



Universiteit
Leiden
The Netherlands

Neural correlates of the motivation to be moral

Nunspeet, F. van

Citation

Nunspeet, F. van. (2014, May 27). *Neural correlates of the motivation to be moral*. Kurt Lewin Institute Dissertation Series. Ridderprint B.V., Ridderkerk. Retrieved from <https://hdl.handle.net/1887/25829>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/25829>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/25829> holds various files of this Leiden University dissertation.

Author: Nunspeet, Félice van

Title: Neural correlates of the motivation to be moral

Issue Date: 2014-05-27

Part III

**The need for confirmation of
one's own morality**

Chapter 6

Affective and attentional responses to positive and negative feedback about one's own moral behavior

Collaborators on the research described in this chapter are:

Naomi Ellemers, Eveline Crone, Belle Derks, and David Amodio. Their contribution can be specified as follows: Design of the studies: FvN, NE, EC, BD, DA. Performing the experiments: FvN. Analysis of skin conductance data: FvN. Analysis of fMRI data: FvN, EC. Writing of the paper: FvN, NE, EC, BD. *Manuscript in preparation.*

A general principle in psychology is that bad is stronger than good (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Although applicable to many types of judgments and situations, this is also an essential mechanism in judging someone else's moral integrity. This was established by Skowronski and Carlston (1987), who examined positive and negative extremity biases for morality and competence judgments during impression formation. Their findings revealed that negative rather than positive behaviors are perceived as more diagnostic for someone's 'true character' when these refer to the moral domain. In contrast, however, positive rather than negative behaviors are perceived as more diagnostic for someone's personality when these behaviors relate to their competence. In other words, we assume that everyone can act in a moral way, for instance when criminals pretend to be upright citizens - so this is non-diagnostic. However, only *immoral* people should do immoral things. Conversely, we tend to think that everyone can do something *incompetent* once in a while - even a professor can be confused or forgetful - but only competent people should be able to behave competently.

This negative extremity bias concerning morality (e.g., Lupfer, Weeks, & Dupuis, 2000) and the differential diagnosticity of moral and competent behaviors (e.g., Martijn, Spears, Van der Pligt, & Jakobs, 1992) have been observed in empirical research. However, prior studies have focused on impression *formation* about others - examining this from a perceiver's perspective. Thus far, it has remained unclear whether a similar asymmetry in the value attached to moral vs. competent behaviors is also evident in impression *management* - in the concerns people have about the image of the self in the eyes of others (from an actor's perspective). To the extent that positive and negative extremity biases for morality and competence are also associated with impression management about the self, people should be strongly preoccupied with avoiding to display any behavior that might indicate their immorality, and focus on providing confirmation of their competence. Because it is not always possible to act in line with one's ambitions and ideals, people are likely to be confronted from time to time with others who provide negative evaluations of their moral or competent behavior. We argue that the asymmetrical implications of person information concerning morality vs. competence should therefore be visible in the affective states people experience.

That is, they should suffer increased negative states when being confronted with negative information concerning their own morality (as compared to their competence). Conversely, they should experience increased positive states when they receive information about their own competence (as compared to their morality). Previous research concerning people's self-perceptions and impression management has revealed evidence offering partial support for this reasoning, as it has established that people tend to attach greater importance to moral information about the individual or group self than to competence information. That is, overall people indicate they perceive moral traits as more important characteristics of their personal and social identity than traits referring to their competence (and sociability; Leach, Ellemers, & Barreto, 2007). They indicate being motivated to display behavior that is seen as moral as a way to secure inclusion in a group (Ellemers, Pagliaro, Barreto & Leach, 2008) and to earn respect from fellow ingroup members (Pagliaro, Ellemers, & Barreto, 2011). Moreover, this motivation to display moral behavior is also evident at a less explicit level as people tend to inhibit their social bias against Muslims (i.e., display a moral task performance) when the test used to assess this was said to be indicative of their morality instead of their competence (Van Nunspeet, Ellemers, Derks, & Nieuwenhuis, 2014).

Prior research thus underlines the importance of morality over competence in impression management about the self. This is the case when people have to explicitly state their preference or when they are assigned to a task condition that emphasizes either moral or competence implications of task performance. As yet, it still needs to be examined whether the greater value attached to moral information about the self relates to the desire to avoid appearing immoral, or stems from the ambition to demonstrate one's ability to behave morally. The aim of the present research was to directly compare the impact of these different types of information related to the self, as a way to establish whether people differentially welcome information that might confirm their morality or competence in a positive way, or are disturbed by negative information depending on whether it threatens to reveal their lack of morality or competence.

One way of examining the impact of different types of information supposedly relevant to the self, is to ask participants to report how they feel after

receiving this information. Such a method relies on the introspective capabilities of participants and may be affected by people's explicit preferences for a particular type of information over the other, as well as their willingness to reveal these to the experimenter. Thus, such self-report measures do not necessarily provide a reliable picture of their internal states. Psychophysiological measures seem to offer a solution for these difficulties associated with self-report measures. For example, electrodermal activity, often measured as skin conductance, is an automatic response from the sympathetic nervous system caused by arousing stimuli (for an overview see Dawson, Schell, & Filion, 2000). Indices of skin conductance can thus not easily be adapted by the participant for self-presentational reasons, and can be measured online (i.e., to monitor changed states *while* participants receive relevant information, instead of relying on retrospective reports). Combining self-reports with skin conductance data can thus elucidate how people respond to information about their own behavior and compare this to what they report when thinking back about the information.

In addition, previous neuroscientific research has been able to disentangle different cognitive processes associated with processing self-relevant information (most often using functional magnetic resonance imaging; fMRI). That is, processes associated with the *detection* of self-relevant information seems to be associated with different parts of the brain (i.e., the ventral medial prefrontal cortex, vMPFC) than the *evaluation* of self-relevant information (i.e., the dorsal medial prefrontal cortex, dMPFC; for reviews, see for example Northoff & Bermpohl, 2004; Van der Meer, Costafreda, Aleman, & David, 2010). Comparing fMRI responses observed in these two areas allows us to establish the extent to which people detect information as being self-relevant, and separate this from their tendency to relate this to their actual self-views. In the current research, we thus combined these different indicators of the way participants process self-relevant information: We measured participants' self-reported affective reactions *after* having received either positive or negative feedback about their scores on a measure of their morality and competence. In addition, we measured their skin conductance to assess physiological arousal (Study 6.1) and used fMRI to examine mental processing (Study 6.2) *while* receiving morality and competence feedback.

