**Moral motivated perception: The possible role of spontaneous self-referential process**

Chuan-Peng Hu

Department of Psychology, Tsinghua University, 100084 Beijing, China

Neuroimaging Center, University Medical Center Mainz, 55131, Mainz, Germany

Kaiping Peng

Department of Psychology, Tsinghua University, 100084 Beijing, China

Jie Sui

Department of Psychology, Tsinghua University, 100084 Beijing, China

Department of Psychology, University of Bath, Bath, United Kingdom

Author Note: The data reported here is part of C-P Hu’s PhD dissertation, part of the results was reported at the 17th National Academic Congress of Psychology, Oct 2014, Beijing, China. This work was supported by CSFN 31371017 for J Sui and 31471001 for K Peng. We thank John Sakaluk for sharing his code for plotting APA style graph in previous version of this manuscript.

We embrace the values of openness and transparency in science ([Schönbrodt, Maier, Heene, & Zehetleitner, 2015](#_ENREF_24)). Therefore, we followed the 21-word solution provided by [Simmons, Nelson, and Simonsohn (2012)](#_ENREF_26) or referred to the complete project documentations in the OSF. We will furthermore publish all raw data necessary to reproduce the reported results and are providing scripts for all data analyses reported in this manuscript (see the additional materials in the OSF and github).

Corresponding to: Chuan-Peng Hu, [hcp4715@gmail.com](mailto:hcp4715@gmail.com)

## General method

**Participants.** The experiments (except experiment 3b) reported in the current study were first conducted between 2014 to 2016 in Tsinghua University, Beijing, participants of these experiments were recruited in Tsinghua University community. To increase the power by adding collecting more data so that each experiment has 50 or more valid data ([Simmons, Nelson, & Simonsohn, 2013](#_ENREF_27)) , we recruited additional participants in Wenzhou University, Wenzhou, China in 2017. However, duo to the limited time and resources, additional data were not collected for experiment 2, 3, and 4b.

**Stimuli and procedure.** In the current study, we used the social associative learning paradigm ([Sui, He, & Humphreys, 2012](#_ENREF_28)), in which participants first learn 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 learning phase, participants finish 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 were matched. When participants can get 60% or higher accuracy at the end of the practicing session, they can start the experimental task which is the same as in the practice phase.

If not noted, E-prime 2.0 was used in all experiments. For participants recruited in Tsinghua University, they finished the experiment individually in a dim-lighted chamber, stimuli were presented on 22-inch CRT monitors, with a chin-rest brace. The visual angle of geometric shapes was about 3.7º × 3.7º, the finxation cross is of (0.8º × 0.8º of visual angle) at the center of the screen. The words were of 3.6º × 1.6º visual angle. The distance between the center of the shape or the word and the fixation cross was 3.5º of visual angle. Participant fixed their head on a chin-fixation, about 60 cm from the screen.

For participants recruited in Wenzhou University, they finished the experiment in a group consist 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.

**Data analysis.** We reported all the measurements, analysis and results in all the experiments in the current study. All data were first pre-processed using R (3.1.0, R core teams, 2018) and then analyzed using JASP (0.8.6.0, [www.jasp-stats.org](http://www.jasp-stats.org), ([Love et al., 2019](#_ENREF_19))). 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.

We analyzed accuracy performance using a signal detection approach, as in [Sui et al. (2012)](#_ENREF_28). 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. To control the false positive when conducting the post-hoc comparisons, we used Bonferroni correction. Please note that in the first two experiment (experiment 1a and 1b), we included the variable matchness (matched vs. mismatched) in our ANOVA of reaction times and then examine matched trials and mismatched trials separately when the interaction between matchness and other variables are significant. In both experiments, we found significant interaction between matchness and valence. Then, as previous study, we focused on the matched trial for the rest of the experiment ([Sui et al., 2012](#_ENREF_28)).

We reported the effect size of repeated measures ANOVA (*ω²*) ([Bakeman, 2005](#_ENREF_3); [Lakens, 2013](#_ENREF_17)). Also, we reported Cohen’s *d* and its 95% confidence intervals for the post-hoc comparisons. To provide more information about the results, we also reported the Bayes Factor using JASP ([Hu, Kong, Wagenmakers, Ly, & Peng, 2018](#_ENREF_12); [Wagenmakers et al., 2018](#_ENREF_31)). The Bayes factor is the ratio of the probability of the current data pattern under alternative hypothesis (H1) and the probability of the current data pattern under null hypothesis (H0), which index the relative evidence for these two hypotheses from the current data. The BF10 represents the evidence for alternative hypothesis (H1) vs. evidence for null hypothesis (H0); in contrast, BF01 represents that evidence for null hypothesis over the evidence for althernative hypothesis. We used the default prior in JASP for all the Bayes Factor analyses, and used [Jeffreys (1961)](#_ENREF_14)’s convention for the strength of evidence: the BF10 > 3 means there are some evidence for H1 as compared with H0, BF10 ≥ 10 means strong evidence for H1.

To assess the individual difference, we explored correlation between self-reported psychological distance and more objective responses bias (i.e., reaction times and *d* prime). To do this, we first normalized the personal distance by taking the percentage of the mean distance between each two persons in the sum of all 6 distances (self-good, self-normal, self-bad, good-normal, good-bad, normal-bad), and then calculated the bias score (indexed by the differences between good-normal, good-bad). Also, as exploratory analysis, we analyzed the correlation between behavioral response and moral identity, self-esteem, if data are available. As recent study showed that small size leads to unstable correlation estimates ([Schönbrodt & Perugini, 2013](#_ENREF_25)), we only reported the correlation based on data pooled from all experiments, while the results of each experiment were reported in supplementary results.

## Experiment 1a

### Method

**Participants.** 57 college students (38 female, age = 20.75 ± 2.54 years) participated. 39 of them were recruited from Tsinghua University community in 2014; 18 were recruited from Wenzhou University in 2017. All participants were right-handed except one, and all had normal or corrected-to-normal vision. Informed consent was obtained from all participants prior to the experiment according to procedures approved by the local ethics committees. 6 participant’s data were excluded from analysis because nearly random level of accuracy (4 participants from the Tsinghua sample and 2 from the Wenzhou sample), leaving 51 participants (34 female, age = 20.72 ± 2.44 years).

**Stimuli and Tasks.** Three geometric shapes were used in this experiment: triangle, square, and circle. These shapes were paired with three labels (bad person (“坏人”), good person (“好人”) or neutral person (“常人”)). The pairs were counterbalanced across participants.

**Procedure.** As we describe in general method part,this experiment had two phases. First, there was a learning stage. Participants were asked to learn the relationship between geometric shapes (triangle, square, and circle) and different person (bad person, a good person, or a neutral person). For example, a participant was told, “bad person is a circle; good person is a triangle; and a neutral person is represented by a square.” After participant remember the associations (usually in a few minutes), participants started a practicing phase of matching task which has the exact task as in the experimental task.

In the experimental task, participants judged whether shape–label pairs, which were subsequently presented, were correct. Each trial started with the presentation of a central fixation cross for 500 ms. Subsequently, a pairing of a shape and label (good person, bad person, and neutral person) was presented for 100 ms. The pair presented could confirm to the verbal instruction for each pairing given in the training stage, or it could be a recombination of a shape with a different label, with the shape–label pairings being generated at random. The next frame showed a blank for 1100ms. Participants were expected to judge whether the shape was correctly assigned to the person by pressing one of the two response buttons as quickly and accurately as possible within this timeframe (to encourage immediate responding). Feedback (correct or incorrect) was given on the screen for 500 ms at the end of each trial, if no response detected, “too slow” was presented to remind participants to accelerate. Participants were informed of their overall accuracy at the end of each block. The practice phase finished and the experimental task began after the overall performance of accuracy during practice phase achieved 60%.

For pariticpants from the Tsinghua community, they completed 6 experimental blocks of 60 trials. Thus, there were 60 trials in each condition (bad-person matched, bad-person nonmatching, good-person matched, good-person nonmatching, neutral-person matched, and neutral-person nonmatching). For the participants from Wenzhou Univeristy, they finished 6 blocks of 120 trials, therefore, 120 trials for each condition.

**Questionnaires.** After the experiment, part of the participants in Tsinghua University also finished psychological distance, trait social justice ([白福宝, 2013](#_ENREF_35)), cognitive reflection test ([Frederick, 2005](#_ENREF_10)), and disgust senstivity ([谭永红, 丛中, & 鲁晓华, 2007](#_ENREF_36)). The psychological distance measurement finished by indicating the the psychological distance between self, good person, bad person and neutral person, through two points on a horizontal line. This procedure is presented by Matlab. This method had been proven been an effective way to measure the psychological distance ([Enock, Sui, Hewstone, & Humphreys, 2018](#_ENREF_6)).

For all participants from Wenzhou University, they finished following questionnaires online immediately after the experiment: objective and subjective socioeconomic status (the objective SES measured by parents’ education and occupation ([师保国 & 申继亮, 2007](#_ENREF_34)), the subjective SES measured by ladder task ([Ostrove, Adler, Kuppermann, & Washington, 2000](#_ENREF_20))), psychological distance ([Enock et al., 2018](#_ENREF_6)), sensitivity to justice ([Wu et al., 2014](#_ENREF_32)), cognitive reflection test ([Frederick, 2005](#_ENREF_10)), disgust senstivity scale ([谭永红 et al., 2007](#_ENREF_36)), belief in just world (short) ([Wu et al., 2011](#_ENREF_33)), a short version of big five personality ([John & Srivastava, 1999](#_ENREF_15)), trait self-esteem ([Rosenberg, 1965](#_ENREF_23)), locus of control (Levenson，1981), Free will and determinism plus (FAD+) (translated version) ([Liu, Jian, Hu, & Peng, 2015](#_ENREF_18); [Paulhus & Carey, 2010](#_ENREF_21)), moral identity ([Aquino & Reed II, 2002](#_ENREF_2)), and moral self image (translated version) ([Jordan, Leliveld, & Tenbrunsel, 2015](#_ENREF_16)). Only the psychological distance data were analyized in the current study.

**Data analysis.** As we describe in the general method section.

### Results

Table 1 shows the accuracy of paired trials in experiment 1a.

Table 1. Accuracy for each condition (mean ± SD) of experiment 1a.

|  |  |  |
| --- | --- | --- |
| Matchness | Moral valence | Accuracy (%) |
| Match | Good | 86.4 ± 11.3 |
|  | Neutral | 81.6 ± 14.1 |
|  | Bad | 76.4 ± 15.4 |
| Mismatch | Good | 77.2 ± 13.2 |
|  | Neutral | 78.0 ± 13.7 |
|  | Bad | 79.2 ± 13.3 |

A repeated measures ANOVA for *d’* showed a significant effect of moral valence, *F*(2,100) = 6.1859, *p* = 0.00293, *ω²* = 0.091, BF10 = 10.476, suggesting that the data provided strong evidence for the effect of moral valence. Therefore, we conducted post-hoc comparison by paired-wise *t*-test for each pair (with a Bonferroni corrected threshold *p* < 0.0167). The results showed that *d’* was larger for good person (*d’* = 2.11 ± 0.98) than for bad person association (*d*’ = 1.75 ± 1.01), *t*(50) = 3.304, *p* = 0.002, Cohen’s *d*= 0.463, 95% CI[0.172 0.75], BF10 = 17.04, whereas there was no difference between good person and neutral person (*d’* = 1.95 ± 1.11), *t*(50) = 1.54, *p* = 0.13, Cohen’s *dz*= 0.216, 95% CI[-0.063 0.492], BF10 = 0.46; or bad person and neutral association, *t*(50) = -2.109, *p* = 0.04, Cohen’s *dz*= -0.295, 95% CI[0.013 0.574]), BF10 = 1.163 ( see Figure 1A).

A repeated measures ANOVA for the RTs, with moral valence (good, neutral, bad person) and matching judgment (matched vs. nonmatching) as two within-subject factors, showed a significant effect of moral valence, *F*(2,100) = 9.622, *p* = 0.00015, *ω²* = 0.143, BF10 = 14.75, as well as main effect of matching judgment, *F*(1,50) = 232.39 , *p* < 0.0001, *ω²* = 0.817, BF10 = 5.026e + 23. The interaction of moral association and matching judgment was also significant, *F*(2,100) = 8.52, *p* = 0.00038, *ω²* = 0.127, BF10 = 56.4.

