van Gils, S., Otto, T., & Dinartika, N. (2020). Better together? The neural response to moral dilemmas is moderated by the presence of a close other. *Journal of Neuroscience, Psychology, and Economics, 13*(3), 150–163. https://doi.org/10.1037/npe0000126

Better together? The neural response to moral dilemmas is moderated by the presence of a close other

Suzanne van Gils*[a,b], Tobias Otto*[b], Niken Dinartika[b]

[a] BI Norwegian Business School

[b]Maastricht University

*The first authorship is shared by Suzanne van Gils and Tobias Otto

Author note: We thank Hans IJzerman and James Coan for their helpful comments early development of our study.

This research was funded by an internal grant of the Maastricht Brain Imaging Centre, project number F0006B10, awarded to Suzanne van Gils and Tobias Otto.

Better together? The neural response to moral dilemmas is moderated by the presence of a close other

Abstract

We investigated the modulation of neural and behavioral responses to moral dilemmas by the physical presence of a close friend. We argue that the presence of a close other not only changes the moral response, but also the process of decision making, something that can only be discovered by combining insights from social and cognitive psychology. Our participants rated the acceptability of sacrificing ingroup members to save outgroup members and vice versa while being alone or in the presence of a close other. We obtained behavioral and fMRI data from a within-participant fMRI study (N=17, native Dutch women). The behavioral data replicated classical identity theory with regard to higher acceptability to sacrifice the outgroup (vs. the ingroup), but did not show any differences when deciding alone or in the presence of the friend. The imaging results did not reveal main and interaction effects in our hypothesized brain areas. Exploratory analysis however revealed an interaction effect in a region previously found to be related to guilt and shame (SFG), such that reactions to the sacrifice of the outgroup evoked increased activation when being together with a close other.

Better together? The neural response to moral dilemmas is moderated by the presence of a close other

Social relationships are at the core of how people understand the world (Baumeister & Leary, 1995; Beckes & Coan, 2011). This social understanding plays a major role in moral decisions as well. As stated by Haidt (2001), moral systems provide a set of rules that guide people's behavior towards others and make social life possible. Recent studies discuss how the moral rules people deem relevant depend on the type of interpersonal relationship they have with the other person (Giessner & Van Quaquebeke, 2010; Rai & Fiske, 2012). In addition, people rely on and adapt to the social context to maintain their self-perception as a moral person (Rom & Conway, 2018). Thus, the idea that morality is intrinsically social is increasingly recognized.

Yet, past social psychological research has predominantly studied moral decision making in settings that ask participants to make decisions while they are seated alone in an isolated laboratory cubicle. If we accept morality as an inherently social phenomenon (Haidt, 2008; Rom & Conway, 2018), the findings from the literature using single-person paradigms may not accurately reflect moral decision making in its full scope for two reasons. Firstly, participants provide less extreme responses when making moral decisions alone outside their social context (Wilder & Shapiro, 1991), and decisions made outside a social context may not be influenced by concerns how to represent the self to others (Rom & Conway, 2018; Rom, Weiss, & Conway, 2017). Secondly, recent research shows neurological differences between situations in which participants decide alone or together (Beckes & Coan, 2011), suggesting a fundamental difference in decision making between these situations. Following this research, we argue the way in which a moral decision is reached, and the processes involved in the decision, may be fundamentally different depending on whether the decision was made alone or in a social context. Thus, we set out to combine insights from social and cognitive

psychology to trace the process of moral decision making in the presence of a close other versus alone.

The trolley dilemma

One of the most used paradigms in experimental social and cognitive morality research (cf. Cikara, Farnsworth, Harris, & Fiske, 2010; Conway & Gawronski, 2013) is the so-called trolley paradigm. In this dilemma, participants are asked whether they want to save the five people that would be run over by a runaway trolley by switching a lever that reroutes the trolley and sacrifices one person on the alternative track (Foot, 1967; Thomson, 1986). The dilemma consists of the tradeoff between the different people on the railway tracks, on the one hand, and of the need to overcome their resistance to make a decision to commit to a harmful act in order to save the potential victims on the other hand. Neural research shows that this dilemma evokes both affective responses and engages cognitive processes (Boccia et al., 2016; Greene, 2001).

Many studies investigating moral decision making have combined the trolley dilemma with the social context in terms of group membership (e.g., Friesdorf, Conway, & Gawronski, 2015). In these studies, group membership helps the decision makers to make a distinction between “us”, the ingroup, and “them”, the outgroup (Cikara, Van Bavel, Ingbretsen, & Lau, 2017; Fiske & Taylor, 1991). According to social identity theory, categorizing oneself as a member of a group promotes a feeling of belonging to the group, as well as a perception that the members of one’s group are superior to those of other groups (Hogg & Turner, 1987; Tajfel, Turner, & Turner, 1979). Group identification motivates people to behave positively towards the members of the group, and to adhere to the group norms (Abrams & Hogg, 2010; Hogg & Hains, 1996). Applied to the trolley dilemma, this social identity process translates into a strong rejection to sacrifice the ingroup and, consequently, a higher willingness to sacrifice outgroup members to save the ingroup (Conway & Gawronski, 2013).

Research in the neural domain that employs the same paradigm shows that social group memberships do affect neural processes when making decisions about the sacrifice of ingroup and outgroup members in trolley dilemmas (Cikara et al., 2010, 2017). Participants in these studies were presented with trolley dilemmas and asked to decide on the acceptability of sacrificing members of various groups in order to save the ingroup and vice versa. In addition to identifying the anterior insula as a specific region for group classification (Cikara et al., 2017), the researchers also confirmed a higher perceived acceptability to sacrifice extreme outgroup members (e.g. the homeless) in particular, in order to save ingroup members. This higher rating of acceptability was found to engage brain areas related to complex tradeoffs (medial PFC, left lateral OFC, left dorsolateral PFC).

Summarizing, the research on moral decision making in the context of trolley dilemmas does show effects of social relationships within the dilemma, both in terms of behavioral and neurological responses. However, as stated above, all of these studies were conducted with participants that complete the task in isolated cubicles or fMRI settings. In the next section, we discuss how participants' responses to the trolley dilemma may differ when the experimental settings are more similar to a natural social situation, that is, when another person is present.