Mental Processing of Self-relevant Information

Previous neuroscientific research has examined the neural networks involved in processing information relevant for the self. Prior research has addressed the brain regions involved in the assessment of self-relevant information (i.e., processing information that people perceive as related to the self; Northoff & Panksepp, 2008; Schmitz & Johnson, 2007), and reported networks including both subcortical and cortical regions (e.g., caudate nucleus, amygdala, Insula, and anterior singulate cortex [ACC]; Schmitz & Johnson, 2007). Moreover, there is high consensus on the role of the medial prefrontal cortex (MPFC) in processing such self-relevant information (e.g., Abraham, 2013; Ochsner et al., 2005; Northoff & Bermanpohl, 2004; Schmitz & Johnson, 2007). In fact, Moran, Macrae, Heatherton, Wyland, and Kelley (2006) showed that MPFC activation during self-referencing was affected by self-relevance. That is, activation in the MPFC was greater when participants judged personality characteristics (i.e., traits words such as “honest”) as high self-relevant as compared to low self-relevant. In line with these findings, and given that we expect that information concerning morality is more self-relevant than information concerning competence, we will examine whether receiving feedback about one’s morality is associated with greater activation in the MPFC than receiving feedback about one’s competence. appraisal

Although activation in the MPFC is found in many studies concerning self-relevance in general, subregions within the MPFC seem to be associated with more specific processes. For example, in their review, Amodio and Frith (2006) discuss that whereas the posterior rostral region of the MFC is activated during action-monitoring tasks, the anterior rostral MFC is activated during tasks involving self-knowledge, person perception and mentalizing. Moreover, Van der Meer et al. (2010) made a distinction between the ventral and dorsal part of the MPFC and argued that the vMPFC is associated with detecting and labeling self-relevant information, and the dMPFC with evaluation and decision-making processes in self-referential thinking. In the current research, in which participants are only asked to passively view their scores on a measure indicative of their moral and competence behavior, we hypothesize that information concerning morality will be

perceived as more self-relevant than information concerning competence which could thus be associated with activation in the ventral MPFC.

Current Research

The current research aims to investigate whether the differential diagnosticity of morality and competence that is found in impression formation of others is also evident when people are informed about their *own* morality and competence. Based on social psychology research, which has shown that people perceive moral traits as more significant for their social and personal identity than traits concerning competence (e.g., Leach et al., 2007; Ellemers et al., 2008), we predict that receiving information concerning one's own morality (as compared to one's competence) is associated with increased self-reported emotional responses, arousal (assessed by a measure of skin conductance) and greater activation in the MPFC. In addition, impression formation research (e.g., Skowronski & Carlston, 1987) has revealed that negative, rather than positive, information is perceived as a better indication of someone's moral integrity. Conversely, positive rather than negative, information tends to be perceived as a better indication of someone's competence. Drawing on these findings relating to impression formation of others, we predict parallel effects when people receive evaluative information about the self. This is why we anticipate the valence of self-related information to interact with the dimension (competence vs. morality) to which this information pertains.

Study 6.1

Method

Participants.

Thirty three students (six males, $M_{\text{age}} = 18.9$, $SD = 1.45$) from Leiden University participated in the study in return for course credits or money. Five participants were not included in the SCR data analyses because of technical failures in the equipment or software; three other participants were excluded from the SCR data analyses because the signal was extremely noisy, and one other participant was excluded from the SCR data analyses since we could not measure a skin conductance signal. Participants were randomly assigned to one of two conditions: They either received positive or negative feedback (i.e., measured between participants) concerning their morality and competence (measured as a

within-participants factor). To enhance the credibility of the feedback provided, in both experimental conditions the valenced feedback was interspersed with evaluatively neutral feedback.

Procedure.

The feedback participants received was said to be based upon their performance on an Implicit Association Test (IAT; Greenwald et al., 1998) which participants completed in the first part of the experiment. The (non-) Muslim IAT in the current study has previously been used to examine whether people adjust their performance when the test is presented as indicative of their morality (i.e., by informing participants that the test can assess their moral values concerning egalitarianism and discrimination) or of their competence (i.e., by informing participants that the test can assess their ability to process information and learn new tasks; Van Nunspeet et al., 2014). Moreover, since this previous research has shown that participants indeed perceive the test as a credible measure of both properties, we implemented the IAT in the current research as a task on which we could present participants with feedback about their moral values as well as their competence in displaying accurate responses. Importantly, in the current study, participants were informed about these test implications *after* they had finished the IAT, right before they received their feedback to keep task motivation and effort constant across experimental conditions.

The IAT included pictures of female faces with and without a headscarf that had to be associated with positive and negative images (International Affective Picture System; Lang et al., 2005). Congruent IAT trials were trials on which participants were asked to press one response key when viewing both female faces with a headscarf and negative pictures and another key when viewing female faces without a headscarf and positive pictures. Incongruent trials were trials on which the same response key had to be pressed for pictures of female faces with a headscarf and positive pictures and another key when viewing female faces without a headscarf and negative pictures. In order to present participants with several instances of feedback (i.e., necessary for reliable skin conductance data), they performed 20 test blocks of the IAT; each test block consisted of eight trials.

After the IAT participants were informed about the implications of the test. That is, they were led to believe that the test is able to assess both their level of competence (tested as their ability to quickly process new information and to learn new tasks), as well as their level of morality (i.e., their moral values concerning egalitarianism and discrimination). Moreover, participants read that their scores on these two test domains would be provided relative to the scores of other university students and could thus give an indication whether they had scored above average (positive feedback, indicating relatively high moral values or competence), below average (negative feedback, indicating relatively low moral values or lack of competence), or whether their scores were average for the student population (neutral feedback). Neutral feedback was included to enhance credibility of the cover story, and as a control - to be able to check whether above or below average scores affected participants more than average (evaluative neutral) scores. The valence of the feedback was manipulated between-participants - since we did not think it would be credible to provide participants with both above and below average scores on a single measure.

Scores were preprogrammed and represented by colored bars in a normal distribution in which the right hand side displayed above average scores related to morality (or competence) and the left hand side below average scores related to immorality (or incompetence). The participant's score was indicated by a red (negative), green (positive) or yellow (neutral) bar in the normal distribution and the text "your score" right above it (see Figure 6.1).

Participants either received positive (and neutral) or negative (and neutral) feedback. Each round of feedback was provided in two blocks in which one block concerned feedback related to one's morality and the other block feedback related to one's competence. Before each block, participants read the information concerning the nature of the task domain under examination (competence or morality). The order of the feedback blocks was counterbalanced between participants. Each block consisted of ten rounds of valenced (positive or negative) feedback interspersed with ten rounds of neutral feedback. Every feedback round consisted of a screen stating that participants' next test score (concerning their morality or competence) was being computed (9 - 11 sec.), followed by a screen

providing the presentation of the feedback (3 sec.). After viewing their test score for three seconds, participants could press a key to go to the next round of feedback (see Figure 6.1).

Skin conductance was assessed during the IAT as well as the feedback phase to enable participants to get used to the equipment that was attached and to avoid drawing particular attention to a particular part of the experiment as being of special interest. After completing the IAT and before the feedback was provided, the waiting time was used to derive a baseline measure for skin conductance. After having received all the feedback, participants were asked to complete some self-report questionnaires (see details below). The experiment lasted approximately thirty minutes in total, after which participants were properly debriefed about the bogus feedback and the actual goal of the study. They were then thanked and received their incentive.

