

Positivity bias in perceptual matching may reflect a spontaneous self-referential processing

Hu Chuan-Peng<sup>1,2</sup>, Kaiping Peng<sup>1</sup>, & Jie Sui<sup>1,3</sup>

<sup>1</sup> Tsinghua University, 100084 Beijing, China

<sup>2</sup> Leibniz Institute for Resilience Research, 55131 Mainz, Germany

<sup>3</sup> University of Aberdeen, Aberdeen, Scotland

#### Author Note

Hu Chuan-Peng, Department of Psychology, Tsinghua University, 100084 Beijing, China. Kaiping Peng, Department of Psychology, Tsinghua University, 100084 Beijing, China. Jie Sui, School of Psychology, University of Aberdeen, Aberdeen, Scotland.

Authors contribution: HCP, JS, & KP design the study, HCP collected the data, HCP analyzed the data and drafted the manuscript. KP & JS supported this project.

Correspondence concerning this article should be addressed to Hu Chuan-Peng, Langenbeckstr. 1, Neuroimaging Center, University Medical Center Mainz, 55131 Mainz, Germany. E-mail: hcp4715@gmail.com

## Abstract

To navigate in a complex social world, individual has learnt to prioritize valuable information. Previous studies suggested the moral related stimuli was prioritized (Anderson, Siegel, et al., 2011, Science; Gantman & Van Bavel, 2014, Cognition). Using social associative learning paradigm, we found that when geometric shapes, without social meaning, were associated with different moral valence (morally good, neutral, or bad), the shapes that associated with positive moral valence were prioritized in moral matching task. This patterns of results were robust across different procedures. Further, we tested whether this positive effect was modulated by self-relevance by manipulating the self-referential explicitly and found that the positive bias showed a large effect when positive valued stimuli were related to the self. This effect exist also when the self related information were presented as a task-irrelevant information. We also tested the specificity of the positive valence and found that this effect was not limited to moral domain. Exploratory analyses on task-questionnaire relationship found that moral self-image score (how closely one feel they are to the ideal moral image of themselves) is positively correlated to the  $d'$  of morally positive condition in singal detection and the drift rate using DDM, while the self-esteem is negatively correlated with  $d'$  of neutral and morally negative conditions. These results suggest that the positive self prioritization in perceptual decision-making may reflect ...

*Keywords:* Perceptual decision-making, Self, positive bias, morality

Word count: X

Positivity bias in perceptual matching may reflect a spontaneous self-referential processing

## Introduction

XXXX In perceptual matching, same is faster than different [Krueger\_1978; Farrell\_1985]. Automatic processing [Spruyt\_de\_Houwer\_2017]

Trisha Van Zandt, Hans Colonius, Robert W. Proctor: A comparison of two response time models applied to perceptual matching

Yakushijin, Reiko.Jacobs, Robert A (2020), Are People Successful at Learning Sequential Decisions on a Perceptual Matching Task?

Schooler, L. J., Shiffrin, R. M., & Raaijmakers, J. G. W. (2001). A Bayesian model for implicit effects in perceptual identification. *Psychological Review*, 108(1), 257–272. <https://doi.org/10.1037/0033-295X.108.1.257>

## General Methods

### Participants.

Most experiments (1a ~ 6b, except experiment 3b) reported in the current study were first finished between 2014 to 2016 in Tsinghua University, Beijing, China. Participants in these experiments were recruited in the local community. To increase the sample size of experiments to 50 or more (Simmons, Nelson, & Simonsohn, 2013), we recruited additional participants in Wenzhou University, Wenzhou, China in 2017 for experiment 1a, 1b, 4a, and 4b. Experiment 3b was finished in Wenzhou University in 2017. To have a better estimation of the effect size, we included the data from two experiments (experiment 7a, 7b) that were reported in Hu, Lan, Macrae, and Sui (2020) (See Table 1 for overview of these experiments). All participant received informed consent and compensated for their time. These experiments were approved by the ethic board in the Department of Tsinghua University.

## Design and Procedure

This series of experiments started to test the effect of instantly acquired moral valence on perceptual decision-making. For this purpose, we used the social associative learning paradigm (or tagging paradigm)(Sui, He, & Humphreys, 2012), in which participants first learned the associations between geometric shapes and labels of person with different moral valence (e.g., in first three studies, the triangle, square, and circle and good person, neutral person, and bad person, respectively). The associations of the shapes and label were counterbalanced across participants. After remembered the associations, participants finished a practice phase to familiar with the task, in which they viewed one of the shapes upon the fixation while one of the labels below the fixation and judged whether the shape and the label matched the association they learned. When participants reached 60% or higher accuracy at the end of the practicing session, they started the experimental task which was the same as in the practice phase.

The experiment 1a, 1b, 1c, 2, and 6a shared a 2 (matchness: matched vs. mismatched) by 3 (moral valence: good vs. neutral vs. bad) within-subject design. Experiment 1a was the first one of the whole series studies and 1b, 1c, and 2 were conducted to exclude the potential confounding factors. More specifically, experiment 1b used different Chinese words as label to test whether the effect only occurred with certain familiar words. Experiment 1c manipulated the moral valence indirectly: participants first learned to associate different moral behaviors with different neutral names, after remembered the association, they then performed the perceptual matching task by associating names with different shapes. Experiment 2 further tested whether the way we presented the stimuli influence the effect of valence, by sequentially presenting labels and shapes. Note that part of participants of experiment 2 were from experiment 1a because we originally planned a cross task comparison. Experiment 6a, which shared the same design as experiment 2, was an EEG experiment which aimed at exploring the neural correlates of

the effect. But we will focus on the behavioral results of experiment 6a in the current manuscript.

For experiment 3a, 3b, 4a, 4b, 6b, 7a, and 7b, we included self-relevance as another within-subject variable in the experimental design. For example, the experiment 3a directly extend the design of experiment 1a into a 2 (matchness: matched vs. mismatched) by 2 (reference: self vs. other) by 3 (moral valence: good vs. neutral vs. bad) within-subject design. Thus in experiment 3a, there were six conditions (good-self, neutral-self, bad-self, good-other, neutral-other, and bad-other) and six shapes (triangle, square, circle, diamond, pentagon, and trapezoids). The experiment 6b was an EEG experiment extended from experiment 3a but presented the label and shape sequentially. Because of the relatively high working memory load (six label-shape pairs), experiment 6b were conducted in two days: the first day participants finished perceptual matching task as a practice, and the second day, they finished the task again while the EEG signals were recorded. Experiment 3b was designed to separate the self-referential trials and other-referential trials. That is, participants finished two different blocks: in the self-referential blocks, they only response to good-self, neutral-self, and bad-self, with half of the trials was matched and half was not; for the other-reference blocks, they only responded to good-other, neutral-other, and bad-other. Experiment 7a and 7b were designed to test the cross task robustness of the effect we observed in the aforementioned experiments (Hu et al., 2020), and the matching task in these two experiments shared the same design with experiment 3a, only with few conditions for moral valence. We only used two conditions of moral valence, i.e., good vs. bad, in experiment 7a and 7b because we found that the neutral and bad conditions constantly show non-significant results.