Social regulation of moral decision making

Research that provides insight into how the actual physical presence of others influences decision makers' responses is scarce. One of the few studies that asked participants to judge outgroup members in the presence of another person showed that the presence of an ingroup member lead to harsher judgment of outgroup members than when the participants were alone or in the presence of a neutral observer (Wilder & Shapiro, 1991). These findings are explained in accordance with social identity theory (Hogg & Hains, 1996;

Tajfel et al., 1979), which predicts that this harsher judgment is a result of increased salience of the ingroup identity, resulting in increased derogation of the outgroup.

More recent studies that specifically focus on moral dilemmas emphasize people's desire to appear like a moral person to others (Gino, Gu, & Zhong, 2009; Rom & Conway, 2018). This research shows that people engage in meta-cognitions about what an (imagined) other person will think of their response to the moral dilemma, and that these concerns motivate them to strategically present their moral decisions as more in line with what they expect to be desirable. Participants in these studies did not change their private decisions, suggesting that the effect is due to a felt need to change self-presentation (Rom & Conway, 2018). This research shows the role of meta-cognitions when making moral decisions in the presence of others.

A more fundamental insight that helps to understand how making decisions alone differs from making decisions in the presence of others, can be found in social baseline theory (SBT ; Beckes & Coan, 2011). Social baseline theory posits that humans are wired to function in dyads, rather than as individuals. Being alone thus represents a more effortful situation, that may even induce anxiety and existential angst of being separated from society (Rai & Fiske, 2012), especially in threatening situations (Beckes & Coan, 2011), while being together presents a more relaxed state. As a consequence, people in dyads spend less cognitive energetic resources on the same task (Beckes & Coan, 2011). Research on social baseline theory shows that people experience less threat and emotional stress, and feel more able to cope with challenges such as the threat of mild electric shocks, when they are together with another person than when they are alone (Coan, Kasle, Jackson, Schaefer, & Davidson, 2013; Coan, Schaefer, & Davidson, 2006). Extending this to physically challenging situations, Schnall and colleagues (Schnall, Harber, Stefanucci, & Proffitt, 2008) found that participants who were asked to estimate the slant of a hill while wearing a heavy backpack,

estimated the hill to be less steep in the presence of another person, suggesting mental load sharing. Importantly, research on social baseline theory suggests that when making decisions in the presence of another person instead of alone, the difference is not attributable to mere self-presentation concerns, but the cognitive process and state of the decision maker are fundamentally different between the two situations.

Hypothesis development

Building on the research above, we suggest that moral decisions are fundamentally different when they are made alone vs. when they are made in the presence of a close other. These decisions may differ due to the salience of the joint identity (cf. Wilder & Shapiro, 1991) may render moral decisions that are made in the presence of a close other more extreme than when these decisions are made alone. Meta-cognitions about the impression left on the other (Rom & Conway, 2018) could make the decision maker comply to the ingroup norm of derogating the outgroup. In addition, we suggest that solving moral dilemmas may be experienced as demanding and threatening to the self. Indeed, previous research has shown that people prefer to refrain from taking action in moral dilemmas (Cikara & Fiske, 2011; Foot, 1967). Applying social baseline theory to morally demanding situations, the presence of a close other (operationalized through hand-holding as in Coan et al. 2006) could lead to the experience of load sharing and could ease the mental strain of overcoming one's inhibitions when considering the sacrifice of an outgroup member to save an ingroup member. Thus, we hypothesize

Hypothesis 1a: the sacrifice of the outgroup to save the ingroup will be rated as more acceptable when the response is given in the presence of a close other, compared to responding alone.

As the combined theories do not provide an unambiguous prediction, we will also explore a competing hypothesis. Previous studies have shown that engaging in moral decision

making may be demanding in terms of cognitive (Ayal, Gino, Barkan, & Ariely, 2015) and even emotional resources (Zheng et al., 2015). Building on social baseline theory, it can be argued that people experience sharing of the load and experienced risk while resolving the dilemma in the presence of a close other, due to shared emotion regulation (Coan et al., 2006). This would in particular concern a reduction in the processing of threats, such as the salience of the outgroup identity of the to be sacrificed person in the trolley dilemma. Complementing this view, meta-cognitions about moral decision making could lead to a general tendency to refrain from sacrificing anyone in a moral dilemma, reducing the acceptability of sacrificing the outgroup. Based on this reasoning, the competing hypothesis is

Hypothesis 1b: the sacrifice of the outgroup to save the ingroup will be rated as less acceptable when the response is given in the presence of a close other, compared to responding alone.

Not all differences in processing decisions alone vs. in a social context need to be within the decision maker's awareness (Beckes & Coan, 2011), and decision processes may differ even if the behavioral outcome is similar. Therefore, in addition to the hypotheses about behavioral outcomes, we propose that the decision making itself will be different when the response is given in the presence of a close other compared to responding alone. Specifically, we predict differences both in strength of activation and in the brain areas engaged in the decision making process. Firstly, following Cikara and colleagues (2010), we expect that rating the acceptability of the sacrifice of an outgroup target should engage the medial PFC, left lateral OFC, and left dorsolateral PFC. Load sharing by the presence of a close other should potentially alleviate the need to engage in complex deliberations when sanctioning such a utilitarian action of a protagonist, thus leading to less activation of these areas in the presence of a close other. We thus hypothesize

Hypothesis 2a: rating the acceptability of the sacrifice of an outgroup target should lead to an increase of activation in medial PFC, left lateral OFC and left dorsolateral PFC, this increase should be attenuated when responding in the presence of a close other.

Secondly, following research on social baseline theory, the presence of a close other, operationalized by handholding, should result in reduced activation in areas previously associated with threat processing compared to when responding alone (Beckes & Coan, 2011). These areas are the dlPFC, vACC, nucleus accumbens, and hypothalamus. We hypothesize

Hypothesis 2b: rating the acceptability of the trolley dilemma should lead to an increase in are the dlPFC, vACC, nucleus accumbens, and hypothalamus, this increase should be attenuated when responding in the presence of a close other.

In addition to these hypotheses we will explore possible interaction effects of the factors group and handholding on neural activation, as well as effects reflecting meta-cognitions. These effects are however too complex to predict in specific hypotheses.