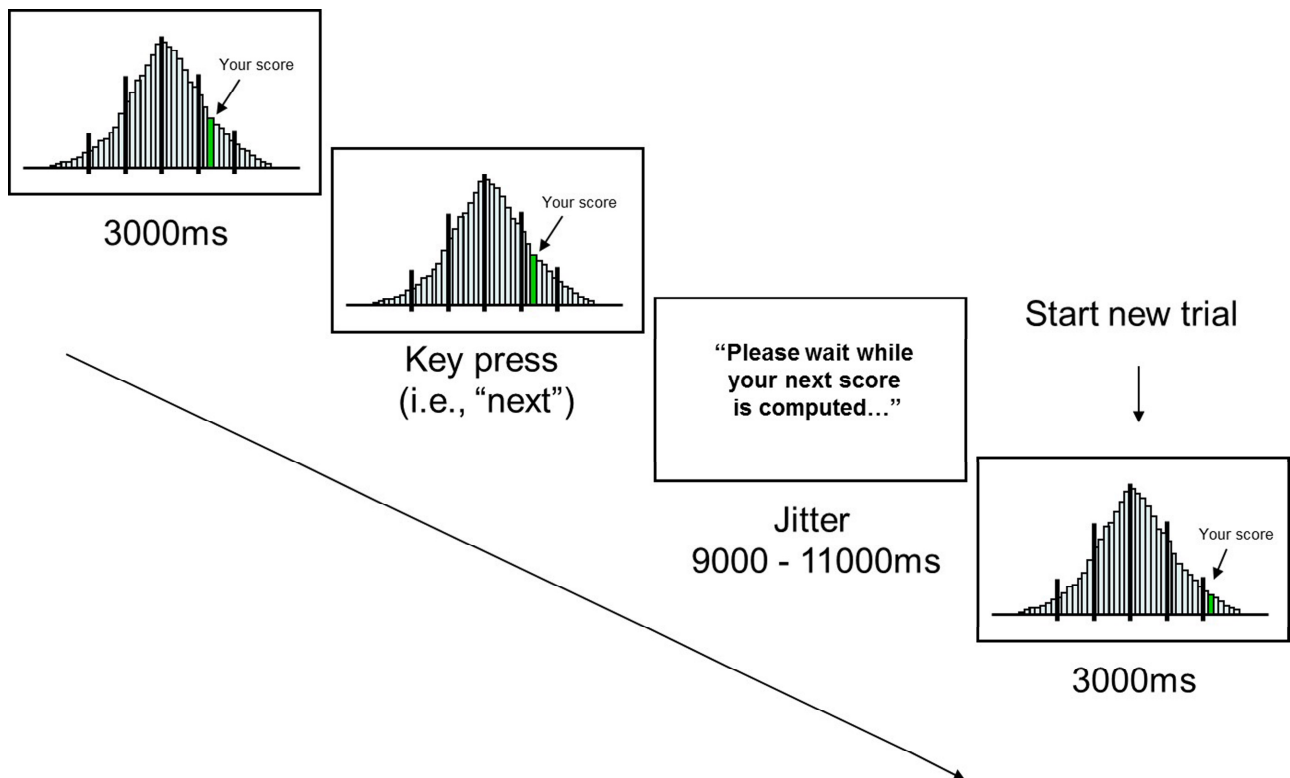


Figure 6.1. Example trial of feedback presented in Study 6.1.

Skin conductance acquisition and processing.

Skin conductance was measured using two pregelled disposable Ag-AgCl electrodes attached to the medial phalanx surfaces of the middle and index fingers of the non-dominant hand. The transponder unit relayed skin conductance data to a host computer running AcqKnowledge software, which logged every feedback stimulus-onset on the skin conductance signal. The data were filtered online with a low pass filter of 2 Hz and offline with a low pass filter of 0.33 Hz. The data was processed in two ways: We measured whether the feedback resulted in an elevated skin conductance level (SCL) compared to baseline, and we determined whether each feedback trial resulted in elevated skin conductance responses (SCRs). For the first measure, we computed difference scores between the average SCL in a 0-6 seconds time window after stimulus-onset in comparison to the average SCL in the final 30 seconds of the baseline measure, separately for each type of feedback (i.e., neutral and valence feedback concerning morality and competence). For the second measure we detected SCRs with a minimum amplitude change of 0.01 μ S after stimulus-onset, and measured the number of SCRs in a time window between 1 and 6 seconds after each stimulus-onset. When there was no SCR associated with the feedback-stimulus, “0” was recorded. The mean number of SCR’s was then calculated separately for feedback indicating scores on morality and competence and separately for neutral and valenced feedback. It should be noted that since many participants failed to generate SCR’s related to the feedback, the mean number of fluctuations was below 1.0 (which is in line with previous research; e.g., Lawrence et al., 2006).

Self-reports.

Checks. We first checked whether participants had experienced the task in a similar way and were equally uncertain about their performance, regardless of whether they had received positive or negative scores. For this purpose, after having received all of their feedback, we asked participants to answer two questions about their experience while performing the IAT (i.e., “I was insecure about my performance on the test” and “During the test, I had the feeling I was able to perform very well” [recoded], $r = .44$, $p = .011$). They could indicate their answers on a 7-point Likert scale (1 = completely agree – 7 = completely disagree).

Self-reported negative emotional response. We then asked participants to reconsider how they felt while receiving the feedback, and to indicate the emotional response this raised. Items asked participants to indicate their general feelings (i.e., “Seeing my scores gave me a bad feeling”; “My scores gave me the idea that I don’t have good qualities”; “Seeing my scores gave me a good feeling”; “My scores made me feel good about myself”) as well as a number of specific emotions (i.e., “When I received feedback concerning the morality/competence domain of the test, I felt: discouraged / nervous / guilty / ashamed / threatened / frustrated / happy / relaxed / motivated / proud / enthusiastic / challenged”). All answers were assessed using 7-point Likert scales (1 = completely agree – 7 = completely disagree). All these questions were asked twice: Once to indicate emotional responses to morality feedback and once to convey emotional responses to competence feedback. Items concerning positive feelings and emotions were recoded so that higher scores always indicated a more negative emotional response. We then combined the items concerning general feelings and specific emotions for each type of feedback, resulting in two overall indicators. One combined score indicated the degree to which participants reported a negative emotional response when viewing their scores on morality ($\alpha = .94$) the other indicated negative emotional responses when viewing their competence scores ($\alpha = .91$).

Results

Checks.

To check whether participants were equally uncertain about their task scores so that the feedback they received seemed credible regardless of experimental condition, we asked participants to indicate their thoughts about their performance during the IAT. Results of a one-sample T-test with the mean of the scale (4) as the test value showed that, overall, participants reported to be quite insecure about their performance ($M = 4.77$, $SD = 1.29$; $t[32] = 3.45$, $p = .002$). There were no differences between experimental conditions, suggesting that below or above average test scores would seem equally plausible.

Skin conductance data.

Skin conductance level (SCL). To test whether the feedback presented during the experiment affected participants’ arousal levels (irrespective of valence

or task domain), we first tested the difference between the average SCL following the feedback (i.e., 0-6 seconds after stimulus-onset, across all types of feedback) and the average SCL during the final 30 seconds of the baseline. Results of a paired sample T-test revealed that, as intended, the feedback significantly increased SCL as compared to baseline, $M_{\text{difference}} = 0.64$, $SD = 1.38$, $t[23] = 2.26$, $p = .03$.

To examine any differences in SCL between the types of feedback, we conducted a repeated measures ANOVA on the difference scores of SCL (0-6 seconds after stimulus-onset minus baseline) with the type of feedback (valenced/neutral) and task domain (morality/ competence scores) as repeated measures, and the context in which feedback was provided (positive/ negative feedback condition) and order (morality/competence block first) as between-groups factors. Results revealed a significant main effect of feedback type; $F(1, 20) = 11.45$, $p = .003$, $\eta_p^2 = .36$, indicating that SCL was greater after valenced ($M = 0.68$, $S.E. = 0.31$) compared to neutral feedback ($M = 0.56$, $S.E. = 0.30$). This main effect was however qualified by a significant feedback type*order interaction effect; $F(1, 20) = 7.46$, $p = .01$, $\eta_p^2 = .27$, revealing that the difference between valenced and neutral feedback was only significant when the scores concerning morality were presented first; $F(1, 20) = 16.33$, $p = .001$, $\eta_p^2 = .45$. The other simple main effects were not significant; all F 's < 1 . There were no interaction effects with task domain, indicating that there were no differences in average SCL between positive/negative or neutral feedback related to morality and competence.