Experiment 4a and 4b were design to test the automaticity of the binding between self/other and moral valence. In 4a, we used only two labels (self vs. other) and two shapes (circle, square). To manipulate the moral valence, we added labels within the shape and instructed participants to ignore the words in the shape during the task. In 4b, we reversed

the role of self-relevance and valence in the task: participant learnt three labels (good-person, neutral-person, and bad-person) and three shapes (circle, square, and triangle), and the words for self-relevance, “self” or “other”, were presented in the shapes. As in 4a, participants were told to ignore the words inside the shape during the task.

Finally, experiment 5 was design to test the specificity of the moral valence. We extended experiment 1a with an additional independent variable: domains of the valence words. More specifically, besides the moral valence, we also added valence from other domains: appearance of person (beautiful, neutral, ugly), appearance of a scene (beautiful, neutral, ugly), and emotion (happy, neutral, and sad). Label-shape pairs from different domains were separated into different blocks.

E-prime 2.0 was used for presenting stimuli and collecting behavioral responses, except that experiment 7a and 7b used Matlab Psychtoolbox (Brainard, 1997; Pelli, 1997). For participants recruited in Tsinghua University, they finished the experiment individually in a dim-lighted chamber, stimuli were presented on 22-inch CRT monitors and their head were fixed by a chin-rest brace. The distance between participants’ eyes and the screen was about 60 cm. The visual angle of geometric shapes was about  $3.7^\circ \times 3.7^\circ$ , the fixation cross is of ( $0.8^\circ \times 0.8^\circ$  of visual angle) at the center of the screen. The words were of  $3.6^\circ \times 1.6^\circ$  visual angle. The distance between the center of the shape or the word and the fixation cross was  $3.5^\circ$  of visual angle. For participants recruited in Wenzhou University, they finished the experiment in a group consisted of 3 ~ 12 participants in a dim-lighted testing room. Participants were required to finished the whole experiment independently. Also, they were instructed to start the experiment at the same time, so that the distraction between participants were minimized. The stimuli were presented on 19-inch CRT monitor. The visual angles are could not be exactly controlled because participants’s chin were not fixed.

In most of these experiments, participant were also asked to fill a battery of

questionnaire after they finish the behavioral tasks. All the questionnaire data are open (see, dataset 4 in Liu et al., 2020). See Table 1 for a summary information about all the experiments reported here.

## Data analysis

We reported all the measurements, analyses, and results in all the experiments in the current study. Participants whose overall accuracy lower than 60% were excluded from analysis. Also, the accurate responses with less than 200ms reaction times were excluded from the analysis. To have a better overview of the effect reported in this series experiment, we reported the synthesized results in the main text and individual experiment in supplementary materials.

**Analysis of individual study.** The individual experiment’s results were reported in supplementary materials. We used the `tidyverse` of `r` (see script `Load_save_data.r`) to exclude the practicing trials, invalid trials of each participants, and invalid participants, if there were any, in the raw data.

Results of each experiment were analyzed as in Sui et al. (2012). That is, the accuracy performance using a signal detection approach, in which the performance in each match condition was combined with that in the nonmatching condition with the same shape to form a measure of  $d'$ . Trials without response were coded either as “miss” (matched trials) or “false alarm” (mismatched trials). For the reaction times (RTs), only RTs of accurate trials were analyzed.

Both signal detection theory analysis of accuracy and RTs were analyzed in Frequentists’ approach and Bayesian approach. In the Frequentists’ approach, we calculated the  $d'$  using Maximum Likelihood approach and then subjected the  $d'$  estimated from each participant to repeated measures analyses of variance (repeated measures ANOVA), via `afex` []; we used the mean RTs of each participant in each condition and

subject these mean value to repeated measures ANOVA too. We reported the results from significance test and effect sizes (including 95% confidence intervals). To control the false positive rate when conducting the post-hoc comparisons, we used Bonferroni correction.

In the Bayesian approach, we used `brms` (Bürkner, 2017, @Bürkner 2017; Carpenter et al., 2017) to implement the Bayesian hierarchical generalized linear model to estimate the effect of valence and self-referential. See supplementary materials for the results of each experiment's method and results.

Finally, we also explored the psychological processes during the perceptual decision-making using the drift-diffusion model (DDM). We used HDDM (Wiecki, Sofer, & Frank, 2013) for this purpose.

**Synthesized results.** We reported the synthesized results from the experiments, because many of them shared the similar experimental design. We reported the results in five parts: valence effect, explicit interaction between valence and self-relevance, implicit interaction between valence and self-relevance, specificity of valence effect, and behavior-questionnaire correlation.

For the first two parts, we reported the synthesized results from Frequentist's approach(mini-meta-analysis, Goh, Hall, & Rosenthal, 2016). The mini meta-analyses were carried out by using `metafor` package (Viechtbauer, 2010). We first calculated the mean of  $d'$  and RT of each condition for each participant, then calculate the effect size (Cohen's  $d$ ) and variance of the effect size for all contrast we interested: Good v. Bad, Good v. Neutral, and Bad v. Neutral for the effect of valence, and self vs. other for the effect of self-relevance. Cohen's  $d$  and its variance were estimated using the following formula (Cooper, Hedges, & Valentine, 2009):

$$d = \frac{(M_1 - M_2)}{\sqrt{(sd_1^2 + sd_2^2) - 2rsd_1sd_2}} \sqrt{2(1 - r)}$$



$$var.d = 2(1 - r)\left(\frac{1}{n} + \frac{d^2}{2n}\right)$$

$M_1$  is the mean of the first condition,  $sd_1$  is the standard deviation of the first condition, while  $M_2$  is the mean of the second condition,  $sd_2$  is the standard deviation of the second condition.  $r$  is the correlation coefficient between data from first and second condition.  $n$  is the number of data point (in our case the number of participants included in our research).

The effect size from each experiment were then synthesized by random effect model using `metafor` (Viechtbauer, 2010). Note that to avoid the cases that some participants participated more than one experiments, we inspected the all available information of participants and only included participants' results from their first participation. As mentioned above, 24 participants were intentionally recruited to participate both exp 1a and exp 2, we only included their results from experiment 1a in the meta-analysis.

**Valence effect.** We synthesized effect size of  $d'$  and RT from experiment 1a, 1b, 1c, 2, 5 and 6a for the valence effect. We reported the synthesized the effect across all experiments that tested the valence effect, using the mini meta-analysis approach (Goh et al., 2016).

**Explicit interaction between Valence and self-relevance.** The results from experiment 3a, 3b, 6b, 7a, and 7b. These experiments explicitly included both moral valence and self-reference.

