154 participants using repeated measure ANOVAs (see Supplementary Materials for sensitivity analysis). 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 (for a more conservative approach excluding participants who failed one check see Supplementary Materials, all analyses hold). 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. Data

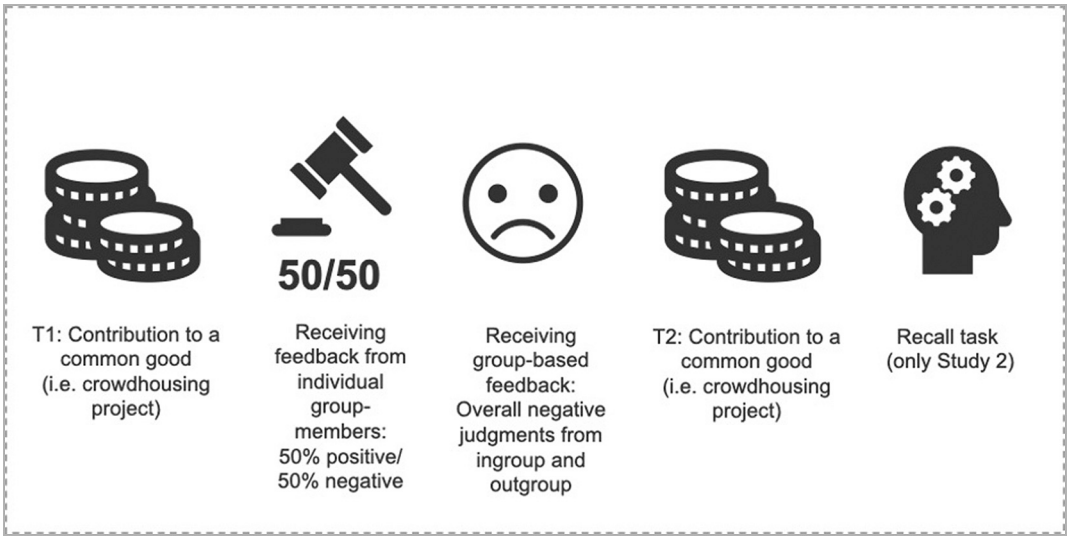

Fig. 1. Overview of experimental procedure.

and code for the experiments reported in this manuscript have been made available at https://osf.io/2953z/.

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 Fig. 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 main task, but we additionally aimed to investigate the affective responses to negative feedback. Therefore, both of these additional 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 min. After completion, participants were fully debriefed and thanked for taking part in the study.

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). We report all measures, manipulations, and exclusions.

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 and 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. Contrary to research in economics that often adds financial incentives to improve task performance, this incentive was aimed at ensuring variability in participant's contribution decisions and enhance credibility of receiving both negative and positive feedback on their behavioral decisions.

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 (250 ms), followed by a sender's face and background color alternating ingroup and outgroup membership (1000 ms), 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 50 % 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 50 % ingroup members (i.e., fellow students, indicated by a blue background color) and 50 % 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.

5.1.2.3. Group-based feedback.

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).[2] Additionally, because we aimed to investigate participants' responses to negative feedback, we presented them with a group-based feedback message. We explained that all delivered feedback messages were 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 (i.e., "Other participants, who are [students/trainees], were able to pick both negative and positive personality traits (related to your [competence/morality]) that they think are indicative of your character, based on the financial contributions you made in the beginning of the experiment. The students, on average, judged your [competence/morality] as: Negative (= [Incompetent/Immoral]): more negative [competence/moral] adjectives were chosen by the group of [students/trainees]

[2] 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.

compared to positive [competence/moral] adjectives").

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

5.1.2.4.1. Manipulation checks. To check our between-participants manipulation of dimension, we asked participants to indicate the nature of the feedback they had received. They 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, $\alpha = 0.90$, outgroup, $\alpha = 0.89$).

5.1.2.4.2. 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., 2021; Watson, Clark, & Tellegen, 1988). The basis of the current cluster of items was based on previous research (Rösler et al., 2021), 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 $\alpha = 0.90$, outgroup $\alpha = 0.92$), three items to the measure *positive emotions after group-based feedback* (e.g., "The feedback from the [students/trainees] made me feel confident", ingroup $\alpha = 0.79$, outgroup $\alpha = 0.85$), 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 $r[192] = 0.75$, $p < .001$, outgroup $r[192] = 0.82$, $p < .001$).

