Positive bias in perceptual matching may reflect an spontaneous self-referential processing

Hu Chuan-Peng[1,2], Kaiping Peng[1], & Jie Sui[1,3]

[1] Tsinghua University, 100084 Beijing, China

[2] German Resilience Center, 55131 Mainz, Germany

[3] 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 contriubtion: 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 soical 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. Interestingly, the better performance in reaction time is not correpsonding to self-rated psychological distance between self and a morally good-person, but with distance between self and morall bad-person. These results may suggest that our participants (College students in two different cities in China) have a positive moral self bias in perceptual processing, which drive the facilitated processing of morally good stimuli because of the spontaneous self-referential processing, and this trendency is not correlated with explicit rating of moral self.

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

Word count: X

Positive bias in perceptual matching may reflect an spontaneous self-referential processing

Introduction

XXXX

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. Participants of these experiments were recruited in the local community. To increase the sample size so that each experiment has 50 or more valid data (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 (2019) (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 self-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 xperiment 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. The experiment 1a was the first one of the whole series studies and 1b, 1c, and 2 were to exclude other confounding variables' influence. More specifically, experiment 1b used different Chinese words as label to test whether the effect only occure with certain familiar words. Experiment 1c manipulated the moral valence indirectly: participants first learnt to associate different moral behaviors with different names, which is neutral at begining, 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 sequently 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 desing 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 trapezois). The experiment 6b was an EEG experiment extended from experiment 3a but presented the lable 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 reponded to good-other, neutral-other, and bad-other. 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. Experiment 7a and 7b were designed to test the cross task robustness of the effect we observed in the aforementioned experiments (Hu et al., 2019). As we found that the neutral and bad conditions constantly show nonsignificant results, we only used two conditions of moral valence, i.e., good vs. bad, in experiment 7a and 7b.

Finally, experiment 5 was design to test the specificity of the moral valence. We extended experiment 1a with an additional independnet 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), apperance 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° 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, Wang, Yan, Peng, & Hu, 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.

All data were first pre-processed using R (Version 3.6.1; R Core Team, 2018) and the R-packages *afex* (Version 0.25.1; Singmann, Bolker, Westfall, & Aust, 2019), *BayesFactor* (Version 0.9.12.4.2; Morey & Rouder, 2018), *boot* (Version 1.3.23; Davison & Hinkley, 1997;

Gerlanc & Kirby, 2015), *bootES* (Version 1.2; Gerlanc & Kirby, 2015), *coda* (Version 0.19.3; Plummer, Best, Cowles, & Vines, 2006), *corrplot2017* (Wei & Simko, 2017), *dplyr* (Version 0.8.3; Wickham et al., 2019), *emmeans* (Version 1.4.3; Lenth, 2019), *forcats* (Version 0.4.0; Wickham, 2019a), *Formula* (Version 1.2.3; Zeileis & Croissant, 2010), *ggformula* (Version 0.9.2; Kaplan & Pruim, 2019), *ggplot2* (Version 3.2.1; Wickham, 2016), *ggstance* (Version 0.3.3; Henry, Wickham, & Chang, 2018), *ggstatsplot* (Version 0.1.3; Patil & Powell, 2018), *here* (Version 0.1; Müller, 2017), *Hmisc* (Version 4.3.0; Harrell Jr, Charles Dupont, & others., 2019), *lattice* (Version 0.20.38; Sarkar, 2008), *lme4* (Version 1.1.21; Bates, Mächler, Bolker, & Walker, 2015), *lsmeans* (Version 2.30.0; Lenth, 2016), *MASS* (Version 7.3.51.4; Venables & Ripley, 2002), *Matrix* (Version 1.2.17; Bates & Maechler, 2019), *MBESS* (Version 4.6.0; Kelley, 2018), *mosaic* (Version 1.5.0; Pruim, Kaplan, & Horton, 2017, 2018), *mosaicData* (Version 0.17.0; Pruim et al., 2018), *multcomp* (Version 1.4.10; Hothorn, Bretz, & Westfall, 2008), *mvtnorm* (Version 1.0.11; Genz & Bretz, 2009), *papaja* (Version 0.1.0.9842; Aust & Barth, 2018), *plyr* (Version 1.8.4; Wickham et al., 2019; Wickham, 2011), *psych* (Version 1.8.12; Revelle, 2018), *purrr* (Version 0.3.3; Henry & Wickham, 2019), *RColorBrewer* (Version 1.1.2; Neuwirth, 2014), *readr* (Version 1.3.1; Wickham, Hester, & Francois, 2018), *reshape2* (Version 1.4.3; Wickham, 2007), *stringr* (Version 1.4.0; Wickham, 2019b), *survival* (Version 3.1.7; Terry M. Therneau & Patricia M. Grambsch, 2000), *TH.data* (Version 1.0.10; Hothorn, 2019), *tibble* (Version 2.1.3; Müller & Wickham, 2019), *tidyr* (Version 1.0.0; Wickham & Henry, 2019), and *tidyverse* (Version 1.3.0; Wickham, 2017). Individual experiment's results were analyzed as in Sui et al. (2012). we analyzed the accuracy performance using a signal detection approach. 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). The *d'* were then analyzed using repeated measures analyses of variance (repeated measures ANOVA). The reaction times of accurate trials were also analyzed using repeated measures ANOVA. These analyses were based on the pre-processed data and finished by using JASP (0.8.6.0, www.jasp-stats.org, Love et al., 2019). To control the false positive rate when conducting the post-hoc comparisons, we used Bonferroni correction. See supplementary materials for the results of each experiment's method and results, which included the significance test resuts, effect size (Bakeman, 2005; Lakens, 2013), and Bayes factor calculated by JASP (Hu, Kong, Wagenmakers, Ly, & Peng, 2018; Wagenmakers et al., 2018).