We then analyzed the data of the matched and non-matched pairs separately. The matched pairs data showed a significant effect of moral valence, *F*(2,100) = 10.52, *p* = 0.00007, *ω²* = 0.156, BF10 = 300.56. Paired *t* tests were then used as post-hoc comparison (with a Bonferroni corrected threshold *p* < 0.0167). The results shows that good-person association (684 ± 81.9) have advantage over bad-person association (728.4 ± 83.4), *t*(50) = -4.41, *p* < 0.0001, Cohen’s *dz* = -0.618, 95%CI [-0.915 -0.315], BF10 = 395.4, but have no differences with neutral-person association (708 ± 79.9), *t*(32) = -2.27, *p* = 0.0278, Cohen’s *dz* = 0.317, 95%CI [ -0.597 -0.034], BF10 = 1.567. Also, the neutral-person associations was slightly faster than bad-person associations, but the evidence is not strong: *t*(32) = -2.49, *p* = 0.016, Cohen’s *dz* = -0.349, 95%CI [-0.631-0.065], BF10 = 2.5. These results suggested that the good-person facilitate the associative learning, and the evidence for the advantage of morally good condition is strong. In contrast, there was no significant effect of moral valence for the non-matched pairs (*F*(2, 100) = 0.959, *p* = 0.3866, BF10 = 0.144) (Fig 1 B).

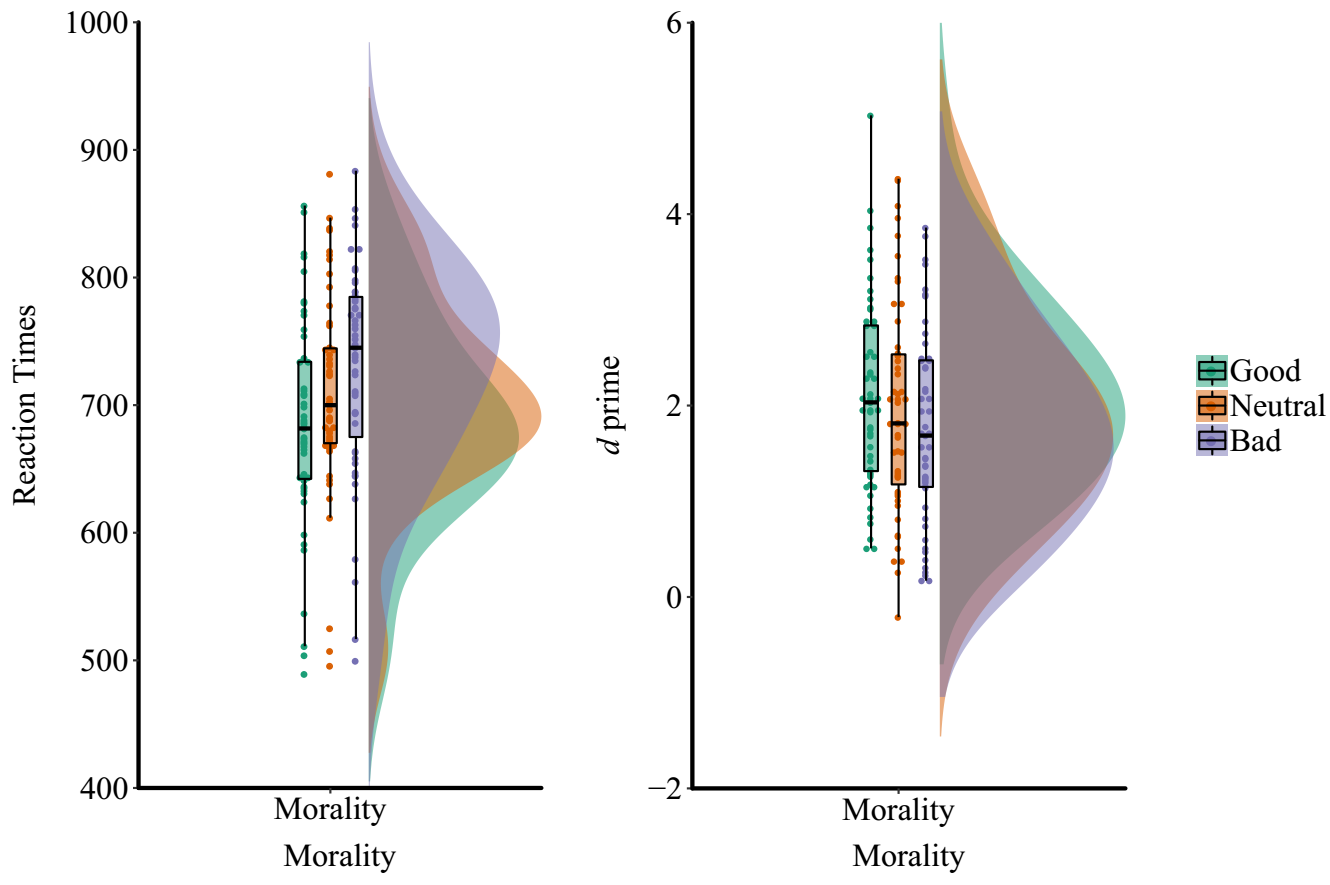


Fig 1a. Reaction times of matached trials and *d* prime of experiment 1a.

**Discussion**

This experiment was set out to examine whether or not the negative moral valenced sitmuli processed faster, as suggested by previously studies ([Anderson, Siegel, Bliss-Moreau, & Barrett, 2011](#_ENREF_1)). However, our results showed a reversed pattern, the stimuli associated with morally positive meaning were processed faster and more accurately.

To replicate this pattern of results, as well as eliminate the confounding factors such as familarity of the words used in this experiment, we conducted experiment 1b, in which the words used for labels have similar familiarity.

## Experiment 1b

In this study, we aimed at excluding the potential confouding factor of the familarity of words we used in experiment 1a, by matching the familiarity of the words.

### Method

**Participants.** 72 college students (49 female, age = 20.17 ± 2.08) participated in the current experiment. Of which, 39 were recruited from Tsinghua University in 2014; 33 college students in Wenzhou University in 2017. All participants were right-handed, and all had normal or corrected-to-normal vision. All participants were compensated for course credit (for most participants from Tsinghua community) or payment. Informed consent was obtained from all participants prior to the experiment according to procedures approved by local ethics committees.

For data from 12 participants from Tsinghua community and 8 participants from Wenzhou community were excluded from analysis due to less than 60% percent overall accuracy, leaving 52 participants (36 females, age ± 2.31).

**Stimuli and Tasks.** Three geometric shapes (triangle, square, and circle, with 3.7º × 3.7º of visual angle) were presented above a white fixation cross subtending 0.8º × 0.8º of visual angle at the center of the screen. The three shapes were randomly assigned to three labels with different moral valence: a morally bad person (“恶人”, ERen), a morally good person (“善人”, ShanRen) or a morally neutral person (“常人”, ChangRen). The order of the associations between shapes and labels was counterbalanced across participants.

Three labels used in this experiment is selected based on the rating results from an independent survey, in which participants rated the familiarity, frequency, and concreteness of eight different words online. Of the eight words, three of them are morally positive (HaoRen, ShanRen, Junzi), two of them are morally neutral (ChangRen, FanRen), and three of them are morally negative (HuaiRen, ERen, LiuMang). An independent sample consist of 35 participants (22 females, age 20.6 ± 3.11) were recruited to rate these words. Based on the ratings (see Figure 1), we selected ShanRen, ChangRen, and ERen to represent morally positve, neutral, and negative person.

Table 1b-1. Ratings of familarity, frequencies and concreteness for different labels that share similar meanings (n = 35).

|  |  |  |  |
| --- | --- | --- | --- |
| Label | Familiarity | frequencies | concreteness |
| HaoRen | 6.31 (1.01) | 5.85 (1.24) | 4.71 (1.83) |
| HuaiRen | 6 (1.44) | 4.68 (1.85) | 4.51 (1.96) |
| ChangRen | 5.49 (1.41) | 4.4 (1.63) | 4.28 (1.88) |
| ShanReng | 4.89 (1.42) | 2.8 (1.6) | 4.57 (1.59) |
| Eren | 5.08 (1.65) | 2.69 (1.85) | 4.6 (1.59) |

**Procedure.** For participants from both Tsinghua community and Wenzhou community, the procedurein the current study was exactly same as in experiment 1a. For participants in Tsinghua community, they finished a survey suite include personal distance, objective and subjective SES, belief in just world ([Wu et al., 2011](#_ENREF_33)), disgust senstivity scale ([谭永红 et al., 2007](#_ENREF_36)), trait justice ([Wu et al., 2014](#_ENREF_32)), and cognitive reflection test ([Frederick, 2005](#_ENREF_10)). For participants from Wenzhou community, they finished exactly the same questionnaires as the participants from Wenzhou University in experiment 1a.

**Analysis.** Data was analyzed as in experiment 1a.

### Results

Table 1-b shows the accuracy of performance in experiment 1b.

We replicated the advantage of positive moral valence of experiment 1a. For *d’*,a main effect of moral valence was significant, *F*(2, 102) = 14.98, *p* < 0.0001, *ω²* = 0.0485, BF10 = 7765. Paired *t* test (Bonferroni corrected threshold *p* < 0.0167) showed that the good-person association (1.872 ± 0.738) is significantly greater than the neutral-person association (1.441 ± 0.727) , *t*(51) = 5.945, *p* < 0.0001, Cohen’s *dz* = 0.824, 95% CI [0.506 1.137], BF10 = 59751, and, marginally, bad person association (1.668 ± 0.794), *t*(51) = 2.265, *p* = 0.028, Cohen’s *dz* = 0.314, 95% CI [0.034 0.591], BF10 = 1.56. Different from experiment 1a, the neutral-person and bad-person conditions also showed difference, *t*(52) = -3.132, *p* = 0.003, Cohen’s *dz* = -0.434, 95% CI [-0.717 -0.148], BF10 = 10.99 (see Figure 1b).

This pattern of results was also shown on reaction times of matched trials: themain effect of morality was significant, *F* (2,102) = 33.975, *p* < 0.0001, *ω²* = 0.386, BF10 = 1.44e+9. Paried-wise *t* test, Bonferroni corrected, for matched trials showed that RT to shapes associated with good-person ( 684.3 ± 63.2) was faster over shapes associated with bad person (728.2 ± 65.9), *t*(51) = -5.72, *p* < 0.0001, Cohen’s *dz* = -0.794, 95%CI[-1.1 -0.497], BF10 = 1.251e + 8, and also faster than neutral person association (739.6 ± 70.9), *t*(54) = -8.17, *p* < 0.0001, Cohen’s *dz* = -1.13, 95% [-1.478 -0.78], BF10 = 28341.9; while there was no difference between the neutral person and immoral person associations, *t*(51) = 1.69 *p* = 0.098, Cohen’d = 0.234, 95%CI[-0.043 0.508], BF10 = 0.565. This pattern of results was not shown on mis-matched trials, as the main effect of moral valence was not significant: *F*(2,102) =1.799, *p* = 0.17, *ω²* = 0.013, BF10 = 0.287.