**Implicit interaction between valence and self-relevance.** In the third part, we focused on experiment 4a and 4b, which were designed to examine the implicit effect of the interaction between moral valence and self-referential processing. We are interested in one particular question: will self-referential and morally positive valence had a mutual facilitation effect. That is, when moral valence (experiment 4a) or self-referential (experiment 4a) was presented as task-irrelevant stimuli, whether they would facilitate

self-referential or valence effect on perceptual decision-making. For experiment 4a, we reported the comparisons between different valence conditions under the self-referential task and other-referential task. For experiment 4b, we first calculated the effect of valence for both self- and other-referential conditions and then compared the effect size of these three contrast from self-referential condition and from other-referential condition. Note that the results were also analyzed in a standard repeated measure ANOVA (see supplementary materials).

**Specificity of the valence effect.** In this part, we reported the data from experiment 5, which included positive, neutral, and negative valence from four different domains: morality, aesthetic of person, aesthetic of scene, and emotion. This experiment was design to test whether the positive bias is specific to morality.

**Behavior-Questionnaire correlation.** Finally, we explored correlation between results from behavioral results and self-reported measures. For the behavioral task part, we derived different indices. First, we used the mean of the RT and  $d'$  from each participants of each condition. Second, we used three parameters from drift diffusion model. The mean value of each parameter's posterior were used. The DDM analyses were finished by HDDM(Wiecki et al., 2013), as reported in Hu et al. (2020). That is, we used the response code approach, matched response were coded as 1 and mismatched responses were coded as 0. To fully explore all parameters, we allow all four parameters of DDM free to vary. We then extracted the estimation of all the four parameters for each participants for the correlation analyses.

For the questionnaire part, we are most interested in the self-rated distance between different person and self-evaluation related questionnaires: self-esteem, moral-self identity, and moral self-image. Other questionnaires (e.g., personality) were not planned to correlated with behavioral data were not included. Note that all data were reported in (Liu et al., 2020).

For the behavioral task part, we derived different indices. First, we used the mean of the RT data from each participants of each condition. Second, we used three parameters of drift diffusion model of each participant as estimated by HDDM. Third, we calculated the differences between different conditions (valence effect: good-self vs. bad-self, good-self vs. neutral-self, bad-self vs. neutral-self; good-other vs. bad-other, good-other vs. neutral-other, bad-other vs. neutral-other; Self-reference effect: good-self vs. good-other, neutral-self vs. neutral-other, bad-self vs. bad-other), as indexed by Cohen's  $d$  and standard error (SE) of Cohen's  $d$ .

$$Cohen's d_z = \frac{(M_1 - M_2)}{\sqrt{(SD_1^2 + SD_2^2)/2}}$$

Given that the task difficulty were different across experiments, we z-transformed all these indices so that they become unit-free.

The DDM analyses were finished by HDDM, as reported in Hu et al. (2020). That is, we used the response code approach, matched response were coded as 1 and mismatched responses were coded as 0. To fully explore all parameters, we allow all four parameters of DDM free to vary. We then extracted the estimation of all the four parameters for each participants for the correlation analyses. However, because the starting point is only related to response (match vs. non-match) but not the valence of the stimuli, we didn't included it in correlation analysis.

For the questionnaire part, we are interested in those questionnaires related to morality: self-rated distance between different person, self-esteem, moral-self identity, and moral self-image. Other questionnaires (e.g., personality) were not planned to correlated with behavioral data were not included.

We used Pearson correlation analysis to quantify the correlation. For those correlation that is significant ( $p < 0.05$ ), we further tested the robustness of the correlation using bootstrap.

## Results

### Effect of moral valence

In this part, we synthesized results from experiment 1a, 1b, 1c, 2, 5 and 6a. Data from 192 participants were included in these analyses. We found differences between positive and negative conditions on RT was Cohen's  $d = -0.58 \pm 0.06$ , 95% CI [-0.70 -0.47]; on  $d'$  was Cohen's  $d = 0.24 \pm 0.05$ , 95% CI [0.15 0.34]. The effect was also observed between positive and neutral condition, RT: Cohen's  $d = -0.44 \pm 0.10$ , 95% CI [-0.63 -0.25];  $d'$ : Cohen's  $d = 0.31 \pm 0.07$ , 95% CI [0.16 0.45]. And the difference between neutral and bad conditions are not significant, RT: Cohen's  $d = 0.15 \pm 0.07$ , 95% CI [0.00 0.30];  $d'$ : Cohen's  $d = 0.07 \pm 0.07$ , 95% CI [-0.08 0.21]. See Figure 1 left panel.

### Interaction between valence and self-reference

In this part, we combined the experiments that explicitly manipulated the self-reference and valence, which includes 3a, 3b, 6b, 7a, and 7b. For the positive versus negative contrast, data were from five experiments with 178 participants; for positive versus neutral and neutral versus negative contrasts, data were from three experiments (3a, 3b, and 6b) with 108 participants.

In most of these experiments, the interaction between self-reference and valence was significant (see results of each experiment in supplementary materials). In the mini-meta-analysis, we analyzed the valence effect for self-referential condition and other-referential condition separately.

For the self-referential condition, we found the same pattern as in the first part of results. That is we found significant differences between positive and neutral as well as positive and negative, but not neutral and negative. The effect size of RT between positive and negative is Cohen's  $d = -0.89 \pm 0.12$ , 95% CI [-1.11 -0.66]; on  $d'$  was Cohen's  $d = 0.61$

$\pm 0.09$ , 95% CI [0.44 0.78]. The effect was also observed between positive and neutral condition, RT: Cohen's  $d = -0.76 \pm 0.13$ , 95% CI [-1.01 -0.50];  $d'$ : Cohen's  $d = 0.69 \pm 0.14$ , 95% CI [0.42 0.96]. And the difference between neutral and bad conditions are not significant, RT: Cohen's  $d = 0.03 \pm 0.13$ , 95% CI [-0.22 0.29];  $d'$ : Cohen's  $d = 0.08 \pm 0.08$ , 95% CI [-0.07 0.24]. See Figure 1 the middle panel.

For the other-referential condition, we found that only the difference between positive and negative on RT was significant, all the other conditions were not. The effect size of RT between positive and negative is Cohen's  $d = -0.28 \pm 0.05$ , 95% CI [-0.38 -0.17]; on  $d'$  was Cohen's  $d = -0.02 \pm 0.08$ , 95% CI [-0.17 0.13]. The effect was not observed between positive and neutral condition, RT: Cohen's  $d = -0.12 \pm 0.10$ , 95% CI [-0.31 0.06];  $d'$ : Cohen's  $d = 0.01 \pm 0.08$ , 95% CI [-0.16 0.17]. And the difference between neutral and bad conditions are not significant, RT: Cohen's  $d = 0.14 \pm 0.09$ , 95% CI [-0.03 0.31];  $d'$ : Cohen's  $d = 0.05 \pm 0.07$ , 95% CI [-0.08 0.18]. See Figure 1 right panel.

### Generalizability of the valence effect

In this part, we reported the results from experiment 4 in which either moral valence or self-reference were manipulated as task-irrelevant stimuli.