To have a better estimation of the effect size, we reported the synthesized results in five parts: valence effect, valence-self-relevance interaction, generalizability, specificity of valence effect, and behavior-questionnaire correlation analysis.

For the first two parts, we synthesized results from individual results by mini-meta-analysis (Goh, Hall, & Rosenthal, 2016). The Mini meta-analyses were carried out in R 3.6, using the "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 Neutral v. Bad 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):

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

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

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. 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 exp 1a in the meta-analysis.

Valence effect. We synthesized effect size of *d* prime 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.

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

Implicit coupling between valence and self-relevance. In the third part, we examined the change of effect size brought by change of design, with a focus on 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 report the comparisons between different valence conditions under the self-referential task, not the other-referential task; for experiment 4b, we reported the comparison between the self- vs. other-referential conditions for positive moral condition, not for the neutral or negative conditions. Note that the results were also analyzed in a standard repeated measure ANOVAs (see supplementary materials).

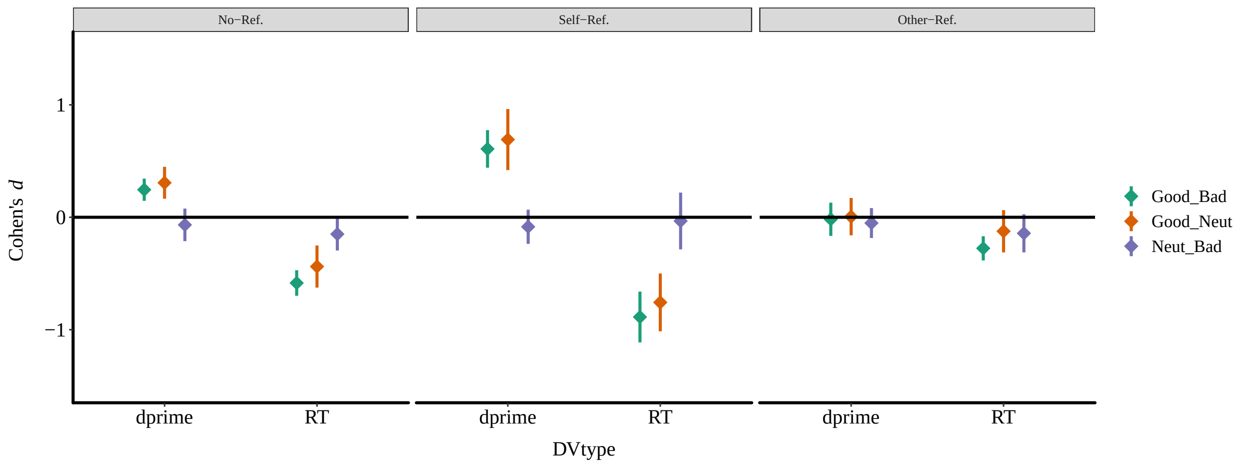
Specificity of the valence effect. In this part, we reported the data from experiment 5, which included positive, neutral, and negative valence from four different domains: morality, aesthetic of person, aesthetic of scence, and emotion. This experiment was design to test whether the positive bias is 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 and SD of the RT data from each participants of each condition. We included the RT variation because it has been shown to be meaningful as individual differences [Jensen, 1992; Ouyang et al., 2017]. Second, we used drift diffusion model to estimate four parameters of DDM for each participants. The DDM analyses were finished by HDDM, as reported in Hu et al., (2019: https://psyarxiv.com/9fczh/). That is, we used the reponse 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 questinnaire 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).