Method

Participants

In the following sections, we report all measures, manipulations, and exclusions. Participants were 21 healthy female volunteers recruited through flyers on the campus. We invited participants to come to the lab accompanied by a close female friend that they would invite to participate with them. As both the participant and the friend would be present in the scanner-area, we pre-screened[1] participants and their friends for contra-indications for MRI

[1] As part of the recruitment, participants and their friends filled out a short survey containing self-report measures. Measures included in this survey were moral identity, relationship quality, and attachment style (when not skipped by the participant). We only used these measures to screen for unusual scores, and as potential control variables for the behavioral data.

research and only invited pairs that met the safety criteria. All participants were right-handed native Dutch speakers, with no history of neurological problems, and had normal or corrected vision. Four participants were eliminated from the final sample, because of movement artifacts (1) or misunderstanding of the instructions (3), leaving a remaining sample of 17 participants. This sample size was determined before any data analysis. Although this is a relatively small sample size, sensitivity analysis based on an α of .05, and a power level of .80 (Faul, Erdfelder, Buchner, & Lang, 2009; Giner-Sorolla, 2018), indicates our sample to be sufficient to detect medium effects, *Cohens f* = .25. A detailed description of the materials and procedure can be found below; all materials were presented in Dutch. Participants took part in the study at least one week after completion of these surveys to avoid priming with morality in the experiment.

Stimuli

The moral dilemma situations were adapted from the footbridge scenario used by Cikara and colleagues (2010). In contrast to Cikara et al.'s paradigm that concerned four different types of groups, we were mainly interested in participants' different reactions to ingroup versus outgroup others – in our case competent but less warm (cf. Cikara et al., 2010) peers. For this reason, participants in the current study were presented with dilemmas that ask to either sacrifice a) an outgroup member, to save a group of ingroup members, or b) an ingroup member, to save a group of outgroup members. Sacrificed participants were represented by a single picture; saved participants were represented five facial portraits. Facial stimuli were taken from four databases of standardized portraits: Karolinska Directed Emotional Faces (KDEF**;** (Lundqvist, Flykt, & Öhman, 1998); the Utrecht ECVP database (Hancock, 2008); the Radboud Faces Database (RaFD; (Langner et al., 2010) and the FACE database (Minear & Park, 2004). A list of the included picture numbers can be found in

Appendix A. The facial stimuli featured both male and female faces from different ethnicities, representative of a contemporary student body at an international university. All faces showed a neutral facial expression. Facial stimuli for slides with multiple faces were designed to have similar background, luminosity and face size. Group membership was indicated by a banner that either represents the ingroup (own university, or the outgroup (a German university, pretested to be perceived as most different in the list of Dutch and German universities).

Procedure

On the day of the MRI study, participants arrived with their friend to the lab. Both signed a consent form and were introduced to the task. Next, participants participated in a 2 (alone vs. handholding) by 2 (ingroup vs outgroup sacrifice) within participant experiment. Throughout the experimental procedure participants were alone half of the time, and held the hand of their friend during the other half of the time, in randomized order. To minimize the disturbance of the process, handholding took place in two subsequent blocks; during the first two, the middle two, or the final two blocks. As the duration of each block was approximately 5 minutes (300s), participants held the hand of their friend for approximately 10 minutes in total. A graphic representation of the study setup can be found in Figure 1.

Insert Figure 1 about here

None of the participants expressed discomfort with the procedure, or stopped the handholding during the experiment. The left hand was used for the hand-holding, while the right(dominant) hand was used for responding to the experimental responses. In order to alleviate differences in sensorimotor processing, participants held the pneumatic alarm ball of the MRI in the left hand during alone trials. Importantly, participants were unable to

communicate with their friends during the experiment – their ears were blocked due to the noise protection canceling out the MRI noise, the participant and the partner could not easily make eye contact due to the placement of the participant in the MRI scanner, and the partner was unable to see the participants' decisions.

In each trial, participants were presented with the moral trolley dilemma scenario while brain activation was measured. As described above, the core of the dilemma related to the acceptability of sacrificing one person to save five others. The participants received an extensive explanation and schematic representation of this dilemma before starting the experiment. Participants were told to imagine that the decision to sacrifice the single person has already been made, and that they should rate to what extent the action was morally acceptable on a 4-point scale ranging from 1= not at all, to 4=very on a button response box (See Cikara et al., (2010) for a similar procedure). The stimuli consisted of pictures depicting a single vs. five persons to represent the sacrificed and saved targets respectively. These targets will either be from the same university (ingroup) or from a different university (outgroup) to the participant. As in earlier research, the period of interest was a window of 4 seconds centered on the moment of rating (Cikara et al., 2010). After the scan, participants and friends were debriefed and compensated.

MRI session & acquisition parameters

(f)MRI data were acquired with a Siemens 3.0T Magnetom Prisma fit scanner (Siemens Medical, Erlangen, Germany) using an 64-channel headcoil. Anatomical imaging was carried out with a standard Alzheimer Disease Neuroimaging Initiative (ADNI) complaint T1 weighted sequence, voxel size $1mm^3$, FOV = 256*256mm, 192 slices, TR = 2250ms., TE = 2ms. For the functional data, whole brain Echo-Planar Imaging (EPI) of 180 volumes per run was performed (38 slices without gap, voxel size $3mm^3$, FOV = 216*216mm TR = 2000ms, TE = 28ms). Slice orientation was tilted in order to minimize susceptibility

artefacts in the orbitofrontal regions (Deichmann, Gottfried, Hutton, & Turner, 2003). As described above, the participants completed 4 runs in about 25 minutes. During the handholding runs, the friend of the participant held their left hand while standing next to the MRI scanner. Hand-holding runs were balanced across participants to occur either in the beginning, the middle or the end of the 4-run series.

Analysis

fMRI data

Data was preprocessed and analyzed using Brain Voyager 21.4 (Brain Innovation BV, the Netherlands) Anatomical images were corrected for inhomogeneity, peeled, and transformed to stereotactic coordinate space (Talairach & Tournoux, 1988). Functional images were slice-scan-time corrected, motion-corrected and high-pass filtered (2 cycles per run) and smoothed with a Gaussian Kernel of 8mm. A boxcar predictor modelling the rating process was defined to last from the moment the participant was presented with the five pictures presenting the "saved" group to the button press in the acceptability rating. The predictor was then convolved with a canonical hemodynamic response function.