For experiment 4a, when self-reference was the target and moral valence was task-irrelevant, we found that only under the implicit self-referential condition, i.e., when the moral words were presented as task irrelevant stimuli, there was the main effect of valence and interaction between valence and reference for both  $d$  prime and RT (See supplementary results for the detailed statistics). For  $d$  prime, we found good-self condition ( $2.55 \pm 0.86$ ) had higher  $d$  prime than bad-self condition ( $2.38 \pm 0.80$ ); good self condition was also higher than neutral self ( $2.45 \pm 0.78$ ) but there was not statistically significant, while the neutral-self condition was higher than bad self condition and not significant neither. For reaction times, good-self condition ( $654.26 \pm 67.09$ ) were faster

relative to bad-self condition ( $665.64 \pm 64.59$ ), and over neutral-self condition ( $664.26 \pm 64.71$ ). The difference between neutral-self and bad-self conditions were not significant. However, for the other-referential condition, there was no significant differences between different valence conditions. See Figure 2.

For experiment 4b, when valence was the target and the identity was task-irrelevant, we found a strong valence effect (see supplementary results and Figure 3, Figure 4).

In this experiment, the advantage of good-self condition can only be disentangled by comparing the self-referential and other-referential conditions. Therefore, we calculated the differences between the valence effect under self-referential and other referential conditions and used the weighted variance as the variance of this differences. We found this modulation effect on RT. The valence effect of RT was stronger in self-referential than other-referential for the Good vs. Neutral condition ( $-0.33 \pm 0.01$ ), and to a less extent the Good vs. Bad condition ( $-0.17 \pm 0.01$ ). While the size of the other effect's CI included zero, suggestion those effects didn't differ from zero. See Figure 5.

### Specificity of valence effect

In this part, we analyzed the results from experiment 5, which included positive, neutral, and negative valence from four different domains: morality, emotion, aesthetics of human, and aesthetics of scene. We found interaction between valence and domain for both  $d$  prime and RT (matched trials). A common pattern appeared in all four domains: each domain showed a binary results instead of gradient on both  $d$  prime and RT. For morality, aesthetics of human, and aesthetics of scene, the positive conditions had advantages over both neutral and negative conditions (greater  $d$  prime and faster RT), and neutral and negative conditions didn't differ from each other. But for the emotional stimuli, it was the positive and neutral had advantage over negative conditions, while positive and neutral conditions were not significantly different. See supplementary materials for detailed

statistics. Also note that the effect size in moral domain is smaller than the aesthetic domains (beauty of people and beauty of scene). See Figure 6.

### Correlation analyses

The reliability of questionnaires can be found in (Liu et al., 2020). Then we calculated the correlation between the data from behavioral task and the questionnaire data.

We focused on the task-questionnaire correlation, the results revealed that the score from three questionnaire are related to behavioral responses data. First, the external moral identity is positively correlated with boundary separation of moral good condition,  $r = 0.194$ , 95% CI [0.023 0.350]; the moral self image is positively correlated with the drift rate ( $r = 0.191$ , 95% CI [-0.016 0.354]) of the morally good condition. See the left and middle part of Figure 7.

Second, we found the personal distance between self and good is positively correlated with the boundary separation of neutral condition and the self-neutral distance is negatively correlated with the boundary separation of neutral condition.

Third, we found the self esteem score was negative correlated with the  $d'$  of bad conditions ( $r = -0.16$ , 95% CI [-0.277 -0.038]) and the neutral conditions ( $r = -0.197$ , 95% CI [-0.348 -0.026]). See Figure 8.

We also explored the correlation between behavioral data and questionnaire scores separately for experiments with and without self-referential. For experiments without self-referential (Valence effect), we found the personal distance between Good-person and self is positively correlated with boundary separation of good conditions,  $r = 0.292$ , 95% CI [0.071 0.485]. also personal distance between the bad and neutral person is positively correlated with non-responding time of bad and neutral conditions,  $r = 0.249$ , 0.233, respectively.

For experiments with self-referential (Valence effect for the self), we found self-esteem

is negatively correlated with d prime of neutral condition,  $r = -0.272$ ,  $[-0.468 -0.052]$ , the self-good distance is positively correlated with d prime for Bad condition,  $r = 0.185$ , 95%CI[0.004 0.354].

## Discussion

## References

- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. Journal Article.
- Bürkner, P.-C. (2017). Brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software; Vol 1, Issue 1 (2017)*. Journal Article. Retrieved from <https://www.jstatsoft.org/v080/i01%0Ahttp://dx.doi.org/10.18637/jss.v080.i01>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1). Journal Article. <https://doi.org/10.18637/jss.v076.i01>
- Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). Book, New York: Sage.
- Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass*, 10(10), 535–549. Journal Article. <https://doi.org/10.1111/spc3.12267>
- Hu, C.-P., Lan, Y., Macrae, C. N., & Sui, J. (2020). Good me bad me: Does valence influence self-prioritization during perceptual decision-making? *Collabra: Psychology*, 6(1), 20. Journal Article. <https://doi.org/10.1525/collabra.301>
- Liu, Q., Wang, F., Yan, W., Peng, K., Sui, J., & Hu, C.-P. (2020). Questionnaire data from the revision of a chinese version of free will and determinism plus scale. *Journal of*



383 *Open Psychology Data*, 8(1), 1. Journal Article. <https://doi.org/10.5334/jopd.49/>

384 Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming  
385 numbers into movies. *Spatial Vision*, 10(4), 437–442. Journal Article.

386 Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2013). Life after p-hacking. Conference  
387 Proceedings. <https://doi.org/10.2139/ssrn.2205186>

388 Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: Evidence  
389 from self-prioritization effects on perceptual matching. *Journal of Experimental*  
390 *Psychology: Human Perception and Performance*, 38(5), 1105–1117. Journal  
391 Article. <https://doi.org/10.1037/a0029792>

392 Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical bayesian estimation of  
393 the drift-diffusion model in python. *Frontiers in Neuroinformatics*, 7.  
394 <https://doi.org/10.3389/fninf.2013.00014>