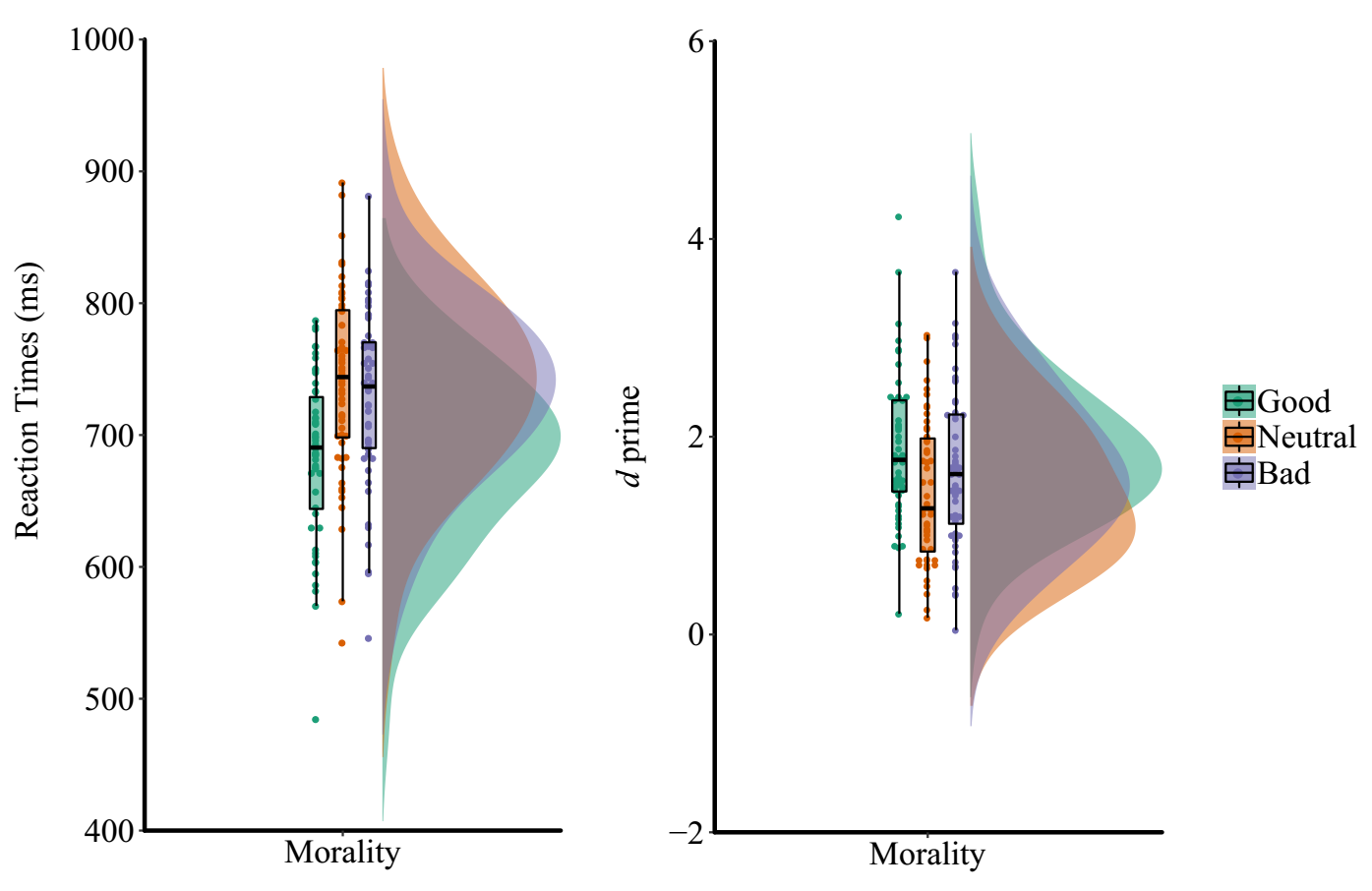


Figure 1b\_2.

**Discussion**

These results confirmed the facilitation effect of positive moral valence on the perceptual matching task. This pattern of results mimic prior results demonstrating self-bias effect on perceptual matching (Sui et al., 2012) and in line with previous studies that indirect learning of other’s moral reputation do have influence on our subsequence behavior ([Fouragnan et al., 2013](#_ENREF_9)).

Combine the experiment 1a and 1b, the facilitation effect of positive moral labels could not be explained by the familiarity of the labels used in the experiment (see table 1b-1).

One might argue that this facilitation effect of positive moral valence was caused by participants different expectation about the label, even three types of matched trials have the same probability. To test the effect of expectation, we conducted experiment 2, in which label-shape pair were presented in a sequential way: each trial the labels were presented first and then the shapes.

## Experiment 1c

In this study, we further control the valence of words using in our experiment. Instead of using label with moral valence, we used valence-neutral names in China. Participant first learn behaviors of the different person, then, they associate the names and shapes. And then they perform a name-shape matching task.

### Method

**Participants.** 23 college students (15 female, age = 22.6 ± 2.6) participated in the current experiment. All of them were recruited from Tsinghua University in 2014 (October to November). All participants were right-handed, and all had neutral or corrected-to-neutral vision. All participants were compensated for payment. Informed consent was obtained from all participants prior to the experiment according to procedures approved by local ethics committees.

**Stimuli and Tasks.** Three geometric shapes (triangle, square, and circle, with 3.7º × 3.7º of visual angle) were presented above a white fixation cross subtending 0.8º × 0.8º of visual angle at the center of the screen. The three most common names were chosen, which are neutral in moral valence before the manipulation.

Three names (Zhang, Wang, Li) were first paired with three paragraphs of behavioral description. Each description includes one sentence of biographic infomration and four sentences that describing the moral behavioral under that name. To assess the that these three descriptions represented good, neutral, and bad valence, we collected the ratings of three person on six dimensions: morality, likability, trustworthiness, dominance, competence, and aggressiviess, from an independent sample (n = 34, 18 female, age = 19.6 ± 2.05). The rating results showed that the person with morally good behavioral description has higher score on morality (M = 3.59, SD = 0.66) than neutral (M = 0.88, SD = 1.1), *t*(33) = 12.94, *p* < .001, and bad conditions (M = -3.4, SD = 1.1), *t*(33) = 30.78, *p* < .001. Neutral condition was also significant higher than bad conditions *t*(33) = 13.9, *p* < .001.

**Procedure.** After arriving the lab, participants were informed to complete two experimental tasks, first a social memory task to remember three person and their behaviors, after tested for their memory, they will finish a perceptual matching task.

In the social memory task, the descriptions of three person were presented without time limitation. Participant self-paced to memorized the behaviors of each person. After they memorizing, a recognition task was used to test their memory effect. Each participant was required to have over 95% accuracy before preceding to matching task.

The perceptual learning task was followed, three names were randomly paired with geometric shapes. Participants were required to learn the association and perform a practicing task before they start the formal experimental blocks. They kept practicing util they reached 70% accuracy. Then, they would start the perceptual matching task as in experiment 1a. They finished 6 blocks of perceptual matching trials, each have 120 trials.

**Analysis.** Data was analyzed as in experiment 1a.

### Results

Table 1-c shows the accuracy of performance in experiment 1c.

We didn’t find the advantage of positive moral valence in experiment 1a. For *d’*, the main effect of moral valence was not significant, *F*(2, 44) = 0.23, *p* = 0.798, *ω²* < 0.001, BF10 = 0.142. (see Figure 1c). For RT, we found the main effect of matchness *F*(1, 22) = 96.92, *p* < 0.0001, *ω²* = 0.148, BF10 = 34.79. There was no significant interaction between moral valence and matchness, *F*(1, 22) = 1.2, *p* = 0.31, *ω²* < 0.001, BF10 = 0.31.

### Discussion

Experiment 1c aimed at further test the robustness of the effect of positive valence. However, the negative results stopped the author (CP Hu) to collect more data to refute or support H0, therefore, this control experiment didn’t has a powerful design. With limited statistical power, this current experiment can not provide information for any conclusions.

## Experiment 2: Sequential presenting

Experiment 2 was conducted for two purpose: (1) to further confirm the facilitation effect of positive moral associations; (2) to test the effect of expectation of occurrence of each pair. In this experiment, after participant learned the assocation between labels and shapes, they were presented a label first and then a shape, they then asked to judge whether the shape matched the label or not (see ([Sui, Sun, Peng, & Humphreys, 2014](#_ENREF_29)). Previous studies showed that when the labels presented before the shapes, participants formed expectations about the shape, and therefore a top-down process were introduced into the perceptual matching processing. If the facilitation effect of postive moral valence we found in experiment 1 was mainly drive by top-down processes, this sequential presenting paradigm may eliminate or attenuate this effect; if, however, the facilitation effect ocured because of button-up processes, then, similar facilitation effect will appear even with sequential presenting paradigm.

### Method

**Participants**. Thirty-five participants (18 male, 18 to 26 years of age, M = 21.6, s.d. = 2.8) were recruited. 24 of them had participated in Experiment 1a (9 male, mean age = 21.9, s.d. = 2.9), and the time gap between these experiment 1a and experiment 2 is at least six weeks. The results of two participants were excluded from analysis, one because of an injured right middle finger (which he used to press button), and the other because of less than 60% overall accuracy, remains 33 participants (17 female, age: 21.82 ± 3.05). The data were analyzed in the same was as in experiment 1a.

**Procedure**. In Experiment 2, the sequential presenting makes the matching task much easier than experiment 1. To avoid ceiling effect on behavioral data, we did a few pilot experiments to get optimal parameters, i.e., the conditions under which participant have similar accuracy as in Experiment 1 (around 70 ~ 80% accuracy).

In the final procedure, the label (good person, bad person, or neutral person) was presented for 50 ms and then masked by a scrambled image for 200 ms. A geometric shape followed the scrambled mask for 50 ms in a noisy background (which was produced by first decomposing a square with ¾ gray area and ¼ white area to small squares with a size of 2 × 2 pixels and then re-combine these small pieces randomly), instead of pure gray background in Experiment 1. After that, a blank screen was presented 1100 ms, during which participants should press a button to indicate the label and the shape match the original association or not. Feedback was given, as in study 1. The next trial then started after 700 ~ 1100 ms blank. Other aspects of study 2 were identical to study 1.

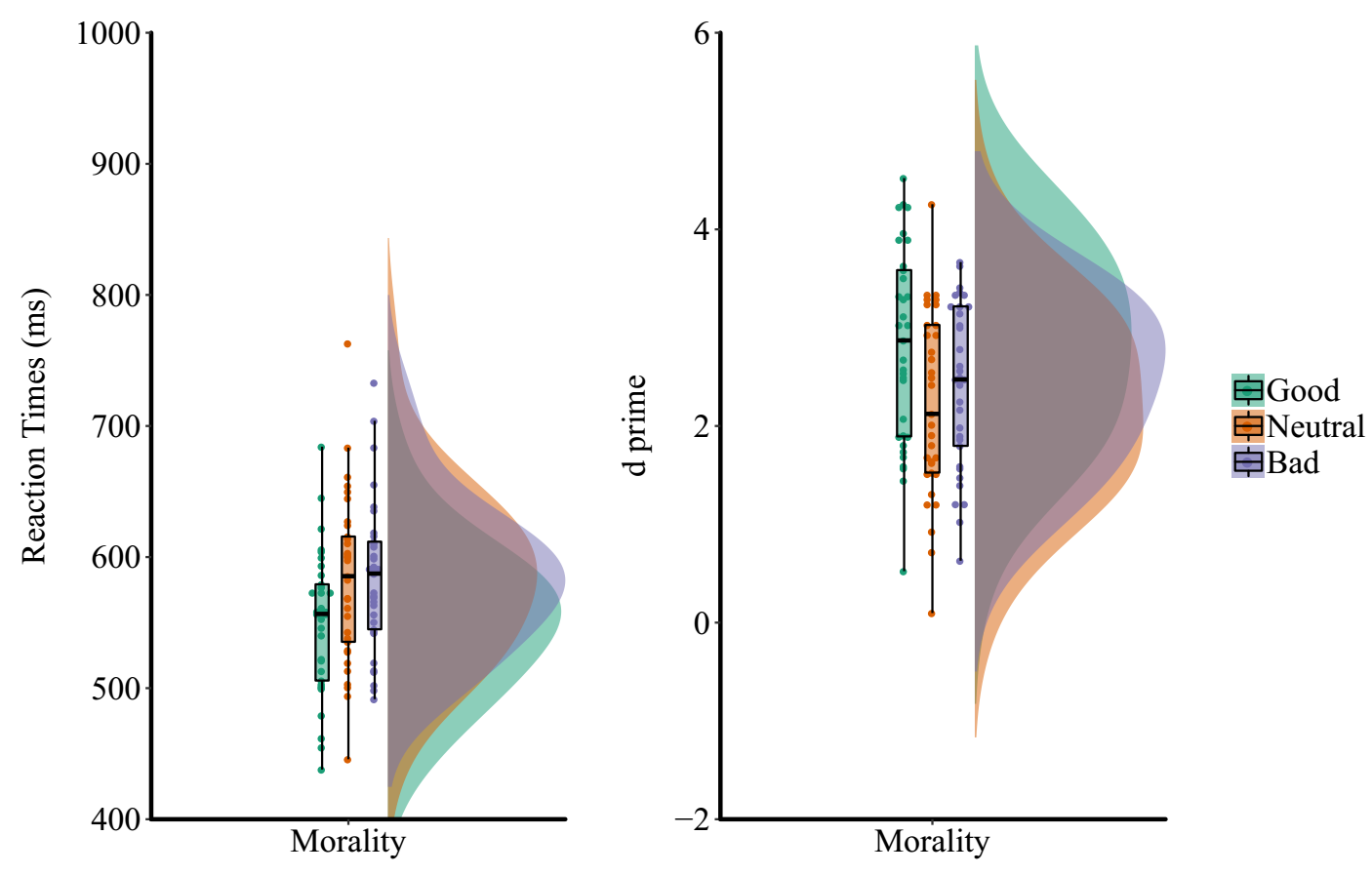
**Analysis.** Data was analyzed as in study 1a.