A full-brain random-effects (RFX) GLM was calculated. Separate contrast maps were calculated for the main effects of Group (sacrificing ingroup vs. sacrificing outgroup) and handholding (together vs. alone). Betas values of the RFX GLM were furthermore used as input for a two-factors repeated measures ANOVA to compute a map of interaction between the group and handholding factors. Contrast maps were adjusted to reflect a confidence level of alpha $< .05$ using a cluster threshold estimation approach (Forman et al., 1995). Minimal cluster sizes were 7 continuous functional voxels for the group contrast map, 9 for the handholding contrast map and 9 for the interaction map, using a primary threshold of $p < .001$.

Results

Behavioral data

Acceptability ratings and reaction times were averaged per condition (ingroup vs. outgroup) and per block (alone vs. together). As the majority of the participants responded too late to the first trial of the first block, these trials were excluded for all participants. Results from a 2 (ingroup vs. outgroup) x 2 (alone vs. together) repeated measures ANOVA with acceptability as the dependent variable showed only that participants judged outgroup sacrifice to be more acceptable (*M*=3.09, *SE*=.08) than ingroup sacrifice (*M*=2.98, *SE*=.09), $F(1, 66)= 11.81$, $p= .001$, $\eta_p=.15$, *Cohens f* = $\sqrt{.15/(1-.15)} = .42$. There were no differences between the handholding (*M*=3.03, *SE*=.12) and alone (*M*=3.04, *SE*=.12) conditions, $F(1, 66)= 0.01$, $p= .919$, nor were there any interactions $F(1, 66)= 0.40$, $p= .527$. These effects can also be observed in Figure 1.

Additional analyses show that results for reaction time were in the same direction but not significant; a 2 (ingroup vs. outgroup) x 2 (alone vs. together) repeated measures ANOVA showed no difference between dilemmas judging outgroup sacrifice (*M*=1080.95, *SE*=52.84) or dilemmas judging ingroup sacrifice (*M*=1138.84, *SE*=58.21), $F(1, 66)= 2.785$, $p= 0.100$, $\eta_p=.04$. There were no differences between the handholding (*M*=1083.87, *SE*=74.70) and alone (*M*=1135.94, *SE*=74.70) conditions, $F(1, 66)= 0.24$, $p= .624$, neither was the interaction significant, $F(1, 66)= 0.60$, $p= .440$[2].

We must thus conclude that handholding does not increase the acceptability sacrificing an outgroup member (rejecting H1a). We furthermore conclude that the presence of a close other does not decrease the acceptability of sacrificing an outgroup member, either (rejecting H1b).

Imaging results

[2] There was no effect for controlling for friendship closeness or attachment. Participants did rate the paradigm as more acceptable on average when they started alone (*M*=3.39, *SE*=.15), compared to when they were alone in the middle two blocks, (*M*=2.92, *SE*=.12), or in the last block (*M*=2.84, *SE*=.15), $F(1, 66)= 4.35$, $p= .017$, $\eta_p=.12$. Controlling for the order did however not change the results.

We first tested for the main effect of the factor group to assess whether this factor evoked a different neural response. This analysis showed effects similar to other studies focusing on ingroup favoritism (Volz, Kessler, & von Cramon, 2009). Specifically, when rating the acceptability of sacrificing an outgroup member in order to save ingroup members, a cluster in the middle Occipital gyrus (OcG) showed an increase in activation (see also Figure 3 and Table 1.

Insert Figure 3 and Table 1 about here

Main effects contrast analysis for the handholding factor showed that when contrasting brain activation of acceptability rating trials while holding the hand of a friend vs. alone, irrespective of identity of the sacrificed person, none of the regions that we hypothesized to be influenced by handholding in H2a (medial PFC, left lateral OFC, left dorsolateral PFC) or H2b (dlPFC, vACC, nucleus accumbens, and hypothalamus) showed any difference in activation. Nonetheless, the analysis did show increased activation when holding hands in one cluster (see Figure 3 and Table 1). This cluster extended in the right posterior cingulate cortex (PCC) to the precuneus. The precuneus has been shown to be commonly recruited over a range of moral decision making tasks, in particular when inferring the mental states of others' moral actions (Eres, Louis, & Molenberghs, 2017), as it was the case as well in our paradigm. In particular, also the PCC has been previously reported by Cikara et al. (2010) during acceptability rating of sacrificing a low-warmth, low- competence target in order to save a high-warmth, high-competence target, as compared to all other sacrificing/saving combinations, albeit in the contralateral (left) hemisphere.

Factorial analysis with both the group identity and hand holding factors did not show an interaction effect in any of the other regions that were expected to differ as a function of

either the group membership of the sacrificed person (e.g., PFC, OFC (H2a)) or as a function of handholding (vACC, nucleus accumbens, hypothalamus (H2b)). However, exploratory analyses revealed an interaction of the two factors in the superior frontal gyrus lateralis (SFGL)/precentral gyrus (PrG) (see Figure 3 and Table 1). This area showed an increased activation when participants were holding the hand of a friend during acceptability rating of sacrificing an outgroup member in order to save ingroup members, as compared to while being alone when performing the same rating. This effect was reversed during the sacrificing of an ingroup member in order to save outgroup members. SFG activation has been associated with increased probability that a sacrifice is necessary in a moral dilemma (Shenhav & Greene, 2010) and in paradigms evoking guilt or shame (Michl et al., 2014).

Discussion

We investigated the behavioral and neurological effects of answering a moral trolley dilemma alone, or while holding the hand of a friend. Results of our behavioral data were in line with findings from the classical social identity theory about harsher treatment of the outgroup than the ingroup (e.g., Hogg & Turner, 1987), but did not show main effects of, or interactions with, the presence of a close other (H1a, H1b). Specifically, participants considered outgroup sacrifices as more acceptable than ingroup sacrifices overall. Thus, in terms of behavior, there was no difference when participants were responding to the moral dilemma alone versus in the presence of a close other.

Imaging results did however show differences in the patterns of brain activation in the different conditions. Our hypotheses concerning the effects of handholding on either regions related to complex tradeoffs (H2a) or threat processing (H2b) were not confirmed, and we did not replicate earlier findings on moral dilemma processing in frontal cortex areas (Cikara et al., 2010; Shenhav & Greene, 2010). The latter finding might be explained by the less extreme nature of the employed ingroup/outgroup contrast in our study. Nonetheless, our

results showed that our employed ingroup/outgroup contrast evoked modulations in brain activation in regions associated previously with ingroup favoritism (Volz et al., 2009).