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 et al., 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 $\alpha = 0.79$, outgroup $\alpha = 0.78$).

5.2. Results and discussion

In line with the timeline of our paradigm (see Fig. 1), we first report manipulation checks, then the self-report results, and finally the behavioral results.

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, −0.65]. 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 = [0.53, 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).

5.2.2. Self-report and behavioral results

5.2.2.1. Perceived valence of feedback from individuals[3]. 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 = negative[2]). 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 ($M = 4.15$, $SD = 0.73$), $t(193) = 2.96$, $p = .003$, CI = [0.05, 0.26]. 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, $F(1, 192) = 6.37$, $p = .012$, $\eta_p^2 = 0.03$. 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 ($M = 4.29$, $SD = 0.76$) compared to on their competence ($M = 4.03$, $SD = 0.68$, between-participants factor). There were no other effects, $Fs < 1$.

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 0.016 (Bonferroni correction, 0.05/3) to correct for multiple comparisons. Results revealed an effect of sender's group-membership at the multivariate level, Pillai's Trace = 0.11, $F(3, 190) = 7.43$, $p < .001$, $\eta_p^2 = 0.11$. 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 ($M = 3.48$, $SD = 1.66$), $F(1, 192) = 20.03$, $p < .001$, $\eta_p^2 = 0.09$, at the univariate level. There were no other effects, $Fs < 1$. 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, $\eta_p^2 = 0.12$. There were no other effects, $Fs < 1$.

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. Participants contributed 30% of all possible contributions (i.e., 30€) for both T1 ($M = 10.17$, $SD = 5.70$) and T2 ($M = 9.99$, $SD = 6.73$, for distributions across factors see Supplementary Materials). We then used averaged contributions as dependent variable 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, $Fs < 1$.

[3] Some participants raised doubts about whether they received feedback in real-time. Using a more conservative sample by excluding them yielded the same results (see Supplementary Materials).

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.

6. Study 2

With Study 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, 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 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 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, 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.

6.1. Method

6.1.1. Design and participants

The experimental design was the same as in Study 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 was determined before any data analyses. 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, $\eta_p^2 = 0.45$, early LPPs (400–650 ms), $\eta_p^2 = 0.31$ and $\eta_p^2 = 0.23$, and on the late LPP (650–900 ms), $\eta_p^2 = 0.18$. Sample size calculation with G*power ($\alpha = 0.05$, $1\text{-}\beta = 0.80$) revealed a minimum sample size between 14 and 40 for the main effects of sender and valence using repeated measure ANOVAs (see Supplementary Materials for sensitivity analysis).

We tested 49 right-handed Dutch students ($M_{age} = 21.20$, $SD_{age} = 2.15$, two participants not indicating age, 39 female) with no history of neurological or psychiatric problems, and normal to corrected vision. Right-handedness was a requirement for taking part in the experiment and we administered this with a short screening questionnaire. 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.

6.1.2. Procedure

The procedure was the same as in Study 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 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 min. After completion, participants were fully debriefed, thanked, and compensated with course credit or €18.

6.1.3. Instruments and stimuli

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

6.1.4. 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 Fig. 2). Each trial thus started with a fixation cross (jittered duration, 350–750 ms), followed by a sender's face and background color indicating 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).

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 ($N = 100$, $M_{age} = 26.4$ 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