### Results

Less than 0.2% correct trials with less than 200ms reaction times were exlucded.

**Effect of moral valence**. As in study 1, we also found the facilitation of positive valence. An repeated measures ANOVA for *d’* revealed a significant effect of moral valence of shapes, *F*(2, 64) = 13.58, *p* < 0.0001, *ω²* = 0.279, BF10 = 1880. *d'* was larger for the good-person association (M = 2.8 ± 1.01) than for associations to the bad-person (M = 2.48 ± 0.84), *t*(32) = 4.168, *p* = 0.00022, Cohen’s *d* = 0.726, 95% CI [0.337 1.105], BF10 = 125.9, and the neutral person condition (M = 2.22 ± 0.943), *t*(32) = 4.597, *p* < 0.0001, Cohen’s *d* = 0.8, 95% CI [0.402 0.1189], BF10 = 382.1. The neutral condition and the bad-personal condition didn’t show differences, *t*(32) = -1.64, *p* = 0.111, Cohen’s *d* = -0.285, 95% CI [-0.631 0.065], BF10 = 0.622 (see Figure 2).

A repeated measures ANOVA on RTs of matched trials showed similar pattern. There was a significant effect of moral valence, *F*(2,64) = 10.56, *p* = 0.00011, *ω²* = 0.222, BF10 = 207.7. Post-hoc *t* test replicated the facilitation effect of the good-person association: responses to the good person association (549.9 ± 54.9) were faster than the neutral person association (582.5 ± 64.3), *t*(32) = -3.77, *p* = 0.00066, Cohen’s *d* = -0.657, 95% CI [-1.029, -0.275] BF10 = 46.6, and the bad person association (584.7 ± 57.9), *t*(32) = -4.005, *p* = 0.00035, Cohen’s *d* = -0.697 , 95% CI [-1.074 -0.311], BF10 = 83.1. But no difference between the neutral and bad person associations, *t*(32) = 0.264, *p* = 0.79 , Cohen’s *d* = -0.046, 95% CI [-0.387 0.296], BF10 = 0.192.



Figure\_2. Behavioral results of experiment 2.

**Cross experiments analysis.** To test the robust of the modulation effect of moral valence on perception across experiments, we conducted a repeated measures ANOVA to the combined data from participants who performed both Experiments 1a and 2, with experiments as within-subject factor. This analysis showed that the modulation effect of moral valence was not influenced by the way stimuli was presented. For the *d* prime data, the interaction between moral valence and experiment was not significant: *F*(1, 22) = 1.854, *p* = 0.1686, *ω²* = 0.0062, BF10 = 0.3633. The main effect of morality: *F*(1, 44) = 10.25, *p* = 0.0002, *ω²* = 0.0606, BF10 = 4.35; the main effect of experiment: *F*(1, 22) = 34.24, *p* < 0.0001, *ω²* = 0.2164, BF10 = 5.72e+8.

For the matched trials of RT data show the same pattern: the interaction between experiment and moral valence was not significant, *F*(1, 44) = 1.98 *p* = 0.1505, *ω²* = 0.007, BF10 = 0.3002. The main effect of morality get contridicting results from Bayesian and NHST: *F*(1, 44) = 9.16, *p* = 0.0005, *ω²* = 0.0869, BF10 = 0.6407. The main effect of experiment was significant, *F*(1, 22) = 232.61, *p* < 0.0001, *ω²* = 0.6682, BF10 = 2.364e+31.

**Discussion**

In this experiment, we repeated the results pattern that the positive moral valenced stimuli has an advantage over the neutral or the negative valenced association. Moreover, with a croass task analysis, we didn’t found evidence that the experiment task interacted with moral valence, suggesting that the effect might not be effect by experiment task.

These findings suggested that the facilitation effect of positive moral valence is robust and not affected by task. This robust effect detected by the associative learning is unexpected. As we mentioned above, there is quit a lot of study in attention and perception found that negative bias, and evolution psychologists proposed a “cheater-detection theory that people would show bias for negative moral association because of the potential harm that this can convey ([Cosmides, 1989](#_ENREF_4); [Cosmides & Tooby, 2000](#_ENREF_5)). It is probably that negative bias show on stimuli that have a threatening meaning in evolutionary history, such as snake and angry human faces [].

However, these results was consistent witht the positive bias in attention literature ([Pool, Brosch, Delplanque, & Sander, 2016](#_ENREF_22)). Given that our paradigm used association task, there was no direct threatening information. Given that this is no rewarding or punishment in both positive or negative association, the positive bias may reflect an intrinsic, self-referential process. We examined this hypothesis in study 3.

## Experiment 3a: valenced-self and other (together)

To examine the modulation effect of positive valence was an intrinsic, self-referential process, we designed study 3. In this study, moral valence was assigned to both self and a stranger. We hypothesized that the modulation effect of moral valence will be stronger for the self than for a stranger.

### Method.

**Participants.** 38 college students (16 females; 19 to 26 years of age, M = 22.03, s.d. = 2.16) participated in study 3. All participants were right-handed, and all had neutral or corrected-to-neutral vision. Informed consent was obtained from all participants prior to the experiment according to procedures approved by a local ethics committee. One female participant’s data was excluded from analysis because less than 60% overall accuracy, another one female and one male student did not finish the experiment. The final data was 35 subjects, with age of 22.17 ± 2.13.

**Design.** Study 3 added the combined moral valence with self-relevance, hence the experiment has a 2 × 3 × 2 within-subject design. The first variable was self-relevance, include two levels: self-relevance vs. stranger-relevance; the second variable was moral valence, include good, neutral and bad; the third variable was the matching between shape and label: match vs. mismatch.

**Stimuli**. The stimuli used in study 3 share the same parameters with experiment 1 & 2, with an increased number. 6 shapes were included (triangle, square, circle, trapezoid, diamond, regular pentagon), as well as 6 labels (good self, neutral self, bad self, good person, bad person, neutral person). To match the concreteness of the label, we asked participant to chosen an unfamiliar name of their own gender to be the stranger.

**Procedure**. After being fully explained and signed the informed consent, participants were instructed to chose a name that can represent a stranger with same gender as the participant themselves, from a name pool that of similar age as the participants. Before experiment, the experimenter explained the meaning of each label to participants. For example, the “good self” mean the morally good side of themselves, them could imagine the moment when they do something’s morally applauded, “bad self” means the morally bad side of themselves, they could also imagine the moment when they doing something morally wrong, and “neutral self” means the aspect of self that doesn’t related to morality, they could imagine the moment when they doing something irrelevant to morality. In the same sense, the “good other”, “bad other”, and “neutral other” means the three different aspects of the stranger, whose name was chosen before the experiment. Then, the experiment proceeded as study 1a. Each participant finished 6 blocks, each have 120 trials. The sequence of trials was pseudo-randomized so that there are 10 matched trials for each condition and 10 non-matched trials for each condition (good self, neutral sef, bad self, good other, neutral other, bad other) for each block.

**Data Analysis**. Data analysis followed strategies described in the general method section. Reaction times and d prime data were analyzed as in study 1 and study 2, except on more within-subject variable (i.e., self-relevance) was included in the repeated measures ANOVA.

### Results

**Effect of moral valence.**

***d’****.* A repeated measures ANOVA for *d’*, with moral valence (good, bad or neutral) and personal association (self vs. other) as within-subjects factors, showed main effect of moral valence, *F*(2,68) = 11.1, *p* < 0.0001 , *ω²* = 0.221, BF10 = 14.2. However, the personal association did not yield significant main effect *F(*1, 34) = 3.219, *p* = 0.096, *ω²* = 0.058, BF10 = 1.26, neither the interaction between moral valence and personal association, *F*(2,68) = 1.304, *p* = 0.039 , *ω²* = 0.063, BF10 = 1.729.

To examine our hypothesis, we conducted two repeated measures ANOVA for self and other condition separately. For the self-condition the results showed a significant effect of moral valence on *d’*: *F*(1.654,56.25) = 13.98, *p* < 0.0001, *ω²* = 0.268, BF10 = 2317.4. Simple effect analysis showed that *d’* was larger for the good-self association condition (1.972 ± 0.846) than for the bad-self association condition (1.431 ± 0.602), *t*(34) = 3.8551, *p* = 0.0005, Cohen’s *d* = 0.652, 95% CI [0.282 1.013], BF10 = 60, and the neutral-self association condition (M = 1.409 ± 0.692), *t*(34) = 4.505, *p* < 0.0001, Cohen’s *d* = 0.762, 95% CI [0.38 1.134], BF10 = 327.8; there was no difference between the neutral-self and bad-self conditions, *t*(34) = -0.371, *p* = 0.814, Cohen’s *d* = -0.04, 95% CI [-0.371 0.292], BF10 = 0.186 (see Figure 4B). For the stranger condition, there was no effect of the morality of the association on *d’*, *F*(2,68) = 0.383, *p* = 0.683, *ω²* < 0.0001, BF10 = 0.12.

We also tested the effect of personal association by comparing *d’* values for difference moral valence level. The results showed that **the bad-self association condition (1.43 ± 0.602) was responded worse than the bad-stranger association condition (1.798 ± 0.702), *t*(34) = -3.195, *p* = 0.003 , Cohen’s *d* = -0.54,** 95% CI [-0.892 -0.181], BF10 = 12**.** The neutral-self (1.409 ± 0.692) was also showed a trendency to be lower than the neutral-stranger (1.72 ± 0.86), *t*(34) = -2.149, *p* = 0.039 , Cohen’s *drm*  = -0.363, 95% CI [-0.703 -0.019], BF10 = 1.382. While the good-self association condition (1.972 ± 0.846) and good-stranger conditions (1.841 ± 0.82) are not differ from each other, *t*(34) = 0.649, *p* = 0.521, Cohen’s *drm*  = 0.11, 95% CI[-0.223 0.441], BF10 = 0.22.

***RTs.*** For the matched trials, a repeated measures ANOVA was carried out with moral valence and personal association as two within-subject variables. The main effect of moral valence was significant, *F*(2, 68) = 35.7, *p* = 2.4394e-11, *ω²* = 0.495, *ηG²* =, BF10 = 227119.9, but not the main effect of self-relevance, *F*(1, 34) = 0.197, *p* = 0.66, *ω ²* < 0.0001, BF10 = 0.162. The interaction between moral valence and self-relevance was significant, but the evidence is not strong, *F*(2, 68) = 3.88, *p* = 0.025, *ω ²*= 0.075, BF10 = 2.98.

Then, the matched trials were analyzed for the self-relevance and stranger-relevance pairs separately. The results showed a significant effect of moral valence for the self-relevance condition, *F*(2,68) = 30.39, *p* < 0.0001, *ω²* = 0 .453, BF10 = 2.1598e + 7. Paired *t* tests showed that responses to the good-self association (M=713, SD = 71) were faster than to bad-self associations (M = 772, SD = 59.8), *t*(34) = -5.67, *p* < 0.001, Cohen’s *d* = -0.956, 95% CI [-1.353 -0.55], BF10 = 7695.69, and to neutral-self association (M = 776, SD = 64.2), *t*(34) = -7.396, *p* < 0.001, Cohen’s *d* = -1.25, 95% CI [-1.689 -0.8007], BF10 = 933766.55. The neutral-self and bad-self associations did not differ significantly, *t*(34) = 0.481, *p* = 0.63 , Cohen’s *d*= 0.081, 95% CI [-0.25111 0.4126], BF10 = 0.202. The effect of moral valence was not significant for the stranger-relevance conditions, *F*(2, 68) = 2.8487, *p* = 0.064849, *ω²*= 0.0495, BF10 = 0.85.

We also analyzed the effect of self-relevance on the different moral valence levels. The results showed that for all three different valence levels, there were no significant differences between self-referential condition and stranger-referential condition: neutral-self (775 ± 69) vs. neutral-stranger (754 ± 62), *t*(34) = 1.998, *p* = 0.0538, Cohen’s *d* = 0.338, 95% CI [-0.0055 0.6761], BF10 = 1.063; good-self (713 ± 71) vs. good-stranger (735 ± 78.6), *t*(34) = -1.37, *p* = 0.179, Cohen’s *d* = -0.232, 95% CI [-0.5657 0.1059]), BF10 = 0.4267; bad-self (772 ± 59.8) vs. bad-other (760 ± 62), *t*(34) = 1.23, *p* = 0.226, Cohen’s *d* = -0.208, 95% CI [-0.128 0.5417], BF10 = 0.3636.