Skin conductance responses (SCRs). To examine whether the different types of feedback affected skin conductance directly after stimulus-onset, we also analyzed SCRs. We assessed differences in SCRs during the feedback round with a repeated measures ANOVA with the type of feedback (valenced/neutral) and task domain (morality/ competence scores) as repeated measures, and the context in which feedback was provided (positive/ negative feedback condition) and task domain (morality/competence scores) as repeated measures, and valence (positive vs. negative feedback) and order (morality vs. competence block first) as between-groups factors. Results revealed no difference in SCR's between valenced and neutral feedback; $F(1,20) = 2.32$, $p = .14$. However, we found evidence in support of our central prediction, indicating that feedback relating to morality had a greater

impact than feedback referring to competence: We observed a marginally significant main effect of task domain; $F(1,20) = 3.90, p = .06, \eta^2_p = .16$, indicating that there were more SCRs when participants were confronted with their morality ($M = 0.36, S.E. = .03$) than competence scores ($M = 0.29, S.E. = 0.04$). This effect was qualified by a significant interaction effect between task domain and order; $F(1,20) = 5.19, p = .03, \eta^2_p = .21$, indicating that a significant difference in SCRs between morality and competence feedback only emerged when the morality scores were presented first (i.e., increased SCR's in the morality [$M = 0.40, S.E. = .05$] compared to the competence block [$M = 0.21, S.E. = .08$], $F[1,20] = 7.90, p = .01, \eta^2_p = .28$). When competence scores were presented first there was no difference in responses to the different task domains ($[M_{\text{morality}} = 0.31, S.E. = .04; M_{\text{competence}} = 0.32, S.E. = .07]; F < 1$). Additionally, we observed a trend towards a three-way interaction between task domain, order and valence; $F(1,20) = 2.99, p = .10, \eta^2_p = .13$. Examination of the repeated measures ANOVA separately for the positive and negative feedback conditions revealed that the task domain x order interaction effect could only be traced to the negative feedback condition; $F(1,11) = 7.36, p = .02, \eta^2_p = .40$, but not the positive feedback condition ($F < 1$; see Figure 6.2).

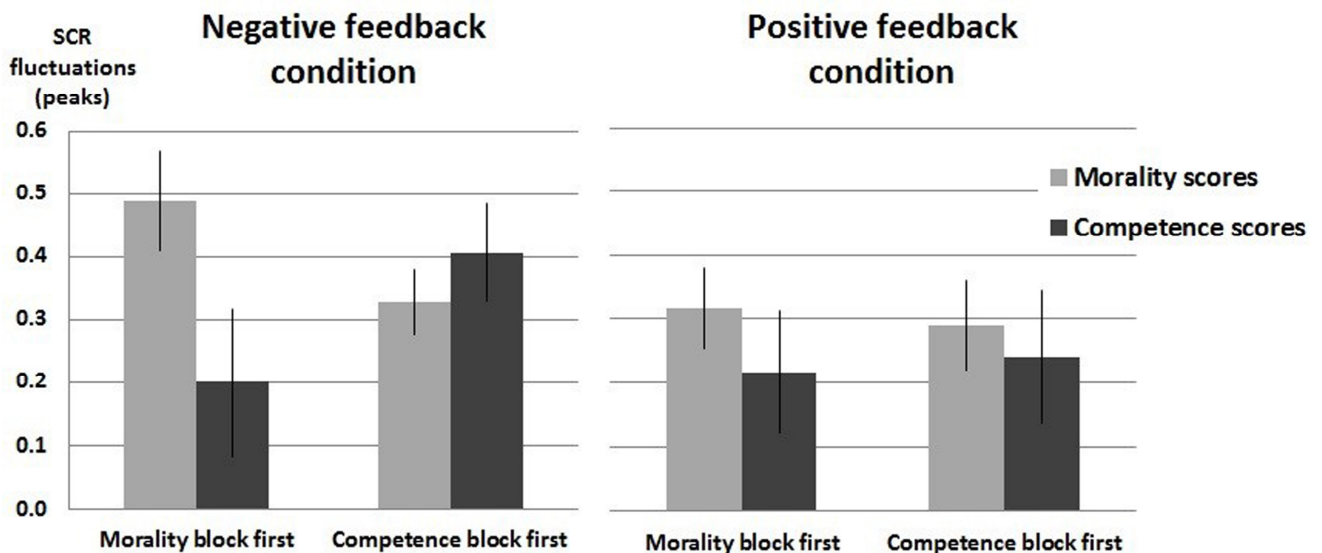


Figure 6.2. Average skin conductance responses (SCRs) in each condition. Whereas there were no differences in SCRs for positive feedback (right), negative feedback concerning morality was associated with increased physiological arousal –in case morality scores were presented first (left).

Self-reported negative emotional response.

After participants had received all of their feedback, we asked them to think back about the moments they received feedback about their morality and competence and to recall and report their emotional response. A repeated measures ANOVA with task domain (morality/competence) as the repeated measure and valence (positive/negative feedback)¹³ as between-participants factor, revealed evidence in support of our reasoning. We observed a significant interaction effect between task domain and valence; $F(1,31) = 4.00, p = .05, \eta^2_p = .11$. The relevant means and analysis of simple main effects confirmed that the difference between positive and negative feedback conditions in self-reported emotional response was more pronounced when participants received feedback regarding their morality; $M_{\text{difference}} = 1.82, S.E. = 0.24; F(1,31) = 60.02, p < .001, \eta^2_p = .66$, rather than their competence; $M_{\text{difference}} = 1.44, S.E. = 0.22; F(1,31) = 42.37, p < .001, \eta^2_p = .58$. Specifically, when participants had received negative feedback they reported a more negative emotional response when the feedback was related to their morality ($M = 3.45, S.E. = 0.17$) rather than their competence ($M = 4.07, S.E. = 0.16$); $F(1,31) = 6.95, p = .01, \eta^2_p = .18$. There was no difference between responses to positive feedback depending on whether this pertained to the morality or the competence domain ($F < 1$).

Taken together, the findings of Study 6.1 offer evidence in line with our reasoning, as they suggest that receiving information related to one's morality has more impact on participants' responses than feedback related to their competence, in particular when people are confronted with negative feedback. To examine whether feedback concerning one's morality (as compared to competence) is also processed differently in the brain, we conducted an fMRI study in which we examined the neural network involved in processing self-relevant information.

¹³ Note that we did not include a factor distinguishing between valenced and neutral feedback in this analysis, because we asked participants how they felt about their feedback overall, which was predominantly negative (negative and neutral) in the negative feedback condition, and predominantly positive (positive and neutral) in the positive feedback condition.

Study 6.2

Method

Participants.

Forty right-handed students (12 males, $M_{\text{age}} = 21.7$ years, $SD = 3.1$) from Leiden University participated in the study in return for course credits or money. None of the participants reported a history of psychiatric or neurological disorders, and current use of any medications. One participant was excluded from the analysis of the behavioral data because she failed to detect the color change of the fixation cross (whereas all other features of the stimuli were clear). Three other participants could not be included in the fMRI analyses because of technical problems. Participants were randomly assigned to the positive or negative feedback condition. All procedures were approved by the medical ethical committee of the Leiden University Medical Center (LUMC) and all participants gave informed consent for the study.

Procedure.