Table 1

*Information about all experiments.*

ExpID	Year	Month	N	DV	Design	Self.ref	Valence	Presenting
Exp_1a_1	2014	4	38 (35)	behav	3 * 2	explicit	words	Simultaneously
Exp_1a_2	2017	4	18 (16)	behav	3 * 2	explicit	words	Simultaneously
Exp_1b_1	2014	10	39 (27)	behav	3 * 2	explicit	words	Simultaneously
Exp_1b_2	2017	4	33 (25)	behav	3 * 2	explicit	words	Simultaneously
Exp_1c	2014	10	23 (23)	behav	3 * 2	explicit	descriptions	Simultaneously
Exp_2	2014	5	35 (34)	behav	3 * 2	explicit	words	Sequentially
Exp_3a	2014	11	38 (35)	behav	3 * 2 * 2	explicit	words	Simultaneously
Exp_3b	2017	4	61 (56)	behav	3 * 2 * 2	explicit	words	Simultaneously
Exp_4a_1	2015	6	32 (29)	behav	3 * 2 * 2	implicit	words	Simultaneously
Exp_4a_2	2017	4	32 (30)	behav	3 * 2 * 2	implicit	words	Simultaneously
Exp_4b_1	2015	10	34 (32)	behav	3 * 2 * 2	implicit	words	Simultaneously
Exp_4b_2	2017	4	19 (13)	behav	3 * 2 * 2	implicit	words	Simultaneously
Exp_5	2016	1	43 (38)	behav	3 * 2 * 4	explicit	words	Simultaneously
Exp_6a	2014	12	24 (24)	behav/EEG	3 * 2	explicit	words	Sequentially
Exp_6b	2016	1	23 (22)	behav/EEG	3 * 2 * 2	explicit	words	Sequentially
Exp_7a	2016	7	35 (29)	behav	2 * 2 * 2	explicit	words	Simultaneously
Exp_7b	2018	5	46 (42)	behav	2 * 2 * 2	explicit	words	Simultaneously

*Note.* DV = dependent variables; Valence = how valence was manipulated; Shape & Label = how shapes & labels were presented.

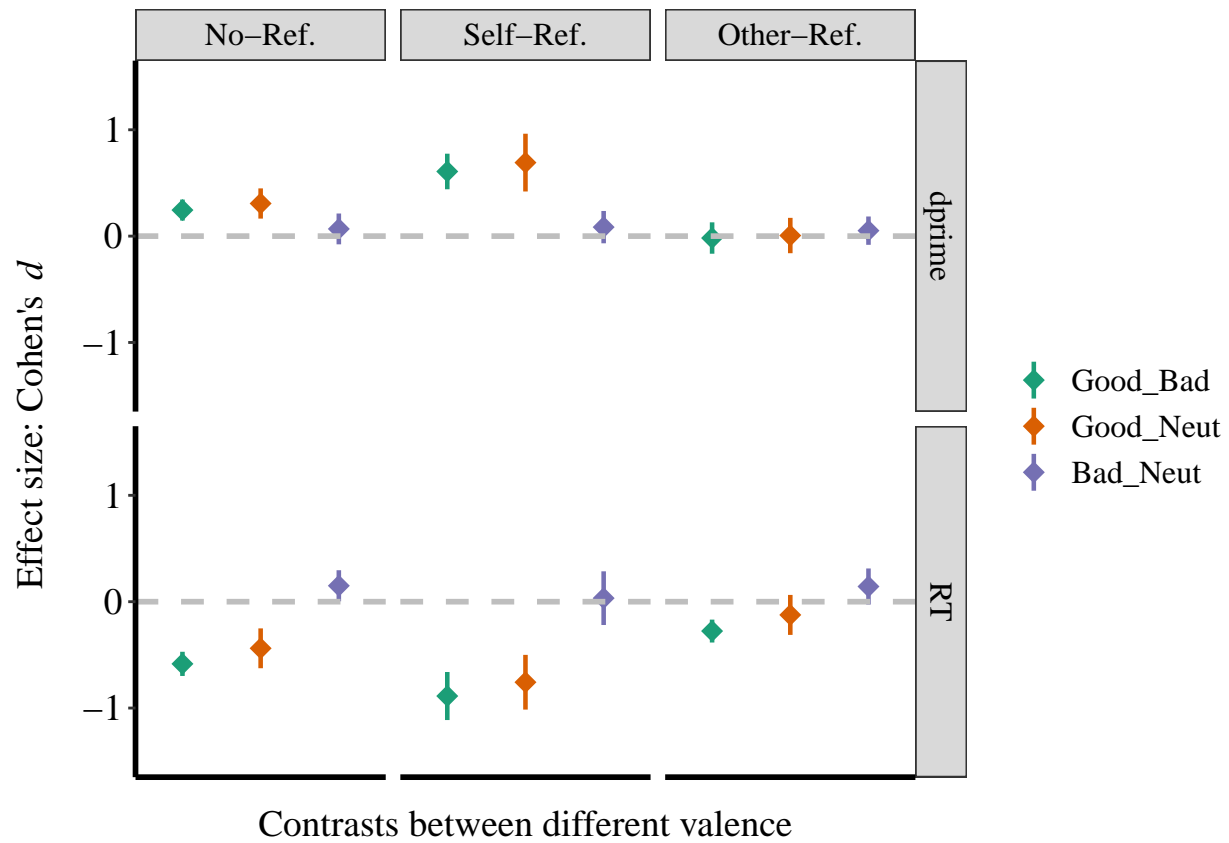


Figure 1. Effect size (Cohen's  $d$ ) of Valence.