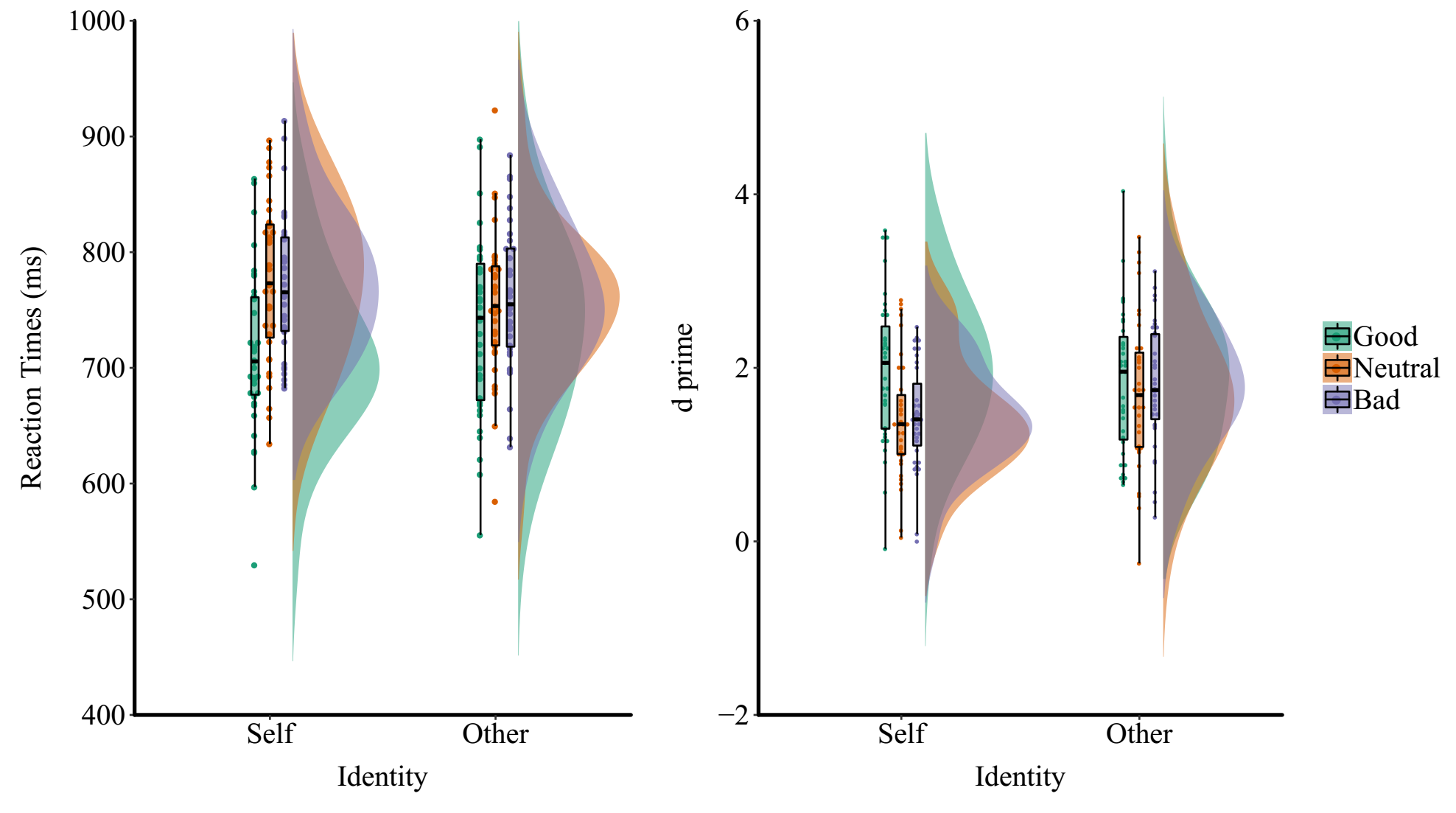


Figure 3a.

**Discussion**

Results of study 3a combined moral valence to self or stranger, to examine the self-relevance effect in moral associations. Notably, the modulation effects of moral association emerged only for self-related stimuli, with RTs and *d*’ to ‘good self’ stimuli being enhanced relative to those linked to ‘bad self’ stimuli.

These data suggest that, for self relevant stimuli, participants prioritized the good self over the bad self. This prioritization did not take place for stimuli associated with other people. In addition, prioritization of the good moral valence for self-related stimuli appeared to come at some trade-off for the other associations to self-stimuli, which tended to be more difficult to respond to compared with the neutral and bad associations to strangers. Given that we found robust effects of moral association in Experiments 1 and 2, there is a contrast with the stranger association here. However, this pattern of results can be interpreted if the self is used as a default when moral associations are made (in Experiments 1 and 2). This default link to the self would lead to prioritization of good associations over associations to a bad moral label. The lack of an advantage for neutral-self and bad-self stimuli, relative to neutral-stranger and bad-stranger associations, also contracdicts the substantial literature showing an advantage for self over other stimuli in perceptual matching (e.g., Sui et al., 2012; Humphreys & Sui, 2015). This in turn suggests that matching is contingent on the relative prioritization of stimuli – if there is prioritization to the good self, then the neutral and bad selves are relatively de-prioritized.

## Experiment 3b: valenced-self and other (separated)

In study 3a, participants had to remember 6 pairs of association, which cause high cogitive load during the whole exepriment. To eliminate the influence of cognitive load, we conducted study 3b, in which participant learn three aspect of self and stranger seperately in to consecutive task. We hypothesize that we will replicate the pattern of study 3a, i.e., the effect of moral valence only occurs for self-relevant conditions.

### Method.

**Participants.** Study 3b were finished in 2017, at that time we have calculated that the effect size (Cohen’s *d*) of good-person (or good-self) vs. bad-person (or bad-other) was between 0.47 ~ 0.53, based on study 1, 2, 3a, and 4. Based on this effect size, we estimated that 54 participants would allow we to detect the effec size of Cohen’s = 0.5 with 95% power and alpha = 0.05, using G\*power 3.192 ([Faul, Erdfelder, Buchner, & Lang, 2009](#_ENREF_7); [Faul, Erdfelder, Lang, & Buchner, 2007](#_ENREF_8)). Therefore, we planned to stop after we arrived this number. During the data collected at Wenzhou University, 61 participants (45 females; 19 to 25 years of age, M = 20.42 ± 1.77, four participants didn’t report their age) came to the testing room and we tested all of them during a single day. All participants were right-handed, and all had neutral or corrected-to-neutral vision. Informed consent was obtained from all participants prior to the experiment according to procedures approved by a local ethics committee. 4 participants’ data were excluded from analysis because their over all accuracy was lower than 60%, 1 more participant waw excluded because of zero hit rate for one condition, leaving 56 participants (44 females; 19 to 25 years old, age = 20.27 ± 1.6, four participants didn’t report their age).

**Design.** Study 3b has the same experimental design as 3a, with a 2× 3× 2 within-subject design. The first variable was self-relevance, include two levels: self-relevant vs. stranger-relevant; the second variable was moral valence, include good, neutral and bad; the third variable was the matching between shape and label: match vs. mismatch.

**Stimuli**. The stimuli used in study 3b share the same parameters with experiment 3a. 6 shapes were included (triangle, square, circle, trapezoid, diamond, regular pentagon), as well as 6 labels, but the labels changed to “good self”, “neutral self”, “bad self”, “good him/her”, bad him/her”, “neutral him/her”, the stranger’s label is consistent with participants’ gender. Same as study 3a, we asked participant to chosen an unfamiliar name of their own gender to be the stranger before showing them the relationship. Note, because of implementing error, the personal distance data didn’t collect for this experiment.

**Procedure**. In this experiment, participants finished two matching tasks, i.e., self-matching task, and other-matching task. In the self-matching task, participants first associate the three aspects of self to three different shapes, and then perform the matching task. In the other-matching task, participants first associate the three aspects of the stranger to three different shapes, and then perform the matching task. The order of self-task and other-task are counter-balanced among participants.

Different from experiment 3a, after presenting the stimuli pair for 100ms, participant has 1900 ms to response, and they were feedbacked with both accuracy and reaction time.

As in study 3a, before each task, the intruction showed the meaning of each label to participants. The self-matching task and other-matching task were randomized between participants. Each participant finished 6 blocks, each have 120 trials.

**Data Analysis**. Data analysis is the same as study 3a.

### Results

**Effect of moral valence.**

***d’****.* A repeated measures ANOVA for *d’*, with moral valence (good, bad or neutral) and personal association (self vs. other) as within-subjects factors, showed the interaction between moral valence and personal association, *F*(2,110) = 16.18, *p* < 0.001, *ω²* = 0.44, BF10 = 549.9, the main effect of moral valence (despite conflicting results from Bayesian and classic statistics), *F*(2,110) = 5.29, *p* = .0064, *ω²* = 0.0134, BF10 = 0.63, and the main effect of self-referential, *F(*1, 55) = 5.08, *p* = 0.0281, *ω²* = 0.0192, BF10 = 20.7.

We conducted two repeated measures ANOVA for self and other condition separately. For the results from self-condition aligned with previous experiments, showed an main effect of moral valence on *d’*: *F*(2, 110) = 18.3, *p* < 0.001, *ω²* = 0.084, BF10 = 98488. Simple effect analysis showed that the *d’* was larger for good self (2.23 ± 0.81) than for bad self (1.66 ± 0.74), *t*(55) = 6.11, *p* < 0.001, Cohen’s *d*= 0.817, 95% CI [0.511 1.117] , BF10 = 8.43, and neutral self (1.91 ± 0.66), *t*(55) = 3.03, *p* = 0.0038, Cohen’s *d* = 0.404, 95%CI [0.13 0.675] , BF10 = 1.33e+5. There was also higher d’ for neutral-self condition than bad-self conditions, *t*(34) = 3.02, *p* = 0.0039, Cohen’s *d* = 0.403, 95% CI [0.129 0.674], BF10 = 8.22.

For the stranger condition, there **was no evidence for the effect of the morality of the association on *d’***, *F*(2,110) = 3.27, *p* = 0.042, *ω²* = 0.0079, BF10 = 0.96.

We also tested the effect of personal association by comparing *d’* values for difference moral valence level. The results showed that **the bad-self association condition was responded worse than the bad-stranger association condition (2.18 ± 0.911), *t*(55) = -4.1, *p* < 0.001, Cohen’s *d* = -0.548,** 95% CI [-0.827 -0.265], BF10 = 167**.** The neutral-self was also worse than the neutral-stranger (2.28 ± 1), *t*(55) = -3.15, *p* = 0.0026 , Cohen’s *d* = -0.422, 95% CI [-0.693 -0.146], BF10 = 11.7. While the good-self association condition and good-stranger conditions (2.03 ± 0.89) are not differ from each other, *t*(55) = 1.394, *p* = 0.169, Cohen’s *d* = 0.186, 95% CI[-0.079 0.449], BF10 = 0.364.

***RTs.*** For the matched trials, a repeated measures ANOVA was carried out with moral valence and personal association as two within-subject variables. The main effect of moral valence was significant, *F*(2, 110) = 24.3, *p* < 0.001, *ω²* = 0.05, BF10 = 186.9, the main effect of self-relevance, *F*(1, 55) = 48.5, *p* < 0.001, *ω²* = 0.174, BF10 = 7.44e+19. The interaction between moral valence and self-relevance was significant, *F*(2, 110) = 6.14, *p* = 0.003, *ω ²*= 0.011, BF10 = 1.17.

Then, the matched trials were analyzed for the self-relevance and stranger-relevance pairs separately. The results showed a significant effect of moral valence for the self condition, *F*(2,110) = 23.97, *p* < 0.001, *ω²* = 0.095, BF10 = 4.54e+6. Paired *t* tests showed that responses to the good-self association (817 **±** 119) were faster than to bad-self associations (915 **±** 132), *t*(55) = -8.78, *p* < 0.001, Cohen’s *d* = -1.173, 95% CI [-1.511 -0.828], BF10 = 1.84e+9, and to neutral-self association (880 **±** 116), *t*(55) = 3.748, *p* < 0.0001, Cohen’s *d* = -0.501, 95% CI [-0.777 -0.221], BF10 = 58.7. The neutral-self was faster than the bad-self associations, *t*(55) = -2.41, *p* = 0.019 , Cohen’s *d*= -0.321, 95% CI [-0.589 -0.051], BF10 = 2.03.