Before the scanning session, participants performed the (non-)Muslim IAT without receiving any information about the implications of the test, similar to Study 6.1. During the scanning session, participants were first informed that the test was able to assess both their level of competence, as well as their level of morality. In contrast to Study 6.1, participants thus read about *both* types of implications before they received any of the feedback stimuli. Participants were presented with the same feedback stimuli as used in Study 6.1.

Participants were informed about both types of test implications at once because the current study used an event-related block design: Feedback was provided in one run in which 6 blocks of feedback concerning morality were alternated with 6 blocks of feedback concerning competence. Each block consisted of 5 feedback trials of which two or three trials provided valenced feedback (positive or negative, depending on experimental condition) and two or three trials provided neutral feedback. The reason for presenting the competence and morality trials in mini blocks was to ensure direct repetition of each task domain, in order for the feedback to have impact on participants (which was similar to the block

design used in Study 6.1). In total there were 15 trials per feedback type (morality-valence/morality-neutral/competence-valence/competence-neutral)¹⁴.

Each feedback round consisted of a screen stating that participants' next test score concerning their morality or competence was being computed (2 sec.), a fixation cross (jittered duration, 4-8 sec.), and the feedback stimulus (3 sec., see Figure 6.3). To ensure that participants were attentive, they were asked to press a key (with their right index finger) whenever the fixation cross changed color, which happened randomly after 1 to 5 seconds.

As part of a larger study, the scanning session lasted approximately one hour. After the scanning session had ended, participants were asked to fill out some questionnaires. The complete study lasted approximately 2 hours, after which participants were properly debriefed, thanked and given their incentive.

fMRI data acquisition and processing.

Scanning was performed at the Leiden University Medical Centre (LUMC) with a standard whole-head coil on a 3.0 Tesla Philips Achieva scanner. Using E-prime 2.0 software, the task instructions and feedback was projected onto a screen at the back of the scanner bore, which participants could view via a window attached to the top of head coil. Participants could respond by pressing a button (using their right index finger) on a box attached to their right leg. The feedback was provided in one run, lasting approximately 15 minutes. Functional data were obtained using T2*-weighted echo-planar imaging (EPI), repetition time (TR) = 2200 ms, echo time (TE) = 30 ms, slice matrix = 80 x 80, slice thickness = 2.75 mm, slice gap = 0.28 mm, field of view [FOV] = 220 mm). A high-resolution T2-weighted anatomical scan (same slice prescription as EPI) was collected at the end of the scanning as well as a high resolution 3D T1-weighted anatomical image (TR = 9.751 ms, TE = 4.59 ms, flip angle = 8°, 140 slices, 0.875 mm x 0.875 mm x 1.2 mm, and FOV = 224.000 x 168.000 x 177.333).

¹⁴ The order of the blocks of feedback was not counterbalanced between participants (i.e., the first five feedback trials always concerned participants' morality and the following five participants' competence), which could have affected the results. We therefore also analyzed the data without the first ten trials to control for the possible high impact of these initial scores. Results of this analysis were similar to the ones described in the current results section.

Data were preprocessed and analyzed using SPM8 software (Wellcome Department of Cognitive Neurology, London) implemented in MATLAB (Mathworks, Sherborn, MA). The functional time series were realigned to compensate for small head movements.

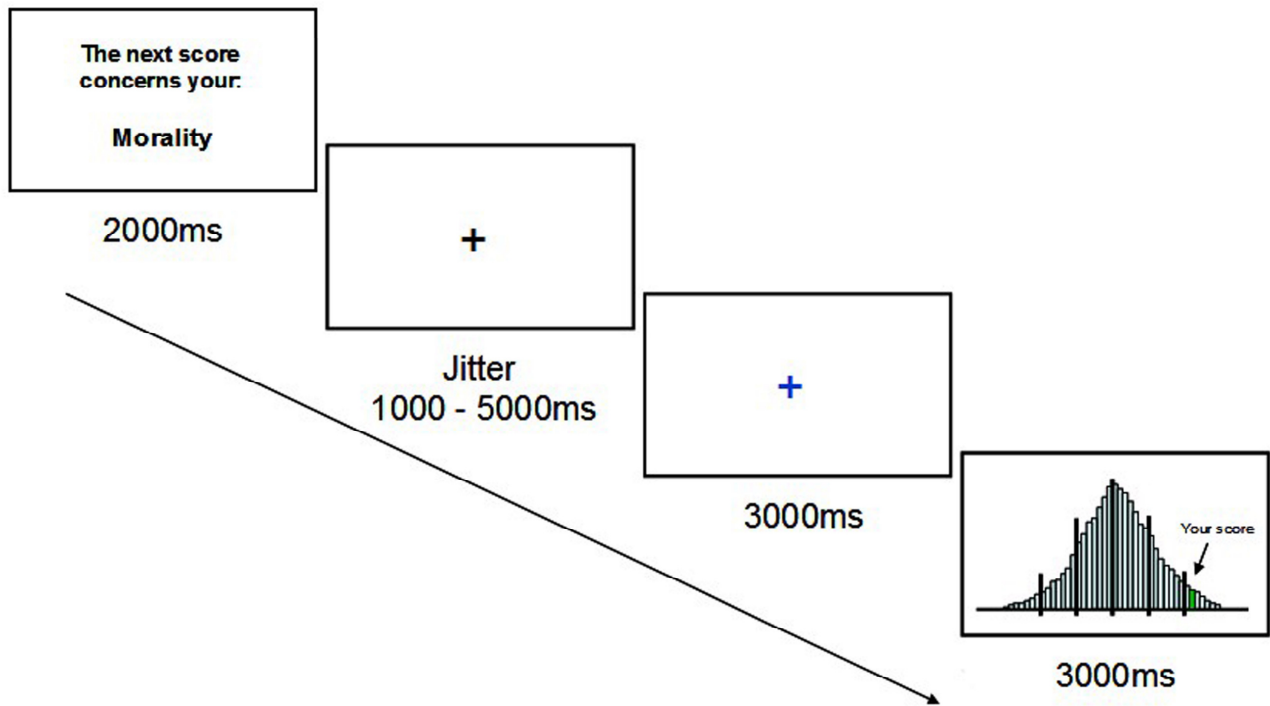


Figure 6.3. Example trial of feedback presented in Study 6.2.

Translational movement parameters never exceeded 1 voxel (< 3 mm) in any direction for any subject or scan. Functional volumes were spatially normalized to EPI templates. The normalization algorithm used a 12 parameter affine transformation together with a nonlinear transformation involving cosine basis functions and resampled the volumes to 3 mm cubic voxels. Functional volumes were spatially smoothed using an 8 mm full-width half-maximum Gaussian kernel. Templates were based on the MNI305 stereotaxic space (Cocosco, Kollokian, Kwan, Pike, & Evans, 1997), and the Montreal Neurological Institute (MNI) atlas was used to refer to the coordinates.

To analyze the data, a canonical hemodynamic response function was convolved at the onset of the feedback stimulus and modeled as a zero-duration event. We distinguished between four conditions *within* participants: Valence versus neutral feedback and feedback related to morality or competence. Whether the valence was positive or negative was a *between*-participants manipulation. These conditions resulted in four 2 X 2 full factorial designs. Two designs were used to examine the effects of valenced and neutral feedback for the positive and negative feedback conditions separately, resulting in two 2 (Feedback: Valence/Neutral) X 2 (Task Domain: morality/competence) ANOVAs which were run separately for the positive feedback condition and the negative feedback condition. Two other designs were used to directly compare the effects of positive versus negative feedback, resulting in a 2 (Valence Feedback: positive/negative) X 2 (Task Domain: morality/competence) ANOVA and a 2 (Neutral Feedback: positive/negative condition) X 2 (Task Domain: morality/competence) ANOVA. These ANOVAs concerned a comparison between groups.