Our exploratory interaction analyses showed that areas previously related to feeling guilt and shame (Boccia et al., 2016; Michl et al., 2014) are engaged more during rating the acceptability of outgroup sacrifice in the handholding condition. It bears emphasizing at this point that our participants showed a clear difference in acceptability of ingroup and outgroup sacrificing. A sanctioning of the choice to sacrifice a person thus mostly happened only in the "sacrifice outgroup" condition. We interpret the increase in activation as a sign that our participants did still make the choice, yet the consequences of their sanctioning of a sacrifice might have become more salient due to the presence of a close other for whom the participants wanted to keep a positive moral image (Gino, Gu, & Zhong, 2009; Rom & Conway, 2018).

Limitations and suggestions for future research

A few limitations should be mentioned: Most research on social baseline theory includes direct physical threats (Coan et al., 2013, 2006). Although the effects of social connectedness have also been found to affect decision making when faced with physical challenges (Schnall et al., 2008) it is not clear whether our focus on moral dilemmas did induce a similar sort of threat to the participants. Between-participant experimental research that assessed the effect of social connectedness on a (one-time) trolley dilemma decision does suggest that the presence (or reminder) of a close other does affect moral decision making (Lucas & Livingston, 2014). We did not find significant differences between the handholding and alone conditions in the neural data with regard to areas associated with decision processing (medial PFC, left lateral OFC and left dorsolateral PFC). This might be a first indirect indication that participants at least did not engage in excessive rumination with regard to the dilemma because of holding the hand of the friend. Future research should

however disentangle the potential motivations for decision making in the presence of a close friend, and more strongly establish whether moral dilemmas can be experienced as threatening.

A second concern with regard to the strength of the manipulation may relate to the choice of group membership in the design. Due to the complexity of the study, we set up our design to resemble the study described by Cikara and colleagues (2010) as closely as possible. Although we pre-tested our chosen outgroup to be similar to the ingroup in terms of competence but somewhat lower in warmth and higher in distance from our participants, it might be that the in- and outgroups in this study were too similar to each other to warrant behavioral effects beyond the main effect of group. Future research could focus on a more extreme outgroup (See e.g. Cikara et al., 2010).

In addition, as the third person perspective on the trolley dilemma that we used in the current study is rather specific, a larger variety of moral dilemmas and variations on the trolley dilemma could be investigated in future studies (Boccia et al., 2016). In this light it is important to recognize recent critiques on the external validity of the trolley dilemma (Bostyn, Sevenhant, & Roets, 2018; Kahane, 2015). We chose to use the trolley paradigm in the current study in order to limit the number of deviations from the paradigm by Cikara and colleagues (Cikara et al., 2010) so that we could identify the effect of the presence of a close other most clearly. Although the trolley paradigm may not fully correspond to daily moral dilemmas, we argue that other types of dilemmas may show similar differences in processing when the decision maker is alone versus in the presence of a close other. It is however up to future research to find support for this suggestion.

In the design of our study, we have taken care to eliminate possible alternative explanations for our findings. Specifically, we invited heterosexual female friends in order to reduce awkwardness, or any romantic interferences. Our participants did not seem to be

uncomfortable with the procedure, or unwilling to do it. None of the participants let go of their friends' hand during the experiment. Despite these precautions, it is however possible that our manipulation of social closeness has induced other thoughts and cognitions. For example, our study does not take into account to what extent participants were aware of their friend's moral values and to what extent the friend's moral values were similar to their own. Importantly, while the presence of the friend may have increased the salience of ingroup membership and the group norms, the decision driven by this salience might have been both more or less moral depending on the content of the norm (cf. van Gils, Hogg, Van Quaquebeke, & van Knippenberg, 2015). This may have produced some noise in the responses in the handholding condition.

Finally, it is unclear to what extent participants engaged in meta-cognitions about expectations of the close other in any of the conditions (Rom & Conway, 2018). The friend was present in the adjacent room during the blocks that the participant completed alone, it might be possible that the participants did not feel completely alone in this condition. In addition, although the friend was not observing the participants' answers in the handholding condition, participants still may have felt the need to improve their moral self-presentation. The design employed in our fMRI study is however not suitable to answer these questions. Future research should test this question with experimental designs that avoid repeated questions about affective processes or meta-cognitions.

Conclusion

Our results extend findings of earlier studies on moral dilemmas, and specifically the trolley dilemma (Cikara et al., 2010; Conway & Gawronski, 2013), by showing that rating the acceptability of the sacrifice of the outgroup evoke more empathic concern when being together with a close other. Given the neurological effect of the presence of others on

responses in moral dilemmas, we suggest that future research should take into account that the social context in morality research may influence the way dilemmas are processed.

References

Abrams, D., & Hogg, M. A. (2010). Social identity and self-categorization. In J. F. Dovidio, M. Hewstone, P. Glick, & V. M. Esses (Eds.), *The SAGE handbook of prejudice, stereotyping and discrimination* (pp. 179–193). London, UK: SAGE.

Ayal, S., Gino, F., Barkan, R., & Ariely, D. (2015). Three Principles to REVISE People's Unethical Behavior. *Perspectives on Psychological Science*, *10*(6), 738–741. https://doi.org/10.1177/1745691615598512

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529.

Beckes, L., & Coan, J. A. (2011). Social Baseline Theory: The Role of Social Proximity in Emotion and Economy of Action. *Social and Personality Psychology Compass*, *5*(12), 976–988. https://doi.org/10.1111/j.1751-9004.2011.00400.x

Bernhardt, B. C., & Singer, T. (2012). The Neural Basis of Empathy. *Annual Review of Neuroscience*, *35*(1), 1–23. https://doi.org/10.1146/annurev-neuro-062111-150536

Boccia, M., Dacquino, C., Piccardi, L., Cordellieri, P., Guariglia, C., Ferlazzo, F., … Giannini, A. M. (2016, February 25). Neural foundation of human moral reasoning: an ALE meta-analysis about the role of personal perspective. *Brain Imaging and Behavior*, pp. 1–15. https://doi.org/10.1007/s11682-016-9505-x

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of Mice, Men, and Trolleys: Hypothetical Judgment Versus Real-Life Behavior in Trolley-Style Moral Dilemmas. *Psychological Science*, *29*(7), 1084–1093. https://doi.org/10.1177/0956797617752640