**The effect of moral valence was significant for the stranger-relevance conditions,** *F*(2, 68) = 5.96, *p* = 0.0035, *ω²*= 0.0114, BF10 = 8.55. the good-other condition (734 ± 158) didn’t differ from neutral-other condition (735 ±160), *t*(55) = -0.07, *p* = 0.946, Cohen’s *d*= -0.009, 95% CI [-0.271 -0.293], BF10 = 0.15, but faster than the bad other condition (776 ± 173), *t*(55) = -3.14, *p* = 0.0027, Cohen’s *d*= -0.419, 95% CI [-0.691 -0.144], BF10 = 11.3. The neutral-other condition also faster than the bad-other condition, *t*(55) = -3.232, *p* = 0.0021, Cohen’s *d*= -0.432, 95% CI [-0.704 -0.156], BF10 = 14.3.

We also analyzed the effect of self-relevance on the different moral valence levels. The results showed that for all three different valence levels, there the self condition was responded slower than other condition: good-self vs. good-stranger, *t*(55) = 4.29, *p* < 0.001, Cohen’s *d* = 0.573, 95% CI [0.288 0.854], BF10 = 297.2; neutral-self vs. neutral -stranger, *t*(55) = 7.17, *p* < 0.001, Cohen’s *d* = 0.958, 95% CI [0.638 1.272]), BF10 = 5.77e+6; bad-self vs. bad-other, *t*(55) = 6.03, *p* < 0.001, Cohen’s *drm* = 0.806, 95% CI [0.5 1.11], BF10 = 100208.03.

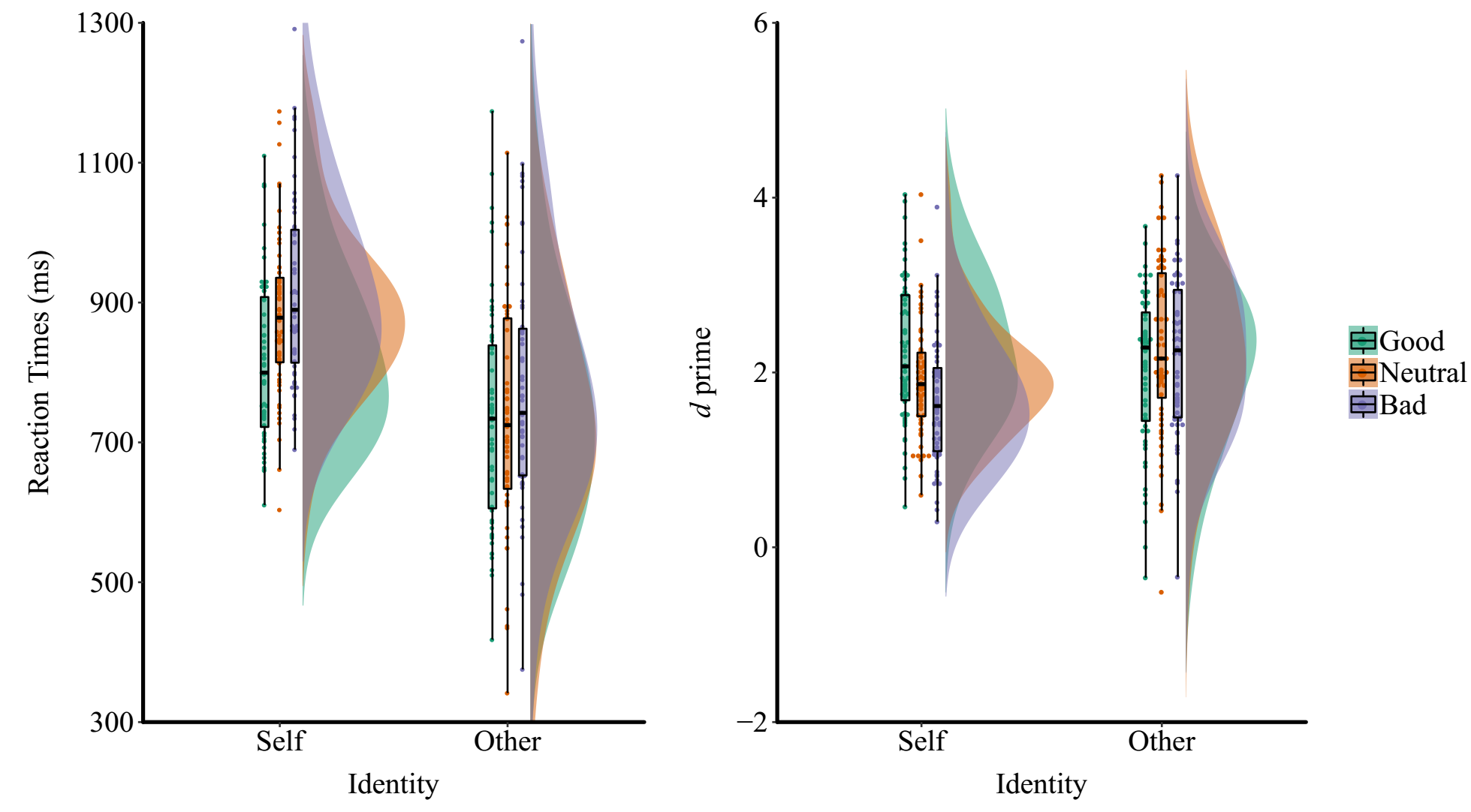


Figure 3b.

**Discussion**

In experiment 3b, we separated two tasks so that the participants have low cognitive load when doing the perceptual matching task. We replicated the effect that the valence effect is stronger when the stimuli were self-referential. however, we found that the self-referential condition is performed worse than other-referential condition. This effect is difficult to explain.

## Experiment 4a: Valence as distractor

In study 1-3 participants made explicit judgements about moral associations. In Experiment 4, we examined whether the effects of moral valence occur even when the moral valence information might not be relevent to the task. In this study participants made perceptual match judgements to associations between self-referential labels and shapes (cf. Sui et al., 2012), but we presented labels of different moral valence levels in the shapes.

### Method

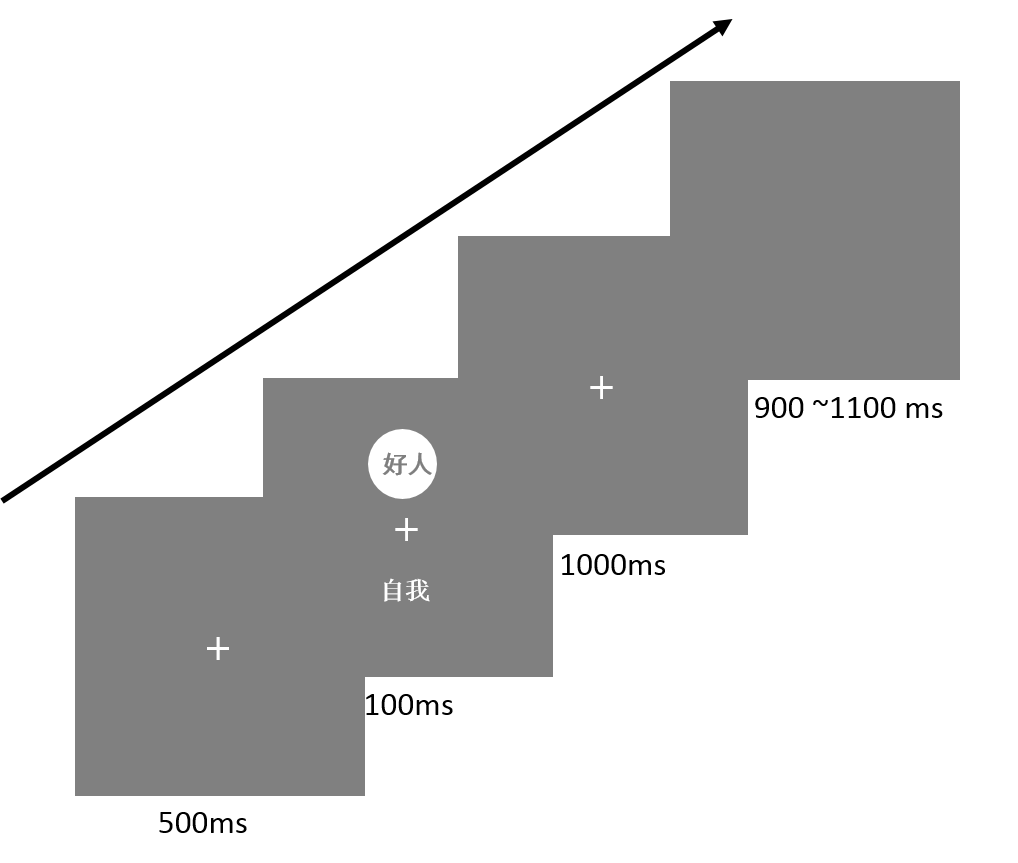
**Participants.** 64 participants (37 female, age = 19.7 ± 1.22) participated the current study, 32 of them were from Tsinghua Universtiy in 2015, the rest were from Wenzhou University parpticipated in 2017. All participants were right-handed, and all had neutral or corrected-to-neutral vision. Informed consent was obtained from all participants prior to the experiment according to procedures approved by a local ethics committee. The data of three participants from Tsinghua site and two participants from Wenzhou site were excluded from analysis because their accuracy was close to chance (< 0.6). The results for the remaining 59 participants (33 female, age = 19.78 ± 1.2) were analyzed and reported.

**Experimental design.** As in Experiment 3, a 2× 3× 2 within-subject design was used. The first variable was self-relevance (self and stranger associations); the second variable was moral valence (good, neutral and bad associations); the third variable was the matching between shape and label (matching vs. non-match for the personal association).

However, in this the task, participants only learn the association between two geometric shapes and two labels (self and other), i.e., only self-relevance were related to the task. The moral valence manipulation was achieved by embeding the personal label of the labels in the geometric shapes, see below. For simplicity, the trials where shapes where paired with self and with a word of “good person” inside were shorted as good-self condition, similarly, the trials where shapes paired with the self and with a word of “bad person” inside were shorted as bad-self condition. Hence, we also have six conditions: good-self, neutral-self, bad-self, good-other, neutral-other, and bad-other.

**Stimuli**. 2 shapes were included (circle, square) and each appeared above a central fixation cross with the personal label appearing below. However, the shapes were not empty but with a two-Chinese-character word in the middle, the word was one of three labels with different moral valence: “good person”, “bad person” and “neutral person”. Before the experiment, participants learned the self/other association, and were informed to only response to the association between shapes configure and the labels below the fixation, but ignore the words within shapes. Besides the behavioral experiments, participants from Tsinghua community also finished questionnaires as Experiments 3, and participants from Wenzhou community finished a series of questionnaire as the other experiment finished in Wenzhou.

**Procedure**. The procedure was similar to Experiment 1. There were 6 blocks of trial, each with 120 trials for 2017 data. Due to procedure error, the data collected in 2015 in Tsinghua community only have 60 trials for each block, i.e., 30 trials per condition.



### Results

Correct responses shorter than 200ms were excluded from the analysis, eliminating fewer than 0.2% of the trials overall. For the 2017 dataset, 0.7% data were excluded. Overall, 0.25% trials were removed.

**The self-referential effect.** ***d’*.** We first check the effect of site (Tsinghua or Wenzhou), and found that site does have significat main effect or interaction with other factors (*F*(1,57) < 0.6, *p* > 0.60, *ω²* < 0.001, BF10 < 1).