The analyses were carried out using the general linear model in SPM8. For each individual, contrast parameter images were computed and the resulting contrast images were submitted to second-level group analyses. Only effects of at least 10 continuous voxels that exceeded a False Discovery Rate (FDR) corrected threshold of $p < .05$ are reported.

Moreover, since we were interested in the –perhaps more subtle– difference between receiving feedback about morality or competence, we extracted parameter estimates from the regions of interest (ROI) that were identified in the whole brain analyses to explore the pattern of the activation across our conditions. We extracted the mean parameter estimate within each ROI for each condition, reducing the ROI to a single data point. This is a common approach in cognitive neuroscience which has two advantages: (1) it reduces the number of comparisons, and (2) collapsing across voxels within the region decreases noise (Poldrack, 2007). We focused specifically on the MPFC in the contrast positive versus negative feedback. However, activation in MPFC was part of a larger network (see Table 6.1). To isolate the activation cluster within the MPFC, we adjusted the threshold to $p < .01$ (FDR corrected, 10 continuous voxels, see Table 6.2). The ROI analysis

was used to gain functional specificity in the regions that were already a priori defined as regions of interest. This region was used to test the hypothesis that valenced feedback would be associated with differential activity in the morality versus competence condition. These regions were extracted using the Marsbar toolbox (Brett, Anton, Valabregue, & Poline, 2002) for SPM8.

Self-reported negative emotional response.

To examine participants' negative emotional response related to the moment they received their feedback, we used the same scales as described in Study 6.1: A scale measuring participants negative emotional response concerning their scores on morality ($\alpha = .90$) and a scale measuring participants negative emotional response concerning their scores on competence ($\alpha = .89$). These self-reports were administered after the scanning session.

Results

Behavioral data.

Since we asked participants to press a key whenever the fixation cross changed color (primarily to keep them attentive during the scanning session), we could test whether their response latencies differed between morality and competence trials. Indeed, a 2 (Feedback Type: positive/negative between-participants factor) x 2 (Task Domain: morality/competence within-participants factor) repeated measures ANOVA revealed a significant interaction effect; $F(1,37) = 9.54, p = .004, \eta_p^2 = .21$. This indicated a significant reversal in the direction of the effects in the morality condition compared to the competence condition (see Figure 6.4). As a result, participants who received negative feedback responded significantly slower on morality ($M = 474.82, SD = 115.81$) than on competence trials ($M = 450.39, SD = 105.35$); $F(1,37) = 6.08, p = .02, \eta_p^2 = .14$. In contrast, participants in the positive feedback condition responded somewhat more slowly on trials concerning competence ($M = 464.86, SD = 95.41$) than morality ($M = 444.24, SD = 66.77$); $F(1,37) = 3.71, p = .06, \eta_p^2 = .09$.

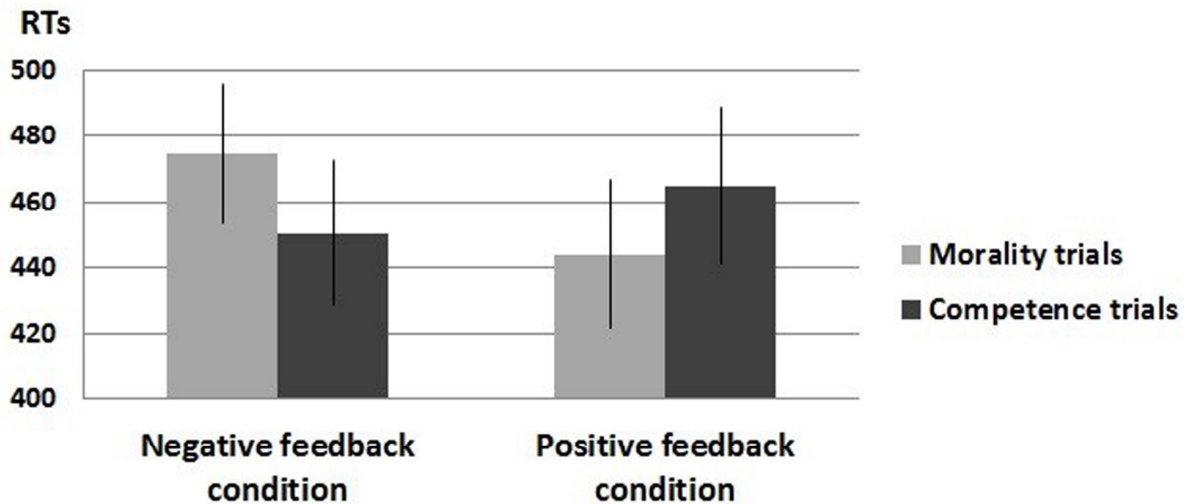


Figure 6.4. Interaction effect between reaction times (RTs, in milliseconds) on morality and competence trials: Whereas participants who received negative feedback responded more slowly on morality as compared to competence trials, participants who received positive feedback responded more slowly on competence as compared to morality trials.

fMRI data.

Whole brain level. To examine neural activation associated with receiving positive or negative and neutral feedback about one's morality and competence, we conducted four ANOVAs. First, we examined the effects of valenced and neutral feedback about morality and competence separately for the positive and negative feedback conditions. The results of these two 2 (Feedback: Valence/Neutral) X 2 (Task Domain: morality/competence) full factorial ANOVAs revealed no significant effects. Second, we examined neural differences between receiving positive versus negative feedback, by selecting only valenced trials. This 2 (Valenced Feedback: positive/negative) X 2 (Task Domain: morality/competence) ANOVA resulted in a main effect of valence (see Table 6.1): Activation in the amygdala, insula, bilateral inferior frontal gyrus, and ventral and dorsal MPFC was greater for participants who received positive feedback than for participants who received negative feedback. There was no main effect of Task Domain, nor an interaction effect. Third, we examined neural activation associated with receiving neutral feedback (i.e., only trials with neutral feedback were selected). This 2 (Neutral Feedback: positive/negative condition) X 2 (Task Domain:

morality/competence) full factorial ANOVA did not show any relevant significant activation (see Table 6.3).

Taken together, the contrast positive versus negative feedback resulted in activation in the expected brain network associated with processing self-relevant information. At the whole brain level the neural activation was not different for morality versus competence trials. In the next section, we describe the results from more fine grained ROI analyses, using the contrast positive > negative feedback as a functional localizer.

Regions of interest. To examine the difference between feedback related to morality and competence, we conducted ROIs analyses of the ventral MPFC (vMPFC), a target brain area showing increased activation for positive compared to negative feedback. Results revealed an interaction effect between feedback and task domain in the vMPFC ($F[1,35] = 4.06, p = .05, \eta_p^2 = .10$). Consistent with our hypothesis that information concerning one's morality has a greater impact than information concerning one's competence, we found that the difference between positive and negative feedback was more pronounced for scores concerning morality; $F(1,35) = 14.90, p < .001, \eta_p^2 = .30$, than for scores concerning competence; $F(1,35) = 7.53, p = .01, \eta_p^2 = .18$ (see Figure 6.5). Moreover, within the positive feedback condition, activation in the vMPFC was greater when participants viewed their scores concerning morality as compared to competence; $F(1,35) = 3.48, p = .07, \eta_p^2 = .09$. This difference was not significant in the negative feedback condition; $F(1,35) < 1$.