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 analysis. We found differences between positive and negative conditions on RT was Cohen's d = -0.58 ± 0.06, 95% CI [-0.70 -0.47]; on d' was Cohen's d = 0.24 ± 0.05, 95% CI [0.15 0.34]. The effect was also observed between positive and neutral condition, RT: Cohen's d = -0.44 ± 0.10, 95% CI [-0.63 -0.25]; d': Cohen's d = 0.31 ± 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 ± 0.07, 95% CI [-0.30 0.00]; d': Cohen's d = -0.07 ± 0.07, 95% CI [-0.21 0.08]. See Figure **??** 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 whith 178 participants; for positive versus neutral and neutral versus negative contrasts, data were from three experiments with 108 participants.

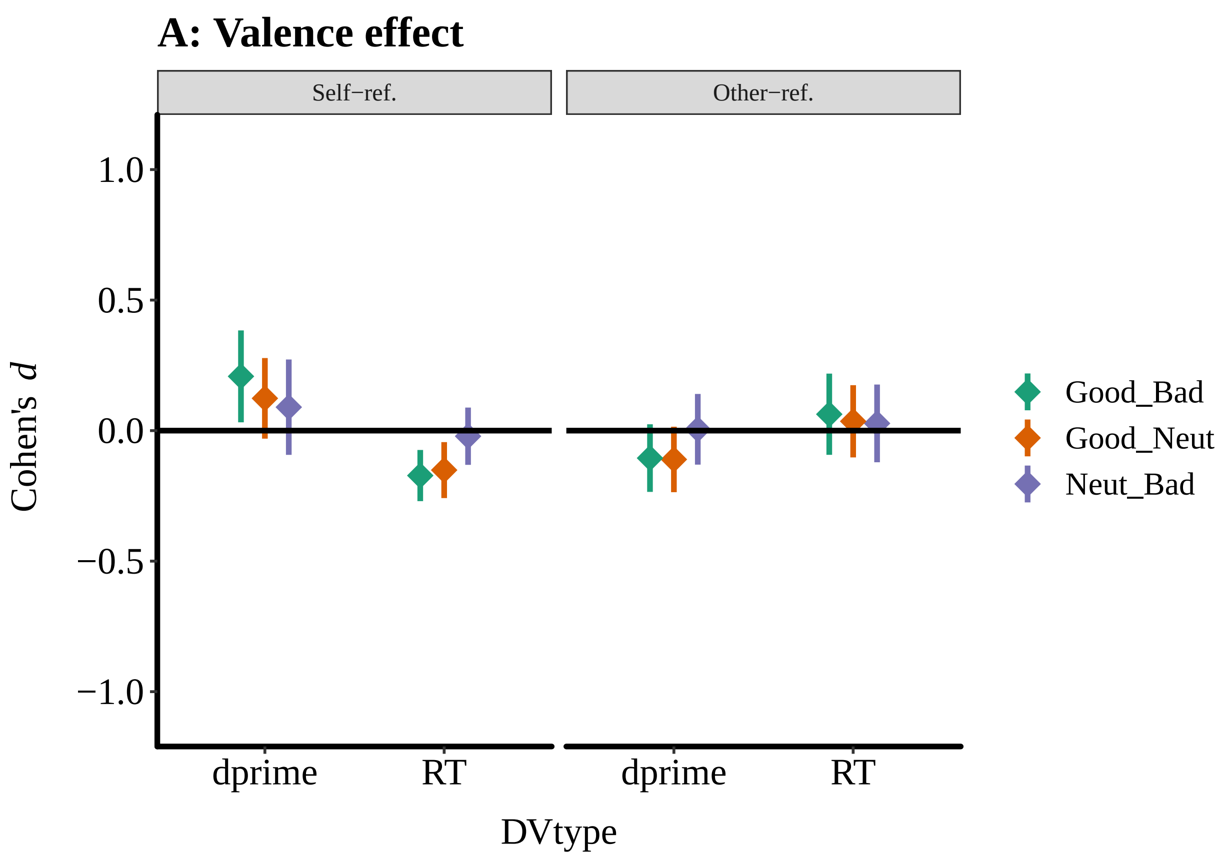
In most of these experiments, the interaction between self-reference and valence was signficant (see results of each experiment in supplementary materials). In the mini-meta-analysis, we analyzed the valence effect for self-referential condition and other-referential condition separately.