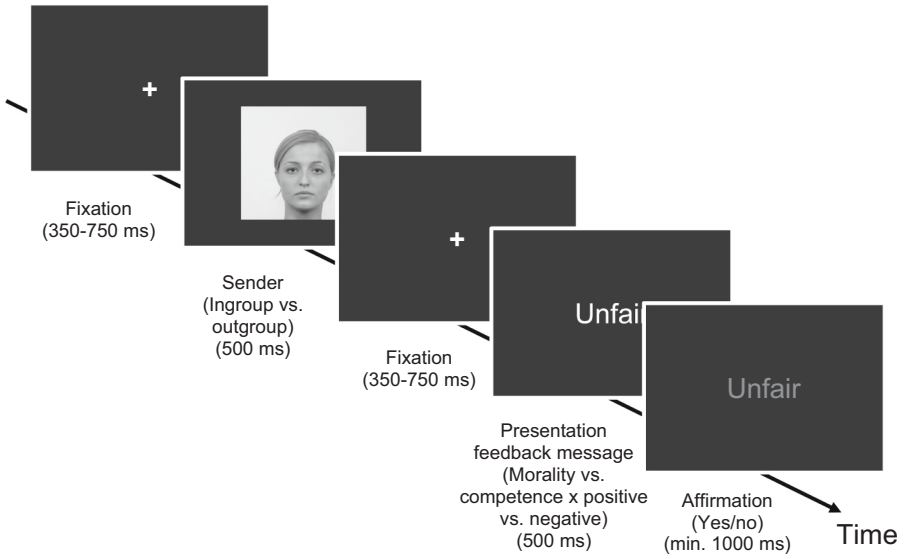

Fig. 2. Example of trial feedback-phase.

(see Table 1).

6.1.5. 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 1. Specifically, for Study 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 $Fs < 1$, $ps \geq 0.896$). 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).

6.1.6. 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 1.[4]

6.1.7. 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. 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 kΩ. 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, Coles, & Donchin, 1983), filtered (0.01-30 Hz), 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).

6.1.8. EEG analyses

Below we detail how we determined the time-windows and localization of our ERP components of interest, based on previous literature. For an alternative approach, which applies temporal-spatial principal component analysis (PCA) to the ERPs to determine their components, please see the Supplementary Materials.

6.1.8.1. P200. P200 peaks were largest at fronto-central electrodes FCz, Cz, and CPz, consistent with prior studies (Herbert, Kissler, Junghöfer, Peyk, & Rockstroh, 2006; León, Díaz, de Vega, & Hernández, 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.

6.1.8.2. 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 and 400 ms 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.

6.1.8.3. LPP. Consistent with prior studies (Klein et al., 2015; Schindler

[4] 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 1 for comparability.

Table 1
Means, Standard deviations, and F-statistics for trait stimuli per condition for Study 2.

	Morality negative ($N = 16$)	Morality positive ($N = 16$)	Competence negative ($N = 16$)	Competence positive ($N = 16$)	$F(1, 60)$
How negative (1) or positive (9) is this trait?	2.39 (0.55)	7.39 (0.67)	2.68 (0.55)	7.49 (0.47)	Morality vs. Competence: $F = 1.80$ Positive vs. Negative: $F = 1201.18$*** Interaction: $F < 1$
Please indicate to what extent this adjective refers to someone being a(n) (im)moral or (in)competent person (1 = morality, 9 = competence)	3.12 (0.87)	3.37 (0.76)	6.58 (1.07)	7.08 (0.82)	Morality vs. Competence: $F = 260.93$*** Positive vs. Negative: $F = 2.85$ Interaction: $F < 1$
How intense do you find this word (1–9)?	5.72 (0.46)	5.71 (0.55)	5.43 (0.53)	5.70 (0.26)	Morality vs. Competence: $F = 1.68$ Positive vs. Negative: $F = 1.23$ Interaction: $F = 1.46$
How familiar are you with this word (1–9)?	7.50 (0.36)	7.59 (0.51)	7.53 (0.32)	7.75 (0.36)	Morality vs. Competence: $F < 1$ Positive vs. Negative: $F = 2.38$ Interaction: $F < 1$
Frequency (million)	21.60 (66.56)	17.67 (43.11)	14.09 (30.83)	6.50 (8.96)	Morality vs. Competence: $F < 1$ Positive vs. Negative: $F < 1$ Interaction: $F < 1$
Length	10.38 (3.10)	8.69 (2.39)	8.75 (2.86)	9.13 (2.33)	Morality vs. Competence: $F < 1$ Positive vs. Negative: $F < 1$ Interaction: $F = 2.35$

Note: The table displays results of a pre-test, in which we asked participants to rate character traits. Each trait belonged to one of the relevant experimental conditions (i.e., negative morality traits, positive morality traits, negative competence traits, positive competence traits). Participants were asked to rate these traits for their valence (i.e., How negative (1) or positive (9) is this trait?), their dimension (i.e., Please indicate to what extent this adjective refers to someone being a (im)moral or (in)competent person (1 = morality, 9 = competence)), their associated arousal, and their familiarity with the trait. Each trait was rated by at least 20 participants. Frequency and length for each trait were taken from the database of Dutch word frequencies 'SUBTLEX-NL'. As intended, traits ratings only differed between the valence and dimension category (i.e., our experimental conditions). ***$p < .001$.