Self-reported negative emotional response. Results of a repeated measures ANOVA with task domain (morality/competence) as the repeated measure and valence (positive/negative feedback) as between-groups factor, supported our reasoning and were consistent with Study 6.1: We observed a significant interaction effect between task domain and valence; $F(1,38) = 4.84, p = .03, \eta_p^2 = .11$. The relevant means and analysis of simple main effects confirmed that the difference between positive and negative feedback conditions in self-reported emotional response was more pronounced when participants received feedback regarding their morality; $M_{\text{difference}} = 1.45, S.E. = 0.22; F(1,38) = 44.24, p < .001, \eta_p^2 = .54$, rather than their competence; $M_{\text{difference}} = 0.82, S.E. = 0.26; F(1,38) = 10.24, p =$

.003, $\eta^2_p = .21$. Specifically, when participants had received negative feedback they indicated a more negative emotional response when the feedback was related to their morality ($M = 4.14$, $S.E. = 0.15$) rather than their competence ($M = 3.59$, $S.E. = 0.18$); $F(1,38) = 7.98$, $p = .01$, $\eta^2_p = .17$. There was no difference between responses to positive feedback when comparing the morality with the competence domain ($F < 1$).

Table 6.1.

Brain regions revealed by the main effect of Valence in the 2 (Valenced feedback: positive/negative feedback) \times 2 (Task Domain: morality/competence) ANOVA at whole brain level.

Anatomical Region	L/R	voxels	Z	MNI coordinates		
				x	y	z
Medial Orbital Prefrontal Cortex	R	51	3.43	30	47	-5
			3.27	36	29	-14
			3.12	27	53	1
Dorsal Medial Prefrontal Cortex	L	97	4.01	-12	26	34
			3.26	-5	38	31
Dorsal Lateral Prefrontal Cortex	L	164	4.53	-36	35	13
			3.09	-45	14	19
			2.72	-57	11	22
Superior Frontal Gyrus	L	27	3.34	-18	26	55
Supplementary Motor Area	R	16	2.85	3	8	61
Middle Temporal Gyrus	R	11	2.74	48	-58	19
ParaHippocampal Gyrus	L	18	3.11	-24	-37	-8
Calcarine/Inual Gyrus	R	5575	5.66	15	-88	10
			5.26	18	-55	-2
			5.07	15	-64	16
Middle Occipital Gyrus	L	44	3.73	-30	-88	19
			3.47	-27	-91	10
			2.74	-33	-73	28

MNI coordinates for main effects, peak voxels reported at $p < .05$, FDR corrected, at least 10 contiguous voxels (voxels size was 3.0 x 3.0 x 3.0 mm).

Table 6.2.

Brain regions revealed by the main effect of Valence in the 2 (Valenced feedback: positive/negative feedback) × 2 (Task Domain: morality/competence) ANOVA at whole brain level.

Anatomical Region	L/R	voxels	Z	MNI coordinates				
				x	y	z		
Ventral Medial Prefrontal Cortex	R	88	4.34	0	59	-2		
			4.29	-12	59	10		
			4.09	9	59	-2		
Dorsal Lateral Prefrontal Cortex	L	27	4.53	-36	35	13		
	R	12	3.97	48	20	28		
Rolandic Operculum/Precentral Gyrus	R	108	4.74	51	-13	19		
			4.37	48	5	37		
Pre-/Postcentral Gyrus	L	19	4.40	-54	2	40		
			3.81	-51	-10	37		
Superior Temporal Gyrus	R	23	4.17	57	-4	4		
Middle Temporal Gyrus	L	19	3.91	-45	-67	19		
Superior Parietal Lobule	L	72	4.72	-24	-58	55		
Superior Parietal Lobule / Cuneus	L	72	4.72	-24	-58	55		
			R	130	4.77	12	-85	31
			4.51	15	-64	55		
Precuneus	L	16	4.29	18	-58	46		
			3.78	-6	-58	37		
			3.49	-12	-58	31		
Calcarine/Inual Gyrus	R	379	5.66	15	-88	10		
			5.26	18	-55	-2		
			5.07	15	-64	16		
Middle Occipital Gyrus	R	16	4.46	30	-79	31		
Insula	L	52	4.12	-33	-16	16		
			3.82	-24	-19	19		
Amygdala	R	13	3.93	39	-28	22		
Hippocampus	R	10	4.13	30	2	-14		
	R	21	3.96	24	-34	-5		

MNI coordinates for main effects, peak voxels reported at $p < .01$, FDR corrected, at least 10 contiguous voxels (voxels size was 3.0 x 3.0 x 3.0 mm).

Table 6.3.

Brain regions revealed by the main effect of Valence in the 2 (Neutral feedback: positive/negative condition) \times 2 (Task Domain: morality/competence) ANOVA at whole brain level.

Anatomical Region	L/R	voxels	Z	MNI coordinates		
				x	y	z
Superior Parietal Lobule	R	21	4.52	18	-61	55
			3.95	18	-58	46
			3.68	21	-55	43
Calcarine Gyrus (Occipital Lobe)	R	47	4.30	24	-61	19
			4.22	15	-64	16

MNI coordinates for main effects, peak voxels reported at $p < .05$, FDR corrected, at least 10 contiguous voxels (voxels size was 3.0 x 3.0 x 3.0 mm).