Cikara, M., Farnsworth, R. A., Harris, L. T., & Fiske, S. T. (2010). On the wrong side of the

trolley track: neural correlates of relative social valuation. *Social Cognitive and Affective Neuroscience, 5*(4), 404–413. https://doi.org/10.1093/scan/nsq011

Cikara, M., & Fiske, S. T. (2011). Bounded empathy: neural responses to outgroup targets' (mis)fortunes. *Journal of Cognitive Neuroscience, 23*(12), 3791–3803. https://doi.org/10.1162/jocn_a_00069

Cikara, M., Van Bavel, J. J., Ingbretsen, Z. A., & Lau, T. (2017). Decoding "Us" and "Them": Neural representations of generalized group concepts. *Journal of Experimental Psychology: General*, *146*(5), 621–631. https://doi.org/10.1037/xge0000287

Coan, J. A., Kasle, S., Jackson, A., Schaefer, H. S., & Davidson, R. J. (2013). Mutuality and the social regulation of neural threat responding. *Attachment & Human Development*, *15*(3), 303–315. https://doi.org/10.1080/14616734.2013.782656

Coan, J. A., Schaefer, H. S., & Davidson, R. J. (2006). Lending a hand: social regulation of the neural response to threat. *Psychological Science*, *17*(12), 1032–1039. https://doi.org/10.1111/j.1467-9280.2006.01832.x

Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: a process dissociation approach. *Journal of Personality and Social Psychology*, *104*(2), 216–235. https://doi.org/10.1037/a0031021

Cox, R. W., Chen, G., Glen, D. R., Reynolds, R. C., & Taylor, P. A. (2017). FMRI Clustering in AFNI: False-Positive Rates Redux. *Brain Connectivity*, *7*(3), 152–171. https://doi.org/10.1089/brain.2016.0475

Deichmann, R., Gottfried, J. A., Hutton, C., & Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *NeuroImage, 19*(2), 430–441.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences, 113*(28), 7900–7905. https://doi.org/10.1073/pnas.1602413113

Eres, R., Louis, W. R., & Molenberghs, P. (2017, July). Common and distinct neural networks involved in fMRI studies investigating morality: an ALE meta-analysis. *Social Neuroscience*, pp. 1–15. https://doi.org/10.1080/17470919.2017.1357657

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York, NY: McGraw-Hill.

Foot, P. (1967). Theories of ethics. Retrieved from https://philpapers.org/rec/FOOTOE

Friesdorf, R., Conway, P., & Gawronski, B. (2015). Gender Differences in Responses to Moral Dilemmas. *Personality and Social Psychology Bulletin*, *41*(5), 696–713. https://doi.org/10.1177/0146167215575731

Giessner, S. R., & Van Quaquebeke, N. (2010). Using a relational models perspective to understand normatively appropriate conduct in ethical leadership. *Journal of Business Ethics*, *95*, 43–55. https://doi.org/10.1007/s10551-011-0790-4

Giner-Sorolla, R. (2018). New Policies at JESP for 2018: The Why and How. Retrieved from https://approachingblog.wordpress.com/2017/12/20/new-policies-at-jesp-for-2018-the-why-and-how/

Gino, F., Gu, J., & Zhong, C.-B. (2009). Contagion or restitution? When bad apples can motivate ethical behavior. *Journal of Experimental Social Psychology*, *45*(6), 1299–1302. https://doi.org/10.1016/j.jesp.2009.07.014

Greene, J. D. (2001). An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science*, *293*(5537), 2105–2108. https://doi.org/10.1126/science.1062872

Haidt, J. (2008). Morality. *Perspectives on Psychological Science*, *3*(1), 65–72. https://doi.org/10.1111/j.1745-6916.2008.00063.x

Hancock, P. (2008). Utrecht ECVP.

Hogg, M. A., & Hains, S. C. (1996). Intergroup Relations and Group Solidarity: Effects of Group Identification and Social Beliefs on Depersonalized Attraction. *Journal of Personality and Social Psychology*, *70*(2), 295–309. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=ovftb&AN=00005205-199602000-00008

Hogg, M. A., & Turner, J. C. (1987). Social identity and conformity: A theory of referent informational influence. In W. Doise & S. Moscovici (Eds.), *Current issues in European social psychology* (Vol. 2, pp. 139–182). Cambridge, UK: Cambridge University Press.

Kahane, G. (2015). Sidetracked by trolleys: Why sacrificial moral dilemmas tell us little (or nothing) about utilitarian judgment. *Social Neuroscience*, *10*(5), 551–560. https://doi.org/10.1080/17470919.2015.1023400

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*(8), 1377–1388.

Lucas, B. J., & Livingston, R. W. (2014). Feeling socially connected increases utilitarian choices in moral dilemmas. *Journal of Experimental Social Psychology*, *53*, 1–4. https://doi.org/10.1016/J.JESP.2014.01.011

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet*, (1998).

Michl, P., Meindl, T., Meister, F., Born, C., Engel, R. R., Reiser, M., & Hennig-Fast, K. (2014). Neurobiological underpinnings of shame and guilt: a pilot fMRI study. *Social Cognitive and Affective Neuroscience*, *9*(2), 150–157. https://doi.org/10.1093/scan/nss114

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior*

Research Methods, Instruments, & Computers, *36*(4), 630–633.

Rai, T. S., & Fiske, A. P. (2012). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 23*, 189–193. https://doi.org/10.1037/a0021867

Rom, S. C., & Conway, P. (2018). The strategic moral self: Self-presentation shapes moral dilemma judgments. *Journal of Experimental Social Psychology, 74*, 24–37. https://doi.org/10.1016/j.jesp.2017.08.003

Rom, S. C., Weiss, A., & Conway, P. (2017). Judging those who judge: Perceivers infer the roles of affect and cognition underpinning others' moral dilemma responses. *Journal of Experimental Social Psychology, 69*, 44–58. https://doi.org/10.1016/j.jesp.2016.09.007

Schnall, S., Harber, K. D., Stefanucci, J. K., & Proffitt, D. R. (2008). Social support and the perception of geographical slant. *Journal of Experimental Social Psychology, 44*(5), 1246–1255. Retrieved from http://www.sciencedirect.com/science/article/pii/S002210310800070X

Shenhav, A., & Greene, J. D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron, 67*(4), 667–677. https://doi.org/10.1016/j.neuron.2010.07.020

Tajfel, H., Turner, N., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worcel (Eds.), *The Social Psychology of Intergroup Relations* (pp. 33–47). Monterey, CA: Brooks/Cole. https://doi.org/10.1016/S0065-2601(05)37005-5

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain.* New York: Thieme.