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 ± 0.12, 95% CI [-1.11 -0.66]; on *d'* was Cohen's d = 0.61 ± 0.09, 95% CI [0.44 0.78]. The effect was also observed between positive and neutral condition, RT: Cohen's d = -0.76 ± 0.13, 95% CI [-1.01 -0.50]; *d'*: Cohen's d = 0.69 ± 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 ± 0.13, 95% CI [-0.29 0.22]; *d'*: Cohen's d = -0.08 ± 0.08, 95% CI [-0.24 0.07]. See Figure **??** 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 ± 0.05, 95% CI [-0.38 -0.17]; on *d'* was Cohen's d = -0.02 ± 0.08, 95% CI [-0.17 0.13]. The effect was also observed between positive and neutral condition, RT: Cohen's d = -0.12 ± 0.10, 95% CI [-0.31 0.06]; *d'*: Cohen's d = 0.01 ± 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 ± 0.09, 95% CI [-0.31 0.03]; *d'*: Cohen's d = -0.05 ± 0.07, 95% CI [-0.18 0.08]. See Figure **??** right panel.

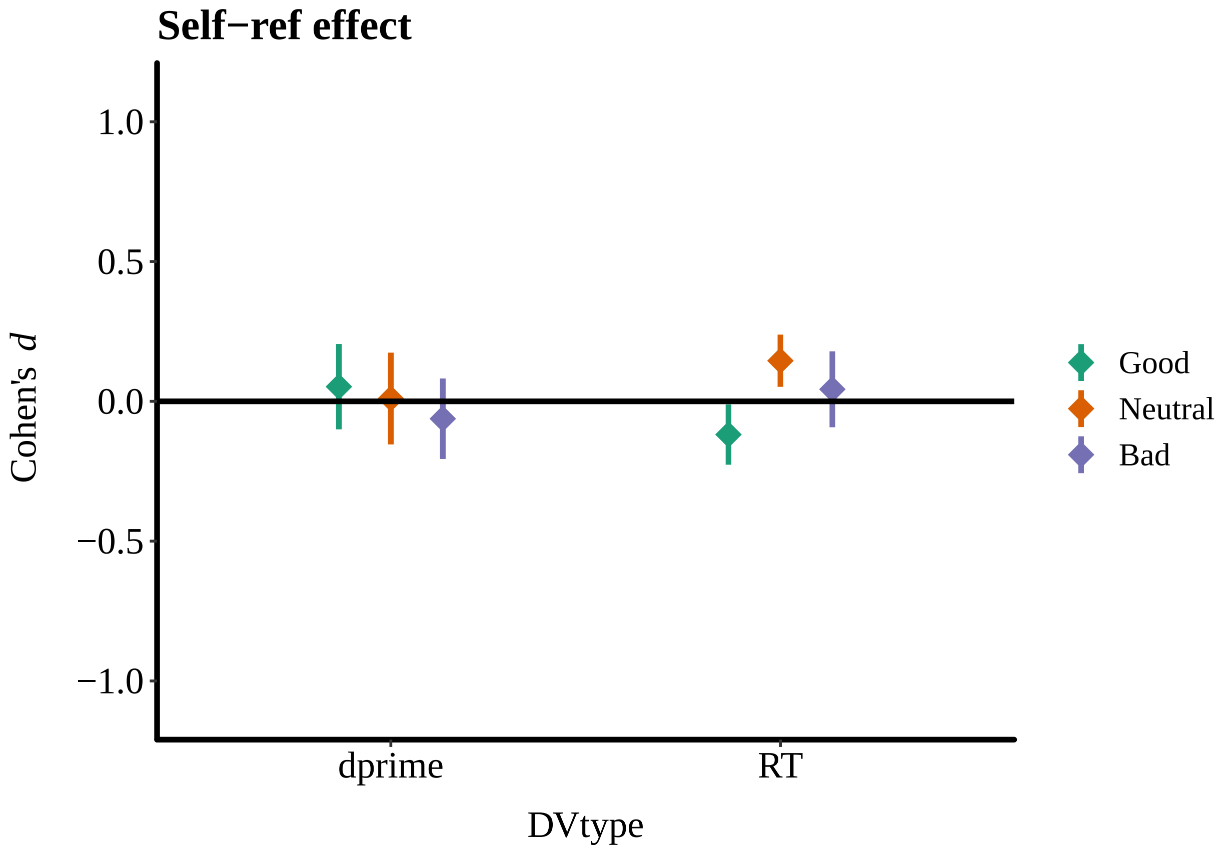
Generalizibility of the valence effect

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 exmperiment 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 resuls for the detailed statistics). For *d* prime, we found good-self condition (2.55 ± 0.86) had higher *d* prime than bad-self condition (2.38 ± 0.80); good self condition was also higher than neutral self (2.45 ± 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 ± 67.09) were faster relative to bad-self condition (665.64 ± 64.59), and over neutral-self condition (664.26 ± 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.