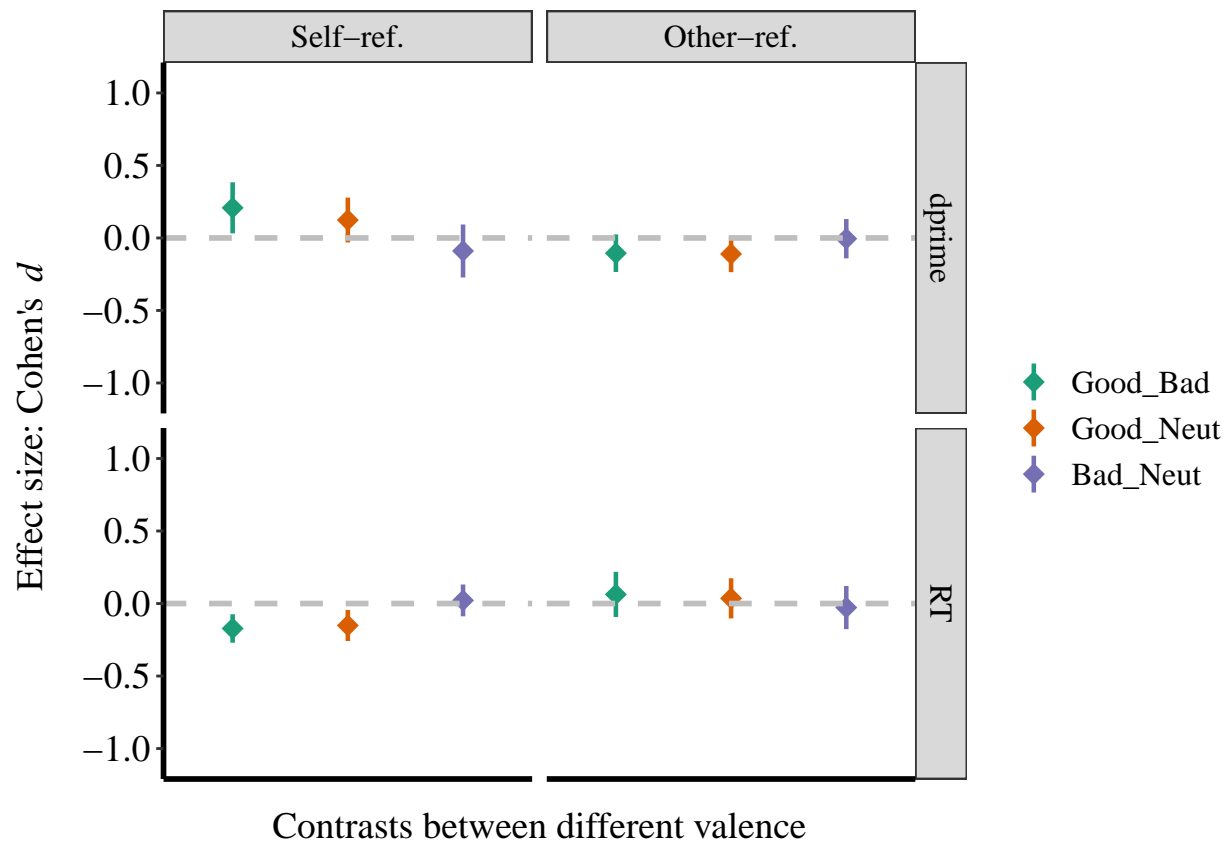


Figure 2. Effect size (Cohen's  $d$ ) of Valence in Exp4a.

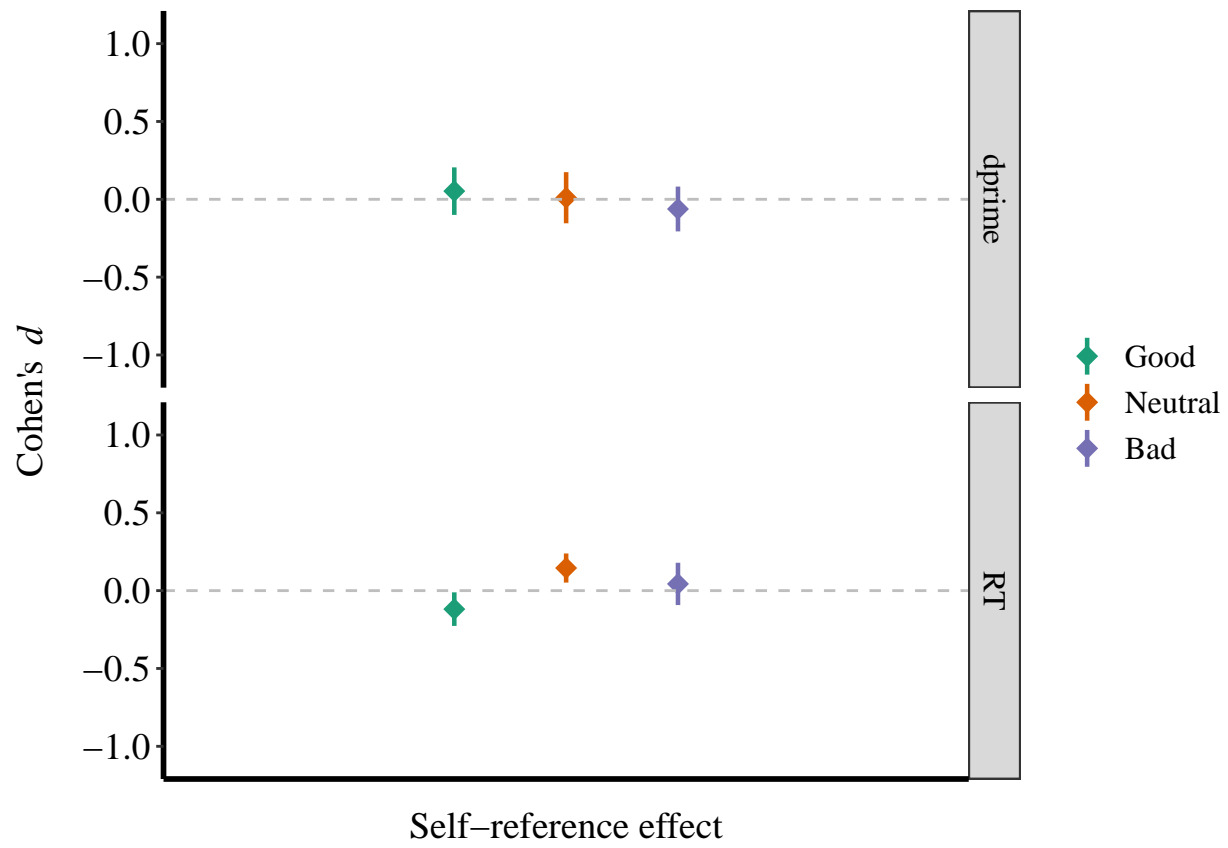


Figure 3. Effect size (Cohen's  $d$ ) of Valence in Exp4b.