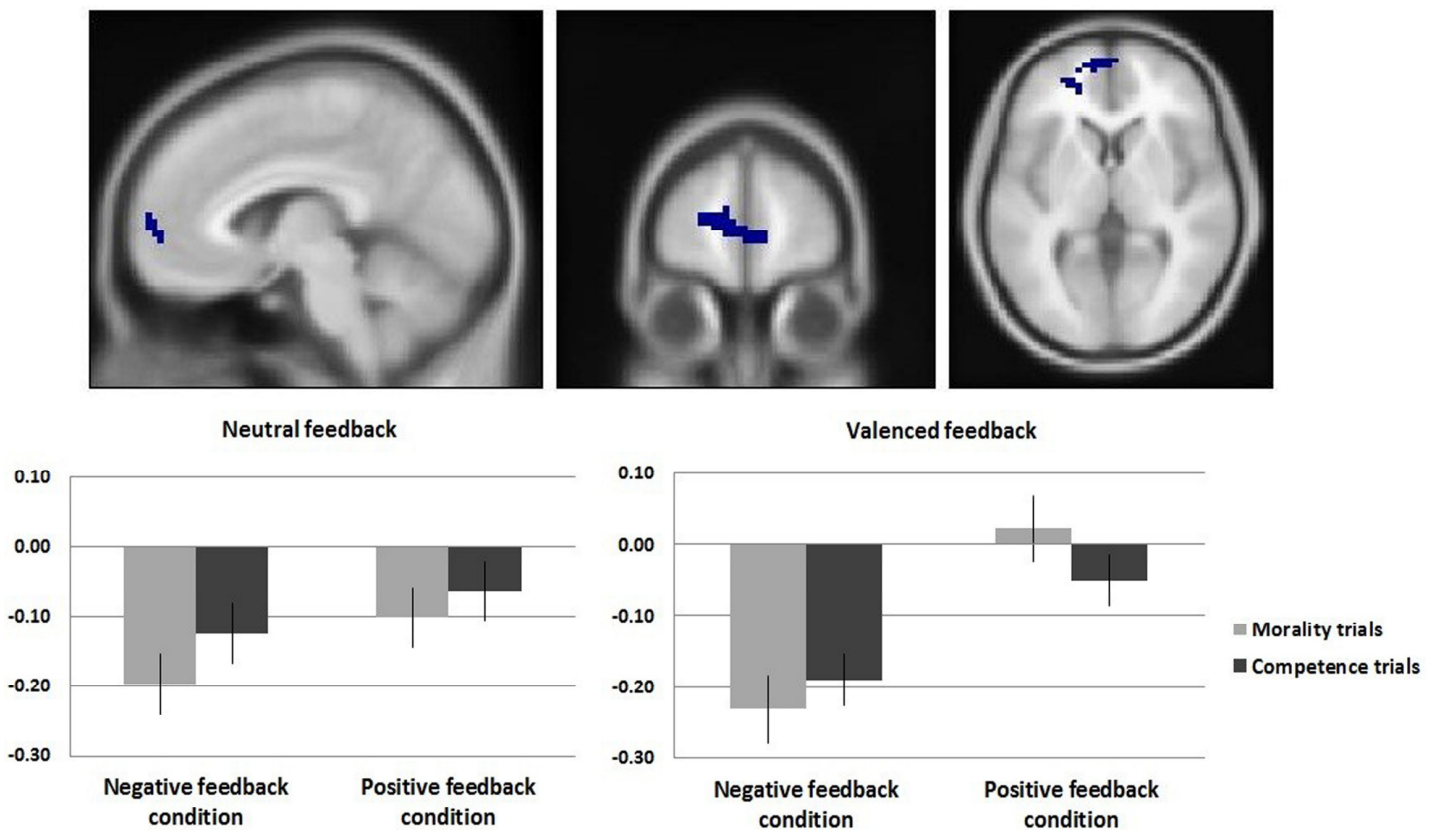


Figure 6.5. Activation in the ventromedial prefrontal cortex (ROI cluster based on a peak voxel, MNI coordinates: x = 0, y = 59, z = -2; $p < .01$, FDR corrected, $p < .01$, at least 10 continuous voxels) revealing the significant interaction between feedback and task domain on valenced feedback trials. There were no effects on neutral feedback trials.

General Discussion

The aim of the present research was to compare the impact of receiving different types of self-relevant information. Specifically, we compared behavioral, self-reported, skin conductance and fMRI responses to information regarding an individuals' own morality and competence. Previous research revealed that when receiving information about another person's morality, negative behaviors are perceived as more informative than positive behaviors. Conversely, in the competence domain, positive information is perceived as more informative than negative information (Skowronski & Carlston, 1987). Importantly, however, this differential diagnosticity has been demonstrated when people form an impression of others. Thus, it is as yet unclear whether a similar asymmetry in the perceived importance of positive and negative information regarding competence and morality is also evident when people process information related to the self.

We examined this in the present research, by confronting participants with information either attesting to or undermining their moral and competent self by giving them positive or negative feedback about their performance on a task that was supposedly indicative of both domains. After having received the feedback, we asked participants to recall their affective responses (i.e., positive and negative emotions) related to the moment of feedback. Additionally, we assessed participants' physiological arousal by assessing their skin conductance levels while they received their feedback (in Study 6.1) and (in Study 6.2) we used fMRI to examine how activation in the neural network involved in processing self-relevant information, was associated with receiving the feedback.

Participants self-reported emotions gave insight into how people reflect upon the information they received about their moral and competent self and thus whether this self-reflection mirrors the asymmetry that has been observed in impression formation of others. The evidence obtained provided partial support for our reasoning regarding the differential diagnosticity of (im)moral and (in)competent information about the self. That is, compared to information concerning competence, information concerning morality had a greater impact upon participants' self-reported emotional response. Especially participants who had received negative feedback reported increased negative affect when the

feedback concerned their morality rather than their competence. These findings extend research about the importance of morality over competence for people's personal and social identity (e.g., Leach et al., 2007; Ellemers et al., 2008).

Interestingly, the results of our (neuro)physiological measures offered additional support for the pattern of differential diagnosticity of (im)moral and (in)competent behaviors found in impression formation research. That is, results of analyses of skin conductance responses revealed that physiological arousal was increased when participants received feedback about their morality as compared to their competence, and this was the case in particular when this feedback had a negative content. (Negative) information about one's own morality thus seemed to be more impactful than information concerning one's competence. These findings thus extend prior research which established the explicit motivation to be (perceived as) moral (e.g., Leach et al., 2007; Ellemers et al., 2008) as they reveal that automatic affective responses are increased when people are confronted with information that calls their morality into question.

In addition, results of the fMRI experiment showed that positive (rather than negative) feedback was associated with greater activation in the amygdala, insula and MPFC. The MPFC has previously been associated with the processing of self-relevant information (e.g., see Abraham, 2013; Moran et al., 2006; Northoff & Berman, 2004; Schmitz & Johnson, 2007). The relative increase in activation in this region for participants who received positive feedback (as compared to participants who received negative feedback) is in line with research showing that people are positively biased when they receive self-relevant information. Specifically, people tend to think they are better than average (especially when the other is a non-specified average student, like in our study; Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995), and expect to receive positive rather than negative feedback in social interactions (Hepper, Hart, Gregg, & Sedikides, 2011). Moreover, prior research has established that positively biased feedback processing is associated with activation in the MPFC (Korn, Prehn, Park, Walter, & Heekeren, 2012). Extending this prior research, our findings thus reveal that positive information concerning one's own behavior is processed as more self-relevant than negative information concerning one's behavior. Moreover, our

results revealed that participants showed more activation in the ventral MPFC (vMPFC) when they received positive feedback concerning their morality as compared to their competence. In line with research suggesting that the vMPFC is associated with the detection and labelling of information relevant to the self (Van der Meer et al., 2010), these findings thus suggest that people detect of the confirmation of one's morality as more self-relevant than confirmation of one's competence.

The findings concerning the impact of negative feedback on affective responses and arousal complement the observed effects of positive feedback in the fMRI results. That is, the skin conductance data in combination with the self-reports suggest that people are emotionally moved by negative feedback concerning their own moral behavior. Additionally, they process positive feedback concerning their own moral behavior as more self-relevant. Across the board, people thus seem more likely to attend and respond to information regarding their morality rather than their competence, which suggests that this process is more complex than the process of impression formation of others: Consistent with impression formation, negative information about the self also has a greater impact when it concerns one's own morality as compared to competence. However, people also seem to be more attentive to positive information concerning their own morality and what this means for their self-view, than that they are focused on possible implications of negative information concerning their own morality. In other words, people are thus particularly attentive to moral information that may help establish a positive self-view. Again, such positive information is most relevant when it concerns one's morality rather than one's competence. At the same time, although people seem to attend less to negative information concerning the self, being confronted with such negative information induces increased arousal and negative emotion. Unfortunately, we cannot directly relate the findings concerning the skin conductance to the fMRI data since we assessed these measures in two separate studies. In order to examine this relation more directly, a measure of skin conductance should be taken *while* participants are being scanned. Nevertheless, different from how we respond to information about others – when negative information is seen as more indicative of another person's morality, our present

observations suggest that we seem to perceive positive information as most relevant to ourselves, especially when this indicates and confirms our moral identity.

Acknowledgements

We thank Thijs Schrama and Maureen Meekel for their assistance with processing the skin conductance data; Maryke Hofman and Lotte van Dillen for their help with the fMRI data collection; and Ilya Veer and Mischa de Rover for their advice concerning the paradigm for the fMRI study.