exp4b-1.bb



For experiemnt 4b, when valence was the target and the reference was task-irrelevant, we found a strong valence effect (see supplementary results). In this experiment, the advantage of good-self conition can only be distangled by comparing the self-referential and other-referential conditions while controling the valence condition. We only found this modulation effect on RT. The RT of good-self (680.49 ± 65.69) were faster relative to good-other condition (688.37 ± 66.94), Cohen's d = -0.12, 95% CI[-0.23 -0.01]. However, neutral-self (712.83 ± 54.95) were faster relative to good-other condition (704.64 ± 57.07), Cohen's d = 0.15, 95% CI[0.05 0.24]. The difference between bad-self and bad-other was not significant. All the differences between self-referential and other-referential were not significant for d prime.

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 gradian 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).

Correlation analyses

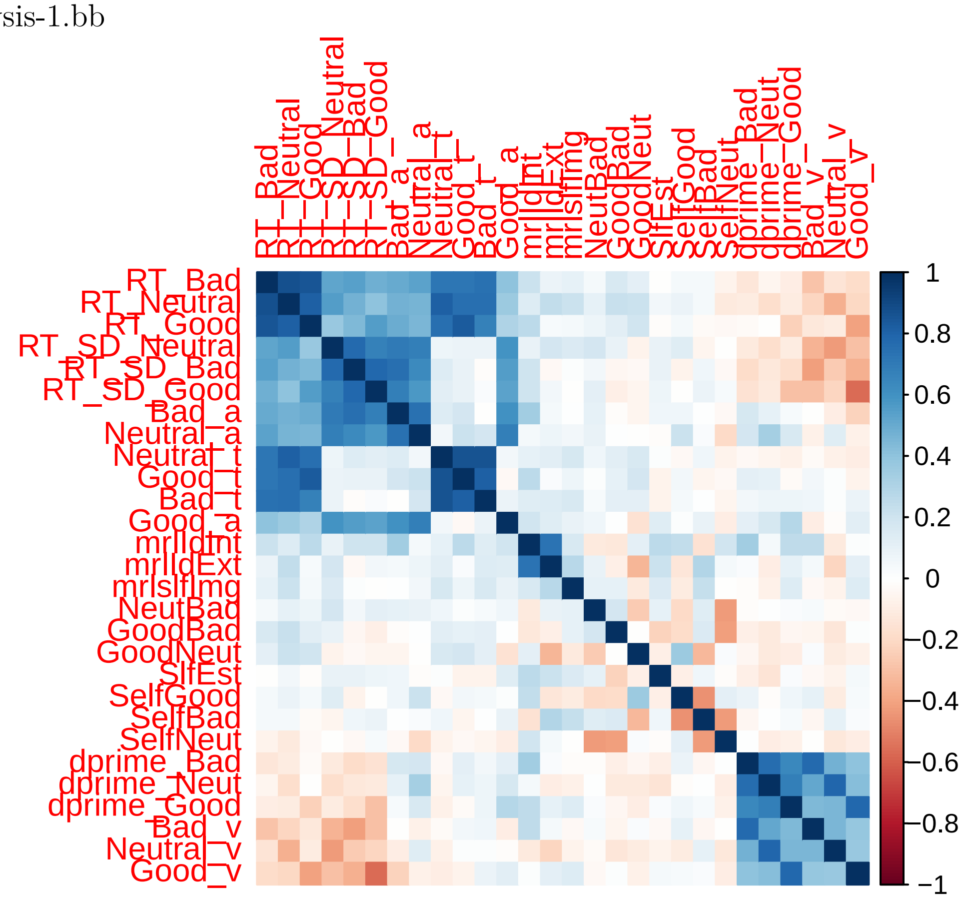
As the reliability of the quesetionnaire can be found in (Liu et al., 2020). Then we calculated the correlation between the data from behavioral task and the questionnaire data.