& Kissler, 2018; Schupp et al., 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 and 700 ms 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.

6.1.9. Behavioral and self-report results

6.1.9.1. 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 = [0.70, 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).

6.1.9.2. Perceived valence of feedback from individuals. A one-sample t-test against the midpoint of the scale (i.e., 4, 1 = positive, 7 = negative[2]) showed that participants perceived the feedback as more negative than positive, $M = 4.33$, SD = 0.75, $t(48) = 3.05$, $p = .004$, CI = [0.11, 0.54], replicating the results of Study 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, $Fs \leq 3.61$, $ps \geq 0.063$.

6.1.9.3. Affective and defensive response to negative group-based feedback[5]. Also in line with Study 1, a RM MANOVA (Sender's group-membership [within]; negative, positive, and moral emotions as dependent variables; Bonferroni correction, $\alpha = 0.05/3$) showed that participants reported stronger negative emotions after receiving an

[5] 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.

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 = 0.43, $F(3, 46) = 11.55$, $p < .001$, $\eta_p^2 = 0.43$, univariate: $F(1, 48) = 35.05$, $p < .001$, $\eta_p^2 = 0.42$. As in Study 1, the other affective responses (i.e., positive and moral) did not show any univariate effects ($Fs \leq 1.85$, $ps \geq 0.180$).

Finally, we also replicated the findings of Study 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 ($M = 5.10$, $SD = 0.82$) than from the outgroup ($M = 4.95$, $SD = 0.95$), $t(48) = 2.07$, $p = .044$, CI = [0.004, 0.29].

6.1.10. Event-related brain potentials (ERPs)

6.1.10.1. P200. 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 ($M = 5.57$, $SE = 0.30$) compared to an outgroup sender ($M = 5.16$, $SE = 0.29$), $F(1, 40) = 7.31$, $p = .010$, $\eta_p^2 = 0.16$ (see Fig. 3). Additionally, we found a three-way interaction effect between electrode, sender's group-membership, and valence, $F(2, 80) = 4.35$, $p = .027$, $\eta_p^2 = 0.10$. 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: $F(1, 40) = 4.36$, $p = .043$, $\eta_p^2 = 0.10$, Cz: $F(1, 40) = 8.95$, $p = .005$, $\eta_p^2 = 0.18$, and CPz: $F(1, 40) = 7.26$, $p = .010$, $\eta_p^2 = 0.15$. In addition, there was a two-way interaction between sender's group-membership and valence on FCz, $F(1, 40) = 4.64$, $p = .037$, $\eta_p^2 = 0.10$. When viewing negative feedback, the P200 was significantly larger when it was delivered by ingroup senders ($M = 6.18$, $SE = 0.34$) compared to outgroup senders ($M = 5.31$, $SE = 0.38$), $p = .006$. 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 ($M = 6.18$, $SE = 0.34$) compared to positive feedback ($M = 5.54$, $SE = 0.39$), $p = .037$. 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 ($M = 5.65$, $SE = 0.32$), $F(1, 40) = 23.06$, $p < .001$, $\eta_p^2 = 0.37$. There were no other effects, $Fs \leq 2.89$, $ps \geq 0.079$.

6.1.10.2. P300. P300-amplitudes in response to feedback messages were larger when the feedback concerned participant's competence ($M = 1.64$, $SE = 0.25$) compared to their morality ($M = 1.35$, $SE = 0.23$), $F(1, 40) = 4.97$, $p = .032$, $\eta_p^2 = 0.11$ (see Fig. 4). Interestingly and similar to the results on the P200, this main effect was qualified by an interaction effect between feedback dimension and valence, $F(1, 40) = 4.62$, $p = .038$, $\eta_p^2 = 0.10$. When viewing negative feedback, the P300 was larger when feedback concerned participant's competence ($M = 1.85$, $SE = 0.23$) as compared to their morality ($M = 1.32$, $SE = 0.25$), $p = .003$. 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 ($M = 1.85$, $SE = 0.23$) compared to positive messages ($M = 1.42$, $SE = 0.24$), $p = .006$, $\eta_p^2 = 0.18$. 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 ($M = 1.85$, $SE = 0.27$), $F(2, 80) = 35.39$, $p < .001$, $\eta_p^2 = 0.47$. There were no other effects, $Fs \leq 2.79$, $ps \geq 0.102$.