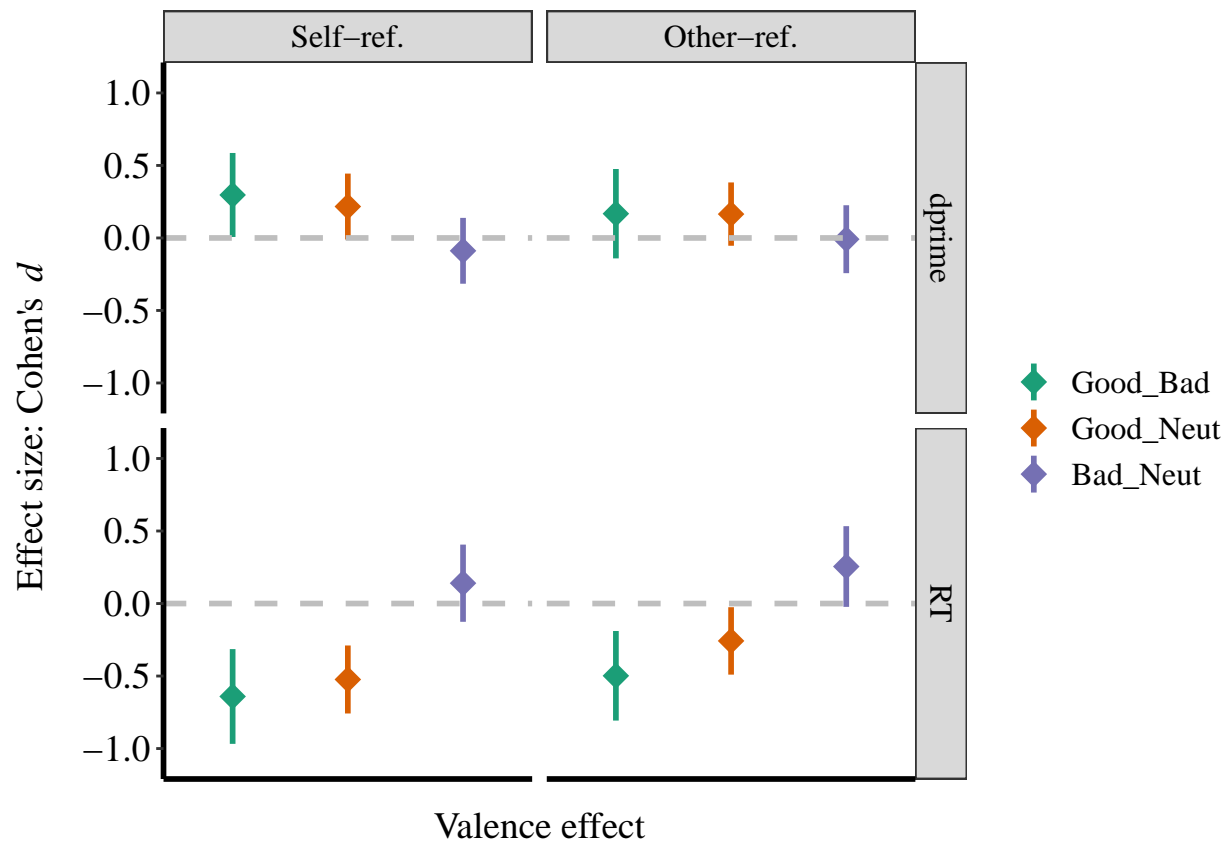


Figure 4. Effect size (Cohen's  $d$ ) of Valence in Exp4b.

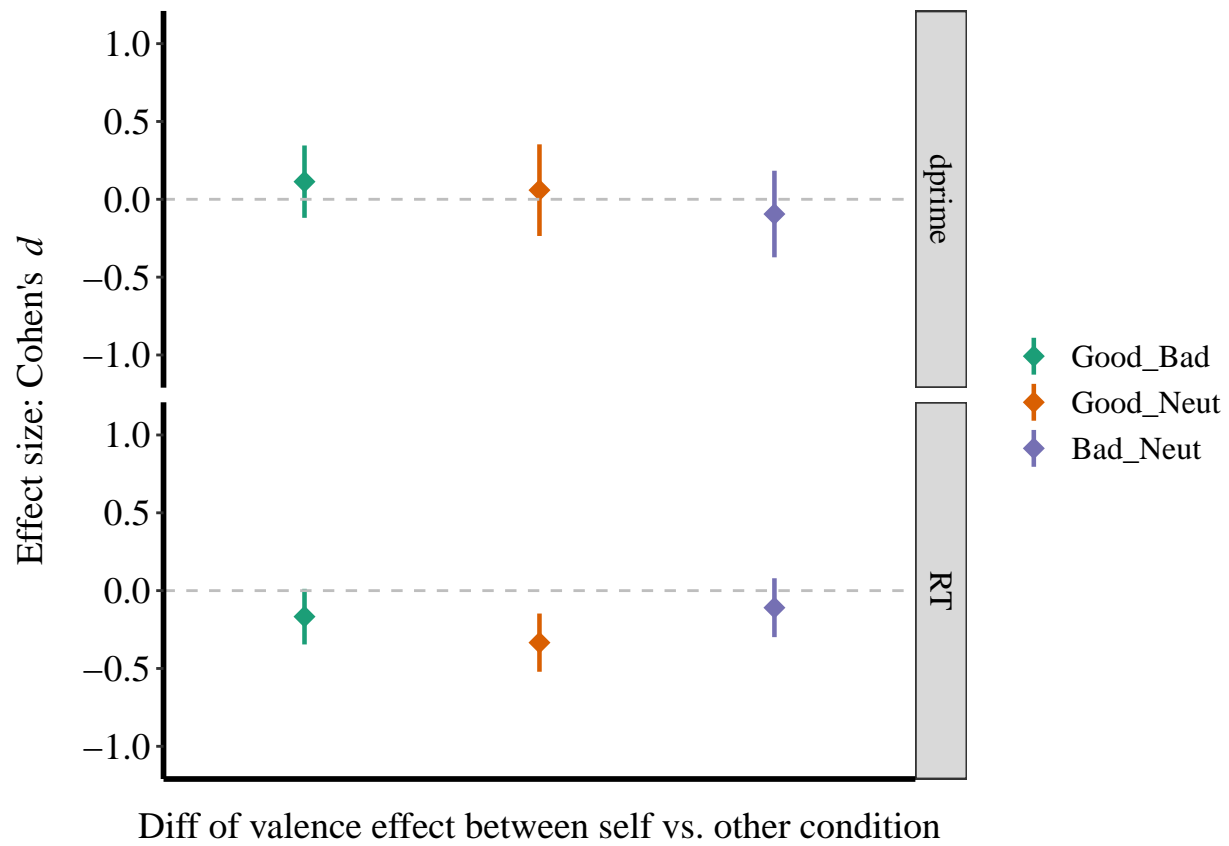


Figure 5. Effect size (Cohen's  $d$ ) of Valence in Exp4b.