For the behavioral task part, we derived different indices. First, we used the mean and SD of the RT data from each participants of each condition. We included the RT variation because it has been shown to be meaningful as individual differences [Jensen, 1992; Ouyang et al., 2017]. Second, we used drift diffusion model to estimate four parameters of DDM for each participants. Third, we also 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 se of Cohen's *d*.

The DDM analyses were finished by HDDM, as reported in Hu et al., (2019: https://psyarxiv.com/9fczh/). That is, we used the reponse 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 questinnaire 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.

analysis-1.bb



We found that data from behavioral task are closely related, but not with self-reported questionnaire data.

Discussion

References

Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown.* Retrieved from https://github.com/crsh/papaja

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379–384. Journal Article. https://doi.org/10.3758/BF03192707

Bates, D., & Maechler, M. (2019). *Matrix: Sparse and dense matrix classes and methods.* Retrieved from https://CRAN.R-project.org/package=Matrix

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436. Journal Article.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). Book, New York: Sage.

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their applications.* Cambridge: Cambridge University Press. Retrieved from http://statwww.epfl.ch/davison/BMA/

Genz, A., & Bretz, F. (2009). *Computation of multivariate normal and t probabilities.* Heidelberg: Springer-Verlag.

Gerlanc, D., & Kirby, K. (2015). *BootES: Bootstrap effect sizes.* Retrieved from https://CRAN.R-project.org/package=bootES

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

Harrell Jr, F. E., Charles Dupont, & others. (2019). *Hmisc: Harrell miscellaneous*. Retrieved from https://CRAN.R-project.org/package=Hmisc

Henry, L., & Wickham, H. (2019). *Purrr: Functional programming tools*. Retrieved from https://CRAN.R-project.org/package=purrr

Henry, L., Wickham, H., & Chang, W. (2018). *Ggstance: Horizontal 'ggplot2' components*. Retrieved from https://CRAN.R-project.org/package=ggstance

Hothorn, T. (2019). *TH.data: TH's data archive*. Retrieved from https://CRAN.R-project.org/package=TH.data

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363.

Hu, C.-P., Kong, X.-Z., Wagenmakers, E.-J., Ly, A., & Peng, K. (2018). Bayes factor and its implementation in jasp: A practical primer (in chinese). *Advances in Psychological Science*, *26*(6), 951–965. Journal Article. https://doi.org/10.3724/SP.J.1042.2018.00951

Hu, C.-P., Lan, Y., Macrae, C. N., & Sui, J. (2019). Good me bad me: Does valence influence self-prioritization during perceptual decision-making? *PsyArxiv*. Journal Article. https://doi.org/10.31234/osf.io/9fczh

Kaplan, D., & Pruim, R. (2019). *Ggformula: Formula interface to the grammar of graphics*. Retrieved from https://CRAN.R-project.org/package=ggformula

Kelley, K. (2018). *MBESS: The mbess r package*. Retrieved from https://CRAN.R-project.org/package=MBESS

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and anovas. *Frontiers in Psychology*, *4*, 863. Journal

Article. https://doi.org/10.3389/fpsyg.2013.00863

Lenth, R. (2019). *Emmeans: Estimated marginal means, aka least-squares means.* Retrieved from https://CRAN.R-project.org/package=emmeans