6.1.10.3. LPP. LPP-amplitudes were larger when feedback concerned participant's competence ($M = 1.01$, $SE = 0.23$) compared to their morality ($M = 0.72$, $SE = 0.22$), $F(1, 40) = 12.17$, $p = .001$, $\eta_p^2 = 0.23$ (see Fig. 4). 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 ($M = 1.15$, $SE = 0.24$) as compared to their morality ($M = 0.75$, $SE = 0.24$), $p = .026$. This difference was not present for positive feedback, $p = .275$. There were no effects of valence for feedback on morality and competence, $ps \geq 0.165$. There was a main effect of electrode with the largest amplitudes at Cz ($M = 1.10$, $SE = 0.25$), $F(4, 160) = 6.37$, $p = .005$, $\eta_p^2 = 0.14$. There were no other effects, $Fs \leq 1.78$, $ps \geq 0.190$.

6.1.10.4. Validation ERP effects: Recall task. In line with our results for the P300 and LPP which suggested that participants perceived feedback on their morality less motivationally relevant than feedback on their competence, a RM ANOVA (Feedback dimension [within], Feedback valence [within]) revealed a main effect of dimension on recalled feedback messages, $F(1, 48) = 8.77$, $p = .005$, $\eta_p^2 = 0.15$ (see Fig. 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 [$M = 2.36$, $SE = 1.55$]), $F(1, 48) = 3.44$, $p = .070$, $\eta_p^2 = 0.07$. There was, however, a significant interaction between dimension of the feedback and valence, $F(1, 48) = 4.84$, $p = .033$, $\eta_p^2 = 0.09$. For negative feedback, messages related to competence were recalled significantly more often ($M = 2.88$, $SE = 0.19$) than messages related to morality ($M = 1.84$, $SE = 0.21$), $p < .001$. For positive feedback, there was no significant difference for dimension, $p = .283$. For feedback on participant's morality, positive feedback was recalled more often ($M = 2.50$, $SE = 0.22$) than negative feedback ($M = 1.84$, $SE = 0.21$), $p = .004$. For feedback on participant's competence, there was no

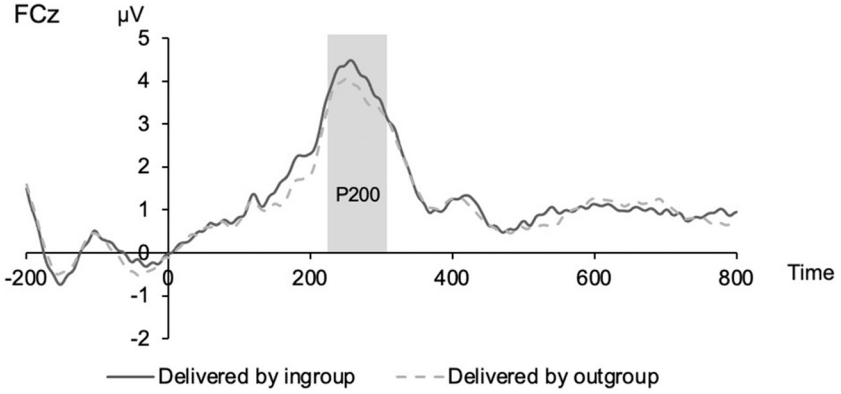

Fig. 3. ERP-waveforms at fronto-central midline electrode FCz for feedback messages delivered by ingroup and outgroup senders (grand averages, rather than peak amplitude as reported in result section).