Thomson, J. (1986). *Rights, restitution, and risk: essays, in moral theory*. Harvard University Press. Retrieved from http://books.google.nl/books?hl=en&lr=&id=sLh4oBgJEtEC&oi=fnd&pg=PR7&dq=tho

mson+1986+moral&ots=7HNWT2wnwJ&sig=UqV4cMtpQe3cFS7oY5v9vXW6CGY

van Gils, S., Hogg, M. A., Van Quaquebeke, N., & van Knippenberg, D. (2015). When Organizational Identification Elicits Moral Decision-Making: A Matter of the Right Climate. *Journal of Business Ethics*. https://doi.org/10.1007/s10551-015-2784-0

Volz, K. G., Kessler, T., & von Cramon, D. Y. (2009). In-group as part of the self: In-group favoritism is mediated by medial prefrontal cortex activation. *Social Neuroscience, 4*(3), 244–260. https://doi.org/10.1080/17470910802553565

Wilder, D. A., & Shapiro, P. (1991). Facilitation of outgroup stereotypes by enhanced ingroup identity. *Journal of Experimental Social Psychology*, *27*(5), 431–452. https://doi.org/10.1016/0022-1031(91)90002-N

Zheng, D., Witt, L. A., Waite, E., David, E. M., van Driel, M., McDonald, D. P., … Crepeau, L. J. (2015). Effects of ethical leadership on emotional exhaustion in high moral intensity situations. *The Leadership Quarterly*, *26*(5), 732–748. https://doi.org/10.1016/j.leaqua.2015.01.006

Figure 1. Graphical representation of the study setup

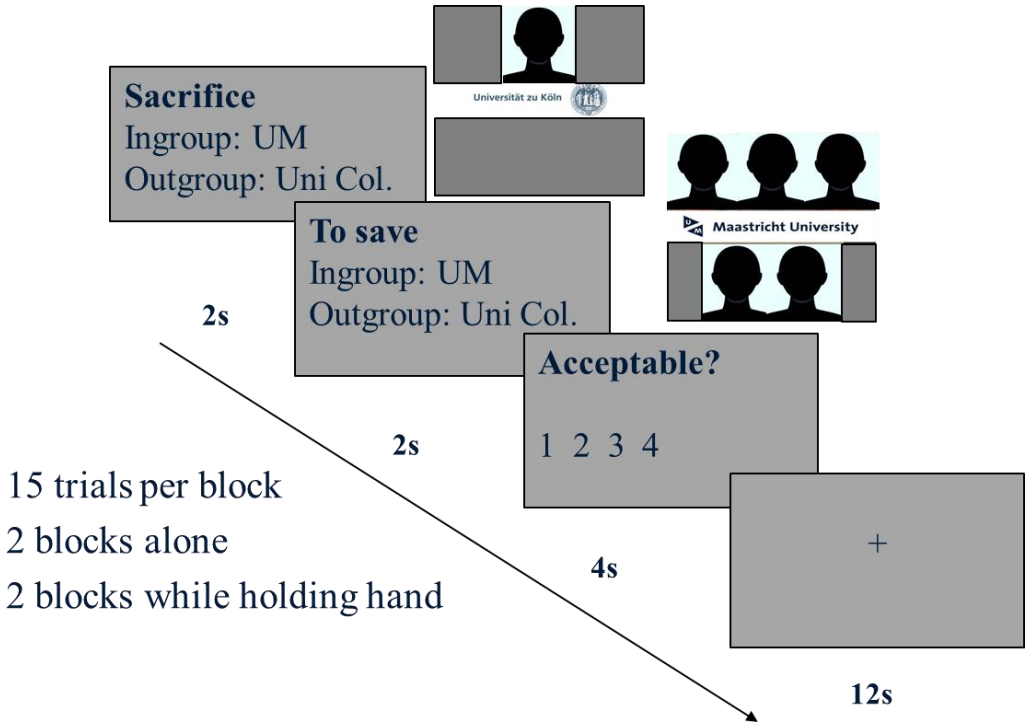


Figure 2. Behavioral results for the acceptability ratings of sacrifice per condition. Error bars represent standard errors.