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33. https://doi.org/10.18637/jss.v069.i01

Liu, Q., Wang, F., Yan, W., Peng, K., & Hu, C.-P. (2020). Questionnaire data from the revision of a chinese version of free will and determinism plus scale. *Journal of Open Psychology Data*, *8*(1), 1. Journal Article. https://doi.org/10.5334/jopd.49/

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., … Wagenmakers, E.-J. (2019). JASP: Graphical statistical software for common statistical designs. *Journal of Statistical Software*, *1*(2), 1–17. Journal Article. https://doi.org/10.18637/jss.v088.i02

Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of bayes factors for common designs.* Retrieved from https://CRAN.R-project.org/package=BayesFactor

Müller, K. (2017). *Here: A simpler way to find your files.* Retrieved from https://CRAN.R-project.org/package=here

Müller, K., & Wickham, H. (2019). *Tibble: Simple data frames.* Retrieved from https://CRAN.R-project.org/package=tibble

Neuwirth, E. (2014). *RColorBrewer: ColorBrewer palettes.* Retrieved from https://CRAN.R-project.org/package=RColorBrewer

Patil, I., & Powell, C. (2018). *Ggstatsplot: 'Ggplot2' based plots with statistical details.* https://doi.org/10.5281/zenodo.2074621

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

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for mcmc. *R News*, *6*(1), 7–11. Retrieved from https://journal.r-project.org/archive/

Pruim, R., Kaplan, D., & Horton, N. (2018). *MosaicData: Project mosaic data sets*. Retrieved from https://CRAN.R-project.org/package=mosaicData

Pruim, R., Kaplan, D. T., & Horton, N. J. (2017). The mosaic package: Helping students to 'think with data' using r. *The R Journal*, *9*(1), 77–102. Retrieved from https://journal.r-project.org/archive/2017/RJ-2017-024/index.html

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Revelle, W. (2018). *Psych: Procedures for psychological, psychometric, and personality research*. Evanston, Illinois: Northwestern University. Retrieved from https://CRAN.R-project.org/package=psych

Sarkar, D. (2008). *Lattice: Multivariate data visualization with r*. New York: Springer. Retrieved from http://lmdvr.r-forge.r-project.org

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

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2019). *Afex: Analysis of factorial experiments*. Retrieved from https://CRAN.R-project.org/package=afex

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

Terry M. Therneau, & Patricia M. Grambsch. (2000). *Modeling survival data: Extending* *the Cox model*. New York: Springer.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with s* (Fourth). New York: Springer. Retrieved from http://www.stats.ox.ac.uk/pub/MASS4

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., … Morey, R. D. (2018). Bayesian inference for psychology. Part ii: Example applications with jasp. *Psychonomic Bulletin & Review*, *25*(1), 58–76. Journal Article. https://doi.org/10.3758/s13423-017-1323-7

Wei, T., & Simko, V. (2017). *R package "corrplot": Visualization of a correlation matrix*. Retrieved from https://github.com/taiyun/corrplot

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, *21*(12), 1–20. Retrieved from http://www.jstatsoft.org/v21/i12/

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, *40*(1), 1–29. Retrieved from http://www.jstatsoft.org/v40/i01/

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H. (2017). *Tidyverse: Easily install and load the 'tidyverse'*. Retrieved from https://CRAN.R-project.org/package=tidyverse

Wickham, H. (2019a). *Forcats: Tools for working with categorical variables (factors)*. Retrieved from https://CRAN.R-project.org/package=forcats

Wickham, H. (2019b). *Stringr: Simple, consistent wrappers for common string operations*. Retrieved from https://CRAN.R-project.org/package=stringr

Wickham, H., François, R., Henry, L., & Müller, K. (2019). *Dplyr: A grammar of data manipulation*. Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., & Henry, L. (2019). *Tidyr: Easily tidy data with 'spread()' and 'gather()' functions*. Retrieved from https://CRAN.R-project.org/package=tidyr

Wickham, H., Hester, J., & Francois, R. (2018). *Readr: Read rectangular text data.* Retrieved from https://CRAN.R-project.org/package=readr

Zeileis, A., & Croissant, Y. (2010). Extended model formulas in R: Multiple parts and multiple responses. *Journal of Statistical Software*, *34*(1), 1–13. https://doi.org/10.18637/jss.v034.i01

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.