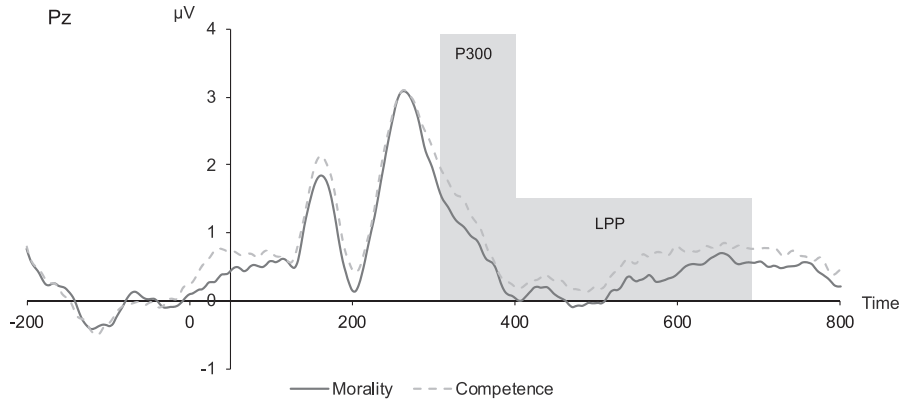

Fig. 4. ERP-waveforms at parietal midline electrode Pz for feedback messages (i.e., character traits) on participant's morality and competence.

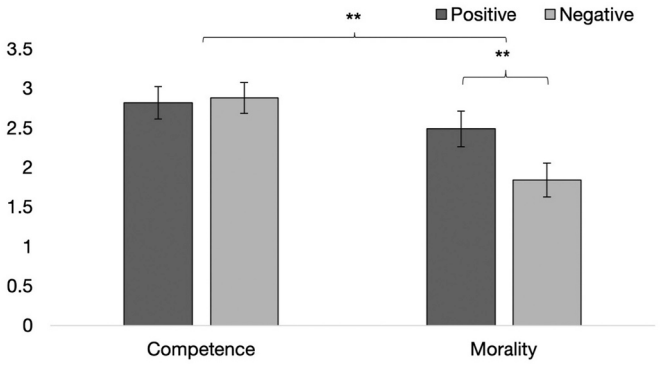

Fig. 5. Averages of recalled feedback messages on recall task per condition (Feedback Dimension x Feedback Valence).

difference for valence, $p = .801$. Thus, negative feedback messages on participants' morality were recalled the least.

6.2. 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, Wieser, Gerdes, Mühlberger, & Pauli, 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, Ellemers, Fiske, Koch, & Yzerbyt, 2021) can modulate the cognitive processing of feedback messages in a similar fashion. 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; Esposo et al., 2013; Hornsey et al., 2002, 2004; Pagliaro et al., 2011; Rösler et al., 2021). 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., 2021).

Extending past research and as expected, participants in Study 1 perceived feedback on their morality as more negative than feedback on their competence. Complementing this finding, participants in Study 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, 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.

6.2.1. 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, Wegrzyn, Steppacher, & Kissler, 2015, Schindler, Miller and Kissler, 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., 2021), 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) and different social group-memberships (Study 1: Students/non-students, Study 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).

Additional analyses add more nuance to this finding. When using the degree of identification with a feedback sender, rather than their group-membership to predict participants' negative assumptions about the sender, lower identification for both ingroup and outgroup senders was associated with more defensiveness (Study 1: ingroup: $r(192) = -0.41$, $p < .001$, outgroup: $r(192) = -0.32$, $p < .001$; Study 2: ingroup: $r(47) = -0.30$, $p = .037$, outgroup: $r(47) = -0.34$, $p = .018$). This is in line with our initial prediction and previous research (Hornsey et al., 2002). It shows that even though participants seem to be more negatively affected by negative feedback coming from ingroup (vs. outgroup) members, they are less defensive towards such feedback the more they identify with either.

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.

6.2.2. *Valence of and feedback dimension addressed in feedback*

In line with previous research (Carretié 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 1). In Study 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 1), remembered morality (vs. competence) messages about themselves less often (Study 2), and remembered (or reported) negative feedback on their morality the least (Study 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).