A repeated measures ANOVA for the *d’* results, with moral valence (good, bad or neutral) and self-relevance (self- vs. other-relevance) as within-subjects factors, showed a significant main effect of self-relevance, *F*(1, 58) = 121, *p* < 0.0001, *ω²*= 0.667 , BF10 = 3.53e+42. The main effect of moral valence was not significant, *F*(2, 116) = 0.5345, *p* = 0.587, *ω²*< 0.0001, BF10 = 0.035. However, there is not enough evidence concerning the interaction between the moral category and self-relevance, *F*(1.52, 88.24) = 4.1235, *p* = 0.0186, *ω²* = 0.0499, BF10 = 0.7033*.*

We conducted the post-hoc analysis as in previous experiment. The results showed **that the good-self condition (2.55 ± 0.856) had higher *d* prime as compared to bad-self condition (2.377 ± 0.803): *t*(58) = 2.339, *p* = 0.0228, Cohen’s *d* = 0.3045, 95% CI [0.0422 0.5643], BF10 = 1.75.** However, there was no evidnece that good self condition and neutral self condition (2.448± 0.776) had differences on *d* prime: *t*(58) = 1.575, *p* = 0.121, Cohen’s *d* = 0.2051, 95% CI [-0.0537 0.462], BF10 = 0.456. For neutral-self condition and the bad-self condition, there is weak evidnece showing that there is no difference between these two conditions: *t*(58) = 0.966, *p* = 0.338, Cohen’s *d* = -0.1258, 95% CI [-0.131 0.381], BF10 = 0.222)

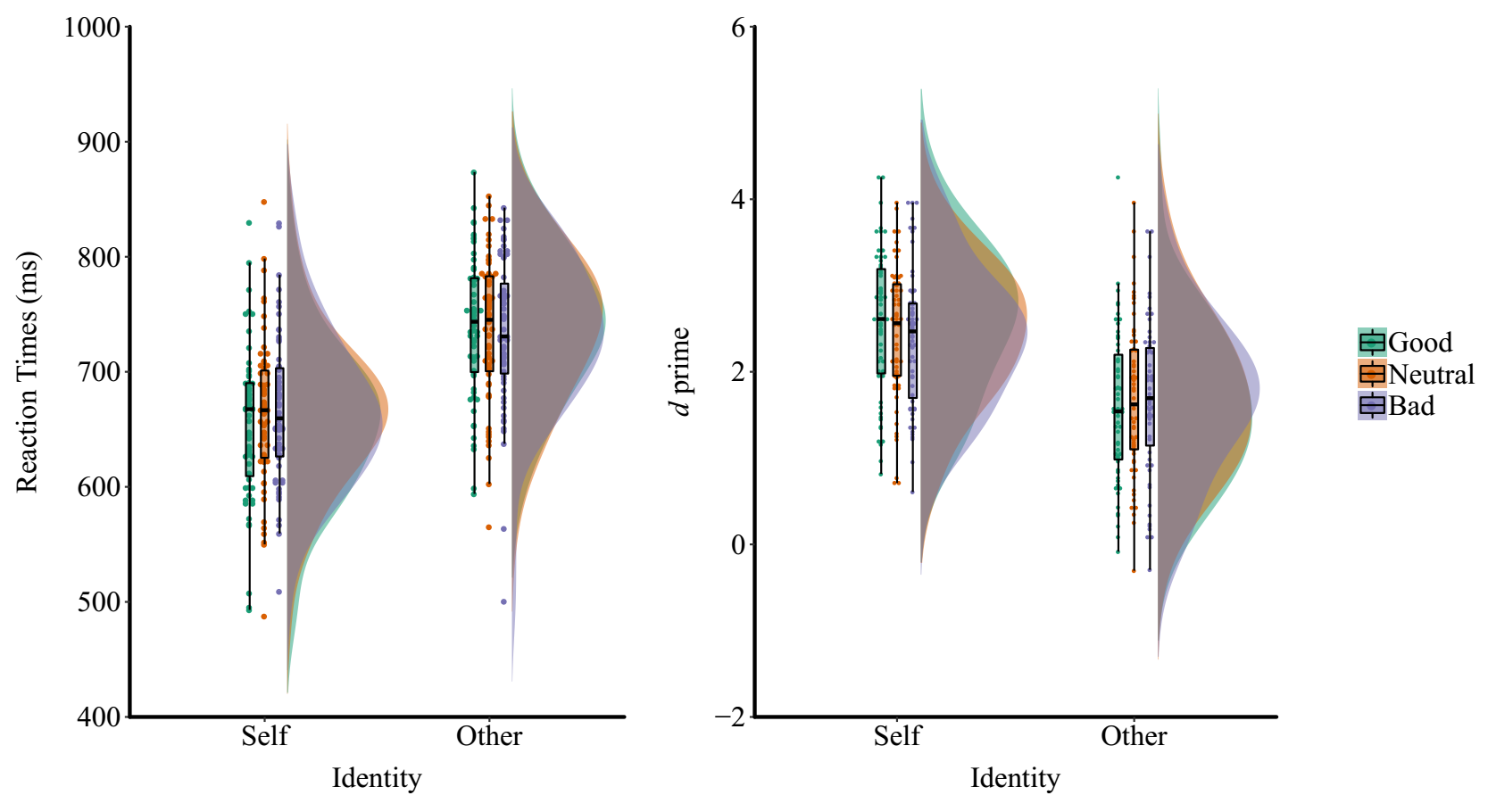
**RTs**.

The site didn’t have main effect of interaction with other factors (*F*(1,57) < 2.43, *p* > 0.126, BF10  < 2).

A repeated measures ANOVA for matched trials revealed a reliable main effect of self-referential, *F*(1, 58) = 124.15, *p* < 0.0001, *ω²* = 0.6724, BF10 = 6.897e+53, while the main effect of moral valence was not significant, *F*(2,116) = 0.942, *p* = .393, *ω²* < 0.0001, BF10 = 0.0367. The interaction between moral valence and self-relevance was also significant but the evidence is ambiguous, *F*(2,116) = 4.41, *p* = 0.014, *ω²* = 0.0542, BF10 = 0.241.

We broke down the interaction by analyzing the data for self and other pairs separately. There was a significant effect of moral valence for self-stimuli, *F*(2,116) = 6.293, *p* = 0.00254, *ω²* = 0.0817, BF10 = 11.16. Paired *t* tests showed that **good-self condition (654 ± 67) were faster relative to bad-self condition (665 ± 64.6), *t*(58) = -3.47, *p* < 0.0001, Cohen’s *drm* = -0.451 CI [-0.718 -0.182], BF10 = 27.0, and over neutral-self condition (664 ± 64), *t*(58) = -2.78, *p* = 0.0073, Cohen’s *drm* = -0.362, 95% CI [-0.624 -0.097], BF10 = 4.63.** The neutral-self and bad-self conditionsdid not differ, *t*(58) = -0.383, *p* = 0.7, Cohen’s *drm* = 0.0499, CI [-0.305 0.206], BF10 = 0.153 (see Figure 5).

For the stranger condition, the results showed that there was no difference among these conditions, *F*(2, 116) = 0.351, *p* = .704, *ω²* < 0.0001, BF10 = 0.077.



## Experiment 4b: Self as distractor

In study 4b we further investigated whether the effects of moral valence would modulate by the presence of self. In this study, participants made perceptual match judgements to associations between different moral valence and shapes as in study 1-3. However, as in study 4, we made the self-referential task as an implicit task, to examine the interaction between self-referential effect and modulation of positive moral valence.

### Method

**Participants**. 53 college students participated in the current study (26 females, age = 21.26 ± 1.78), of which 34 college students (26 females; 18 to 25 years of age, mean age = 21.26 ±1.78) from Tshinghua Univeristy, 19 participants (13 female, age = 19.32 ± 1.06) from Wenzhou University. The data from eight participants and one female participate were excluded from analysis because their accuracy was less than 0.6 (2 from Tsinghua sample and 6 from Wenzhou sample). The final valid dataset includes 45 participants (33 female, age = 20.78 ± 1.76).

**Experimental design**. The experimental design of this experiment is same as experiment 4a: a 3× 2 × 2 within-subject design with moral valence (good, neutral and bad associations), self-relatedness (self vs. other), and matchness between shape and label (match vs. mismatch for the personal association) as within-subject variables. However, in the current task, the participants learned the associations between three shapes and three labels with different moral valence: good-person, neutral-person, and bad-person. While the word “self” or “other” were presented in the shapes (see below).

**Stimuli**. In this task, 3 shapes were included (circle, square, and trapezoid) and were presented above a central fixation cross, as in previous experiments. Similar to experiment 4a, the shapes were not empty but with a two-Chinese-character word in the middle corresponding to the labels “self” and “other”. Before the experiment, we informed participants only response to the relationship between shapes configure and the labels below the fixation, ignoring the wordswithin each shape. Besides the behavioral experiments, participants also finished questionnaires as Experiments 1-3.

**Procedure**. The procedure was similar to Experiment 4 a. Both samples of participants finished 6 blocks of trial, each with 120 trials.

### Results

Correct responses shorter than 200ms were excluded from the analysis, eliminating fewer than 0.04% of the trials overall.

***d’*.** An ANOVA for the *d’* results, with moral category (good, bad or neutral) and shape identity (self vs. other) as within-subjects factors, with the testing sites as between subject factor, we didn’t found the main effect of sites or interaction with other factors (*F*(1,43) <= 2.042, *p* <= 0.136), therefore we combined the data together.

The data showed no interaction between moral valence and self-relevance: *F*(2,88) = 0.5334, *p* = 0.5885 , *ω²* < 0.0001, BF10 = 2.106, or the main effect of moral valence, *F*(2,88) = 2.34, *p* = .1022, *ω²* = 0.0286, BF10 = 1.91, or effect of self-relatedness, *F*(1, 44) < 0.001, *p* = 0.9937,  *ω²* < 0.0001, BF10 = 0.1335*.*

Nonetheless, given that the previous studies found interaction between moral valence and self-relatedness, we exploratorily analyzed the data from self associations and other association as previous experiment. For the self condition, the main effect of morality is not significant, *F*(1.75, 76.86) = 3.0796, *p* = 0.0584, *ω²* = 0.0437, BF10 = 0.8998, nor the other-relevance condition, *F*(1.625, 71.5) = 1.074, *p* = 0.3356, *ω²* = 0.0016, BF10 = 0.1751. No evidence for differences between good-self condition (2.15 ± 0.82) and neutral self condition (1.98 ± 073), *t*(44) = 1.892; *p* = 0.0651, Cohen’d = 0.282, 95%CI [-0.017 0.578], BF10 = 0.826**. There was weak evidence that good self condition is better than bad self condition (1.91 ± 0.77), *t*(44) = 2.0491, *p* = 0.0464, Cohen’d = 0.306, 95%CI [0.005 0.603], BF10 = 1.083.** For the neutral self and bad-self condition, the current showed that here is no differences between the two conditions, *t*(44) = 0.767, *p* = 0.4474, Cohen’d = 0.114, 95%CI [-0.179 0.407], BF10 = 0.213.

For other condition?

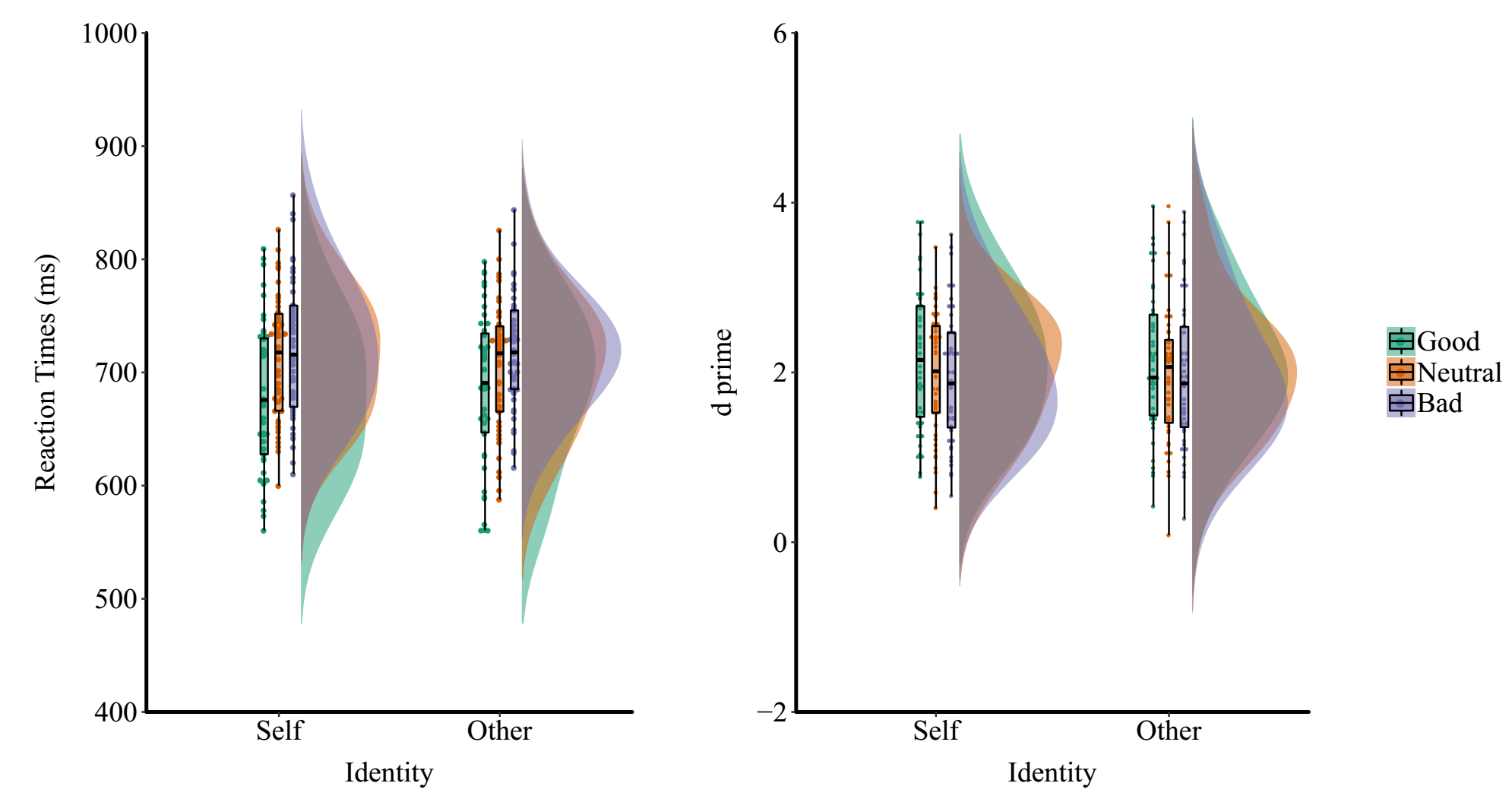
**RTs**. A repeated ANOVA for match trials revealed that location didn’t have main effect or interaction with other factors (*Fs*(1,43) < 2.538, *ps* > = 0.085), therefore we analyze the data together.