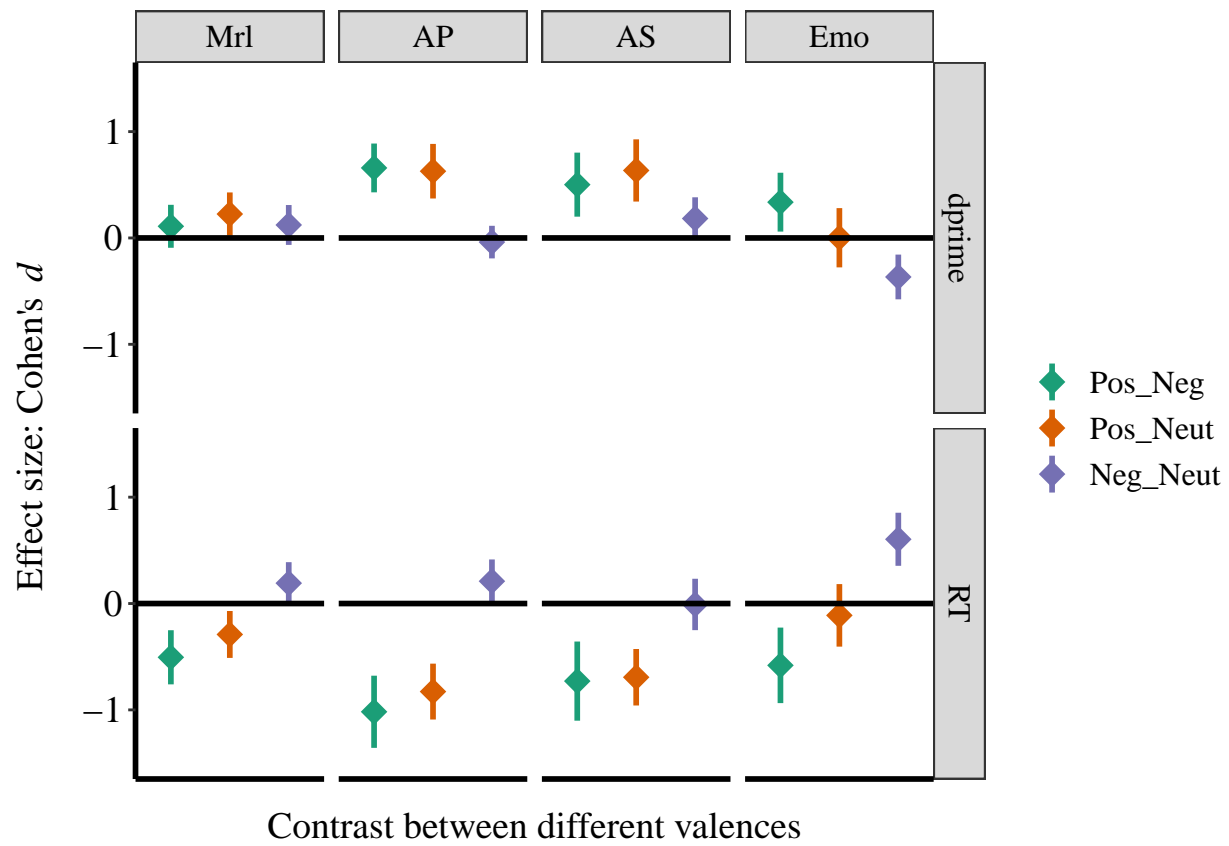


Figure 6. Effect size (Cohen's  $d$ ) of Valence in Exp5.



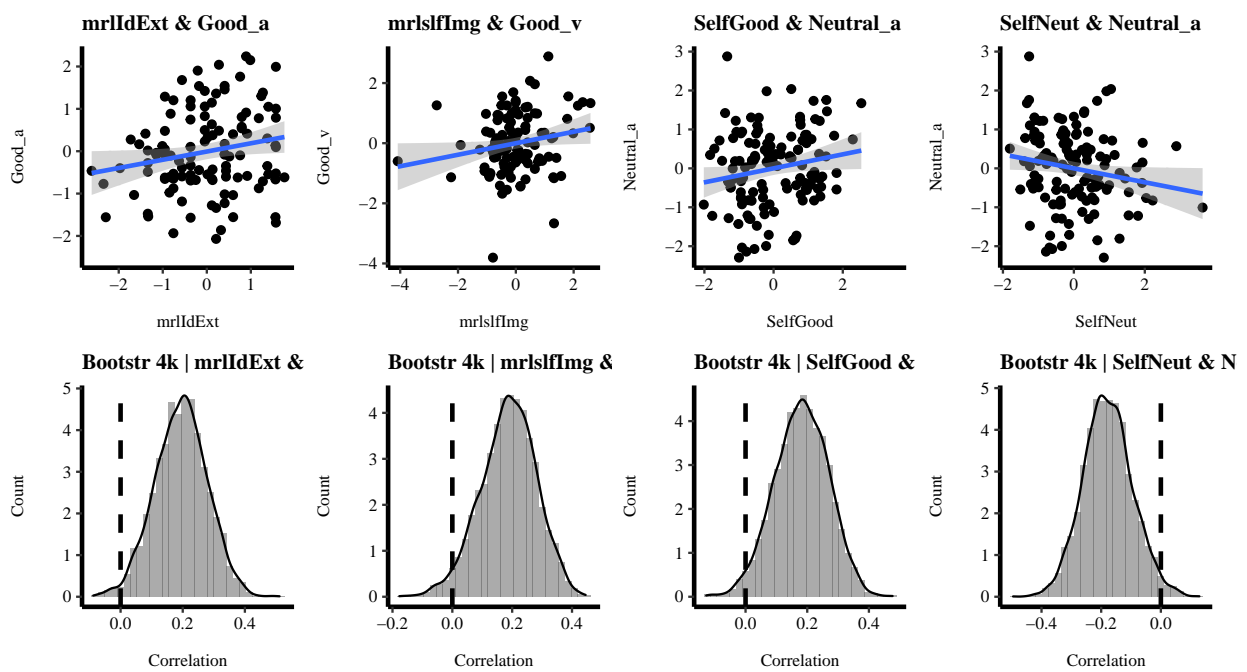


Figure 7. Correlation between moral self-image, self esteem, and behavior

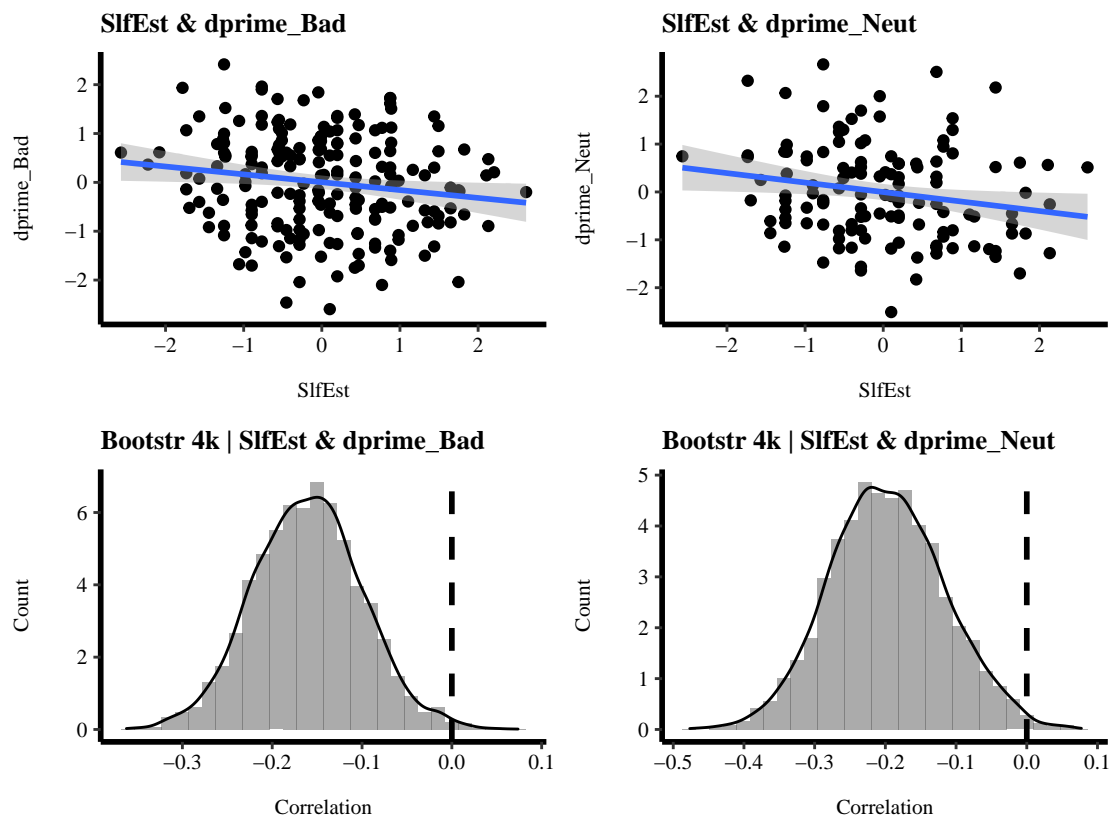


Figure 8. Correlation between moral identity and behavior