A limitation of combining behavioral, self-report, and neuroscientific measures is that each of them has different experimental design requirements to produce reliable data. For self-report findings, many participants are needed to achieve sufficient power. For neuroscientific findings, many trials are needed to achieve a good signal-to-noise ratio. Being aware of this trade-off, we designed Study 1 and Study 2 accordingly, testing fewer participants in Study 2 but increasing the trials. However, this implies that we may have less power for our self-report findings in Study 2. Indeed, we did not replicate the finding that participants perceived feedback as more negative when it concerned their morality as compared to their competence and the effect of sender's group-membership on defensiveness showed lower significance than in Study 1. Not replicating this particular self-report finding may be related to the increased trial number in Study 2 and the change in design. In Study 2, feedback dimension was manipulated within-participants (in Study 1 between-participants) and participants were presented with 448 feedback messages, as compared to 32 in Study 1. This might have made it more difficult to distinguish between moral and competence feedback messages on a conscious level. However we do see differences in the cognitive processing as indicated by differing ERP-amplitudes. Lower significance for the group effect, on the other hand, is most likely related to testing much fewer participants in Study 2 as compared to Study 1. Nevertheless, we had anticipated such lower power and believe that the benefits of combining self-report, behavior, and neuroscientific measures outweigh such potential concerns.

Previous research on how time pressure in economic decision-making games (Rand, Greene, & Nowak, 2012), and cooperative vs. competitive contexts (Wittmann et al., 2016), affects how people make decisions and evaluate their behavior, may offer interesting new avenues for future research in relation to our current findings. Rand et al. (2012) demonstrated that under time pressure, participants made more cooperative than selfish choices—as compared to the more selfish (rather than cooperative) choices they made when they had more time to decide. Future research may investigate whether constricting participants' time to make money contributions to a common good also influences their decisions, and how they then process feedback from others on their (moral) character—as was the case in the current research. Other research has shown that participants use the performance of others for the evaluation of their own performance and that this integration of self- and other-evaluation is influenced by whether they perceive the context as cooperative or competitive (Wittmann et al., 2016). Future research might examine how the processing of feedback received from others on one's (moral) character is affected when the recipient of the feedback considers the context a competitive or a cooperative one. This may reveal important conditions that need to be considered when one wants their moral feedback (and moral criticism in particular) to be heard.

6.3. *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.

Author note

This research was supported by the NWO-Spinoza prize awarded to Naomi Ellemers by the Netherlands Organization for Scientific Research (NWO). The authors declared no conflicts of interest.

Data availability

The link to the data/code is included in the manuscript

Acknowledgements

We would like to thank Tiago O. Paiva for offering help with the PCA analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2022.104419.

References

Abele, A. E., & Wojciszke, B. (2007). *Agency and communion from the perspective of self versus others.*. https://doi.org/10.1037/0022-3514.93.5.751

Abele-Brehm, A., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. *Psychological Review, 128*(2), 290–314. https://doi.org/10.1037/rev0000262

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*. Rinehart, Wins: Holt.

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*(I), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

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

Brañas-Garza, Bucheli, Paz Espinosa, & García-Muñoz. (2013). Moral cleansing and moral licenses: Experimental evidence. *Economics and Philosophy*, 199–212. https://doi.org/10.1017/S0266267113000199

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

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., 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.

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

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

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*. UK: Milton Park.

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

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., & 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 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*. https://doi.org/10.1177/0963721420951584

Esposo, 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

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

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

Gilron, R., & Gutchess, A. H. (2012). Remembering first impressions: Effects of intentionality and diagnosticity on subsequent memory. *Cognitive, Affective, & Behavioral Neuroscience, 12*(1), 85–98.

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

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., & Esposo, 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., 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., 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

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

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

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, & 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

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, P. 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.

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

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.

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

Mende-Siedlecki, P., Baron, S. G., & Todorov, A. (2013). Diagnostic value underlies asymmetric updating of impressions in the morality and ability domains. *Journal of Neuroscience, 33*(50), 19406–19415.

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
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
Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b John. *Clinical Neurophysiology, 118*(10). https://doi.org/10.1016/j.clinph.2007.04.019.Updating
Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature, 489*(7416), 427–430. https://doi.org/10.1038/nature11467
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*(3), 597–609. https://doi.org/10.1002/ejsp.2764
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
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
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
Strohminger, 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*, 33–47. https://doi.org/10.1016/S0065-2601(05)37005-5. January 1979.
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
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
Trauer, S. M., Andersen, S. K., Kotz, S. A., & Müller, M. M. (2012). 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 Toorn, Ellemers, & Doosje. (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 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
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
Wittmann, M. K., Kolling, N., Faber, N. S., Scholl, J., Nelissen, N., & Rushworth, M. F. S. (2016). Self-other mergence in the frontal cortex during cooperation and competition. *Neuron, 91*(2), 482–493.
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