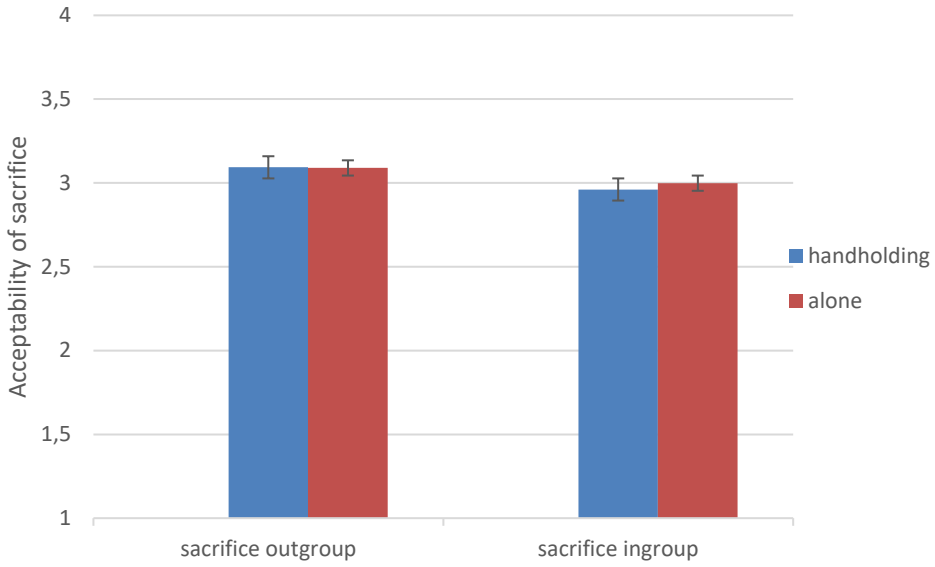
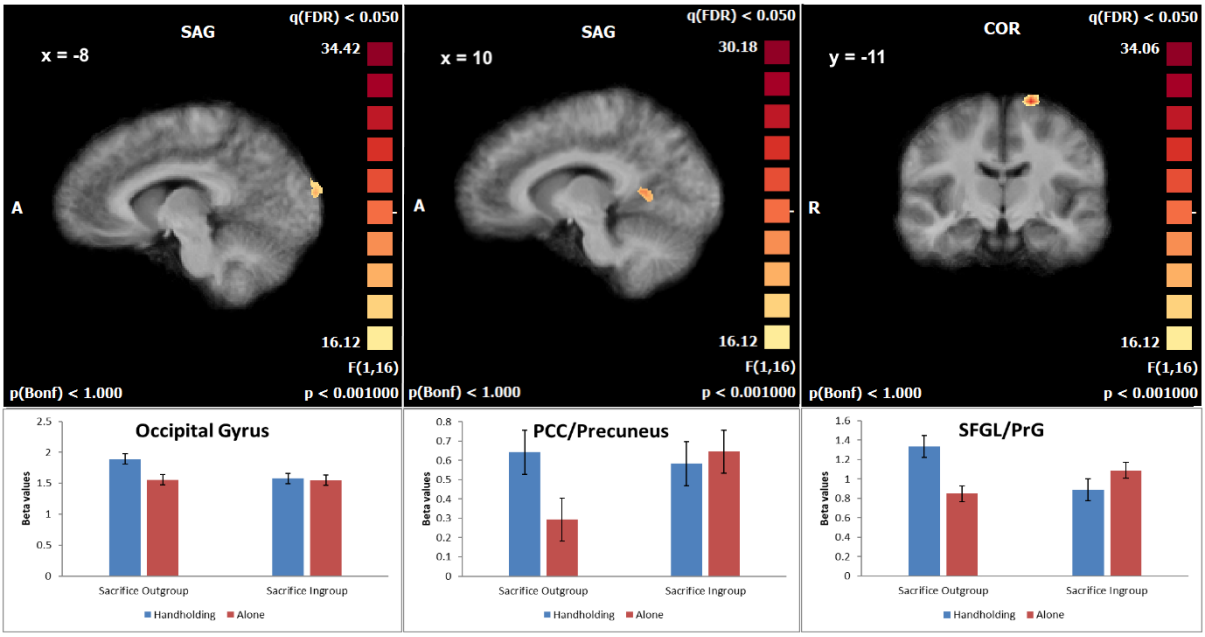


Figure 3. Graphical overview of the main effects for group sacrifice (left), handholding (middle), and the interaction effect (right). Top: Locations of the OcG cluster (left), the PCC/Precuneus cluster (middle) and the SFGL/Prg cluster (right). Bottom: Averaged beta values per cluster.



Note: Standard error bars are for illustrational purpose only.

Table 1.

Brain areas showing either main or interaction effects

Location	**BA**	**Peak x**	**Peak y**	**Peak z**	***F***	***p***	**Voxel**
Sacrifice Outgroup > Ingroup							
Occipital Gyrus	18	-9	-97	16	29,04	0,00	241
Handholding > Alone							
PCC/Precuneus	18 / 31	9	-49	10	29,18	0	266
Interaction							
SFGL/PrG	4	-18	-13	67	33,06	0,00	343

APPENDIX 1 – Pictures used per database

Utrecht ECVP	KDEF	FACE	RaFD*
f4004	AF01NESJ	EMBfemale20-2neutral	Rafd090_01_Caucasian_female_
f4006	AF05NESJ	EMBfemale20neutral	Rafd090_02_Caucasian_female_
f4009	AF07NESJ	EMBfemale21-4neutral	Rafd090_03_Caucasian_male_
f4017	AF19NESJ	EMWfemale18-2neutral	Rafd090_04_Caucasian_female_
f4018s	AF23NESJ	EMWfemale18neutral	Rafd090_05_Caucasian_male_
f4021	AM10NESJ	EMWfemale19neutral	Rafd090_07_Caucasian_male_
f4026	AM18NESJ	EMWfemale20neutral	Rafd090_08_Caucasian_female_
f4027	AM23NESJ	EMWfemale21-2neutral	Rafd090_09_Caucasian_male_
f4029	AM25NESJ	EMWfemale21-4neutral	Rafd090_10_Caucasian_male_
f4030		EMWfemale21-3neutral	Rafd090_12_Caucasian_female_
m4001		EMWfemale22neutral	Rafd090_14_Caucasian_female_
m4011		EMWfemale23neutral	Rafd090_15_Caucasian_male_
m4062s		EMWfemale27-3neutral	Rafd090_16_Caucasian_female_
m4076		EMWfemale27-4neutral	Rafd090_18_Caucasian_female_
		TMAfemale19neutral	Rafd090_19_Caucasian_female_
		TMBfemale18neutral	Rafd090_20_Caucasian_male_
		TMBfemale19-2neutral	Rafd090_22_Caucasian_male_
		TMBfemale 19-3neutral	Rafd090_23_Caucasian_male_
		TMBfemale21neutral	Rafd090_25_Caucasian_male_
		TMWfemale19-3neutral	Rafd090_26_Caucasian_female_
		TMWfemale18neutral	Rafd090_27_Caucasian_female_
		TMWfemale20neutral	Rafd090_31_Caucasian_female_
		TMWfemale23-2neutral	Rafd090_32_Caucasian_female_
		TMWfemale22neutral	Rafd090_37_Caucasian_female_
		TMWfemale25neutral	Rafd090_56_Caucasian_female_
		WIfemale22neutral	Rafd090_57_Caucasian_female_
		WWfemale20-4neutral	Rafd090_58_Caucasian_female_
		WWfemale20neutral	Rafd090_61_Caucasian_female_
		WWfemale21neutral	
		WWfemale22-4neutral	
		WWfemale22-5neutral	
		TMWfemale19neutral	
		EMWmale19-2neutral	
		EMWmale22-3neutral	
		TMWmale20-4neutral	
		TSFWmale22neutral	
		TSFWmale23neutral	
		WImale23neutral	
		WImale24-2neutral	
		WWmale20-2neutral	
		WWmale22-4neutral	
		WImale26neutral	

*note: all picture names for the RAFD database ended on _neutral_frontal, file names were shortened to maintain readability of the table.