For matched trials, the evidence for interaction between moral valence and self-relevance is not clear, as the NHST and Bayes Factor showed contridicting results, *F*(2,88) = 5.22, *p* = .0072, *ω*² = 0.0849, BF10 = 0.199. The main effect of moral valence was significant, *F*(2,88) = 11.04, *p* = .000053, *ω*² = 0.1808, BF10 = 2.371e+7, while there is no main effect of self-relatedness, *F*(1,44) = 0.233, *p* = 0.6317, *ω*² < 0.0001, BF10 = 0.135.

We broke down the interaction by analyzing the data for self and other pairs separately. There was a significant effect of moral valence for self relevant stimuli, *F*(1.74, 76.5) = 13.687, *p* < 0.0001, *ω*² = 0.218, BF10 = 2668.5. **Paired *t* tests showed that good-self association (680 ± 65.7) were faster than bad-self associations (720 ± 60.2), *t*(44) = -4.22, *p* < .001, Cohen’s *d* = -0.629 CI [-0.947 -0.306], BF10 = 200, and neutral-self association (712 ± 54.9), *t*(44) = -4.67, *p* < 0.001, Cohen’s *d* = -0.696, 95% CI [-1.019 -0.367], BF10 = 745.3.** The neutral-self and bad-self associations did not differ, *t*(44) = -1.04, *p* = .31, Cohen’s *d* = -0.155, 95%CI [-0.448 0.14], BF10 = 0.267. RTs in good-self condition were facilitated but without performance being impaired for bad-self associations (relative to the neutral self)(see Figure 5).

For the stranger condition, the main effect of moral valence was also significant, *F*(2,88) = 7.09, *p* = 0.0018, *ω*² = 0.037, BF10 = 21. The RT for good-other association condition (688 ± 66.9) is faster than the bad-other association condition (718 ± 49.7), *t*(44) = -3.353, *p* = 0.0017, Cohen’s *d* = -0.4999, 95%CI [-0.8075 -0.1872], BF10 = 18.84. The RT for good-other condition is slightly faster than neutral-other condition (704 ± 57.1), but the evidence is not strong, *t* (44) = -2.21, *p* = 0.0324, Cohen’s *d* = -0.3294, 95%CI [-6278 -0.0275], BF10 = 1.454. While there is is no strong evidence about the differences between bad-other vs. neutral-other conditions, *t*(44) = -1.8267, *p* = 0.0745, Cohen’s *d* = -0.2723, 95%CI [-0.5685 0.0268], BF10 = 0.743.

We also analyzed the effect of self-relevance on the moral associations. Pair-wise *t* tests showed that **the good self was faster than good other, but the evidence is weak, *t*(44) = -2.17, *p* = 0.036 , Cohen’s *d* = -0.323, 95%CI [-0.621 -0.021], BF10 = 1.3375.** **The neutral-self condition is slower than neutral-other condition, *t*(44) = 3.06, *p* = 0.0037, Cohen’s *d* = 0.457, 95% CI [0.1471 0.762], BF10 = 9.197.** And there is no differences between bad-self (720 ± 60.2) and bad-other (718 ± 49.7), *t*(44) = 0.623, *p* = 0.54, Cohen’s *d* = -0.093 , 95% CI[-0.2004 0.385], BF10 = 0.194.



### Discussion

In experiment 4, we manipulated the task so that the moral valence (experiment 4a) or the self-relatedness (experiment 4b) become irrelevant to the task. We found a robust effect of the task: when the self-relatedness is task related, the results showed a strong effect of self-relatedness; in contrast, when moral valence become task related the main effect of moral valence was strong. However, the task irrelevant stimuli in the shape exerted influence on the performance as well. The good self conditions (the shape associated the self and with a “good person” within the shape) performed better than bad self conditions. These results demonstrated that even when the moral valence and self-reference was separated by the task, participants can still couple the the self with the morally good, and facilitated the perceptual decision making.

## Experiment 5a: EEG study 1

Experiment 5 was conducted to study the neural correlates of the positive prioritization effect. The behavioral paradigm is same as experiment 2.

### Method

**Participants.** 25 young healthy adults (8 female, mean age = 22.04±5.02) participated this experiment. All were right handed and had normal or corrected-to-normal vision. The eeg data of two pariticipants were excluded from analysis because of too much artifacts.

**Experimental design**. The experimental design of this experiment is same as experiment 1a: a 3 × 2 within-subject design with moral valence (good, neutral and bad associations) and matchness between shape and label (match vs. mismatch for the personal association) as within-subject variables. However, in the current task, the participants learned the associations between three shapes and three labels with different moral valence: good-person, neutral-person, and bad-person. While the word “self” or “other” were presented in the shapes (see below).

**Stimuli**. Three geometric shapes (triangle, square and circle, each 4.6º × 4.6º of visual angle) were presented at the center of screen for 50 ms after 500ms of fixation (0.8º × 0.8º of visual angle). The association of the three shapes to bad person (“坏人, HuaiRen”), good person (“好人, HaoRen”) or ordinary person (“常人, ChangRen”) was counterbalanced across participants. The words bad person, good person or ordinary person (3.6º × 1.6º) was also displayed at the center fo the screen. Participants had to judge whether the pairings of label and shape matched (e.g., Does the circle represent a bad person?). The experiment was run on a PC using E-prime software (version 2.0). These stimuli were displayed on a 22-in CRT monitor (1024×768 at 100Hz).

We used backward masking to avoid over-processing of the moral words, in which a scrabmled picture were presented for 900 ms after the label. Also, to avoid the celling effect on accruacy, shapes were presented on a noisy background based on our pilot studies. The noisy images were made by scrambling a picutre of 3/4gray and ¼ white at resoluation of 2 × 2 pixel.

**Procedure**. The procedure was similar to Experiment 2. Participants finished 6 blocks of trial, each with 180 trials. In total, participants finished 180 trials for each combination of condition.

As in experiment 2 ([Sui et al., 2012](#_ENREF_28)), subjects first learned the associations between labels and shapes and then completed a shape-label matching task (e.g., good person-triangle). In each trial of the matching task, a fixation were first presented for 500 ms, followed by a 50 ms label; then, a scramled picture presented 900 ms. After the backward mask, the shape were presented on a noisy background for 50ms. Participant have to response in 1000ms after the presentation of the shape, and finnally, a feedback screen was presented for 500 ms (see figure 1). The inter-trial interval (ITI) were randomly varied at the range of 1000 ~ 1400 ms.

All the stimuli were presented on a gray background (RGB: 127, 127, 127). E-primed 2.0 was used to present stimuli and collect behavioral results. Data were collected and analyzed when accuracy performance in total reached 60%.

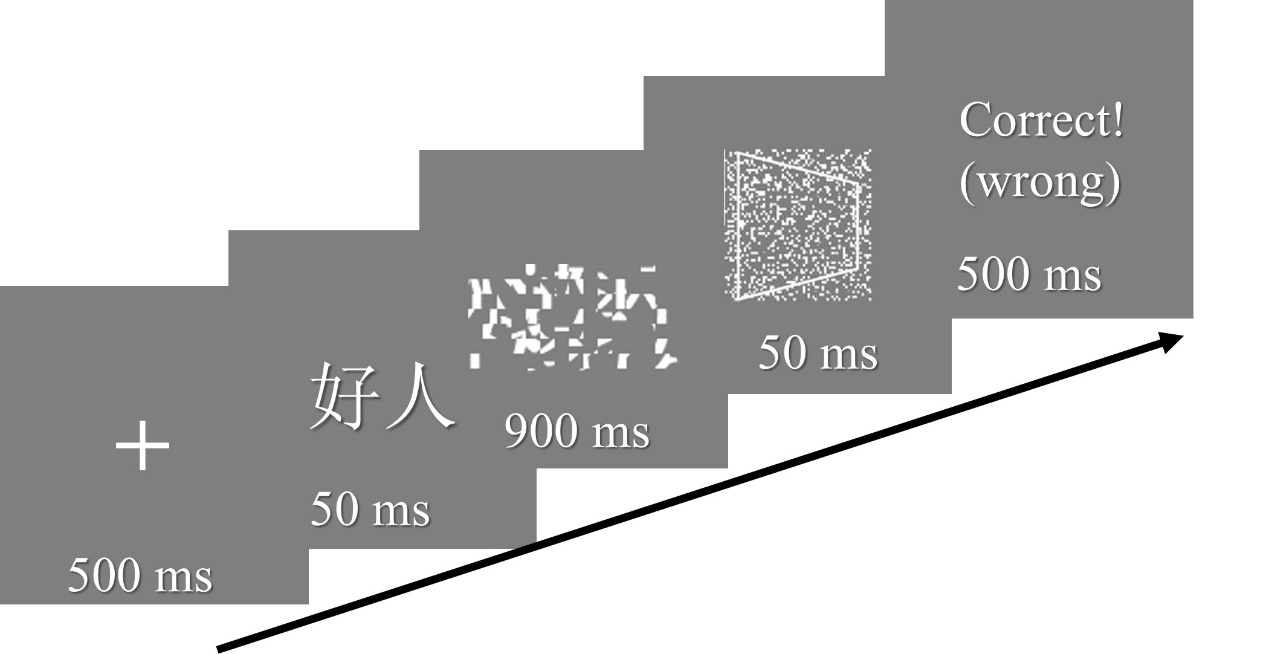


Figure 1. Examples of stimuli and procedures in the experiment

### Results



Only the behavioral results were reported here.

ANOVA for the reaction times showed a significant interaction of shape category and matching judgment, *F*(2,48) = 21.9, *p* < .001, *ω²* =0.021, BF10 = 13567. The main effect of moral valence also significant, *F*(2,48) = 28.8, *p* < .001, *ω²* = 0.019, BF10 = 5.6. The main effect of match, *F*(1,24) = 64.3, *p* < .001, *ω²* = 0.103, BF10 = 1.74e+16. Test of matching trials were followed. The data showed a significant effect of shape category for matched pairs, *F*(2, 48) = 34.7, *p* < .001, *ω²* = .075, BF10 = 1.37e+7. Paired-wise *t*-tests, with Boferroni corrected at p < 0.0167, showed that good- person (470 ± 69) was repsonded faster than bad-person (M = 523, SD = 81), *t*(24) = -8.14, *p* < 0.001, Cohen’s *dz* = -1.63, 95%CI[-2.22 -1.01], BF10 = 584172, and normal person (501± 73), *t*(24) = -5.17 ,*p* < 0.001, cohen’s *dz*= -1.03, 95%CI[-1.52 -0.54], BF10 = 877.4; Neutral person association is also faster than bad person, *t*(24) = -3.28, *p* = 0.003, cohen’s dz = -0.66, 95%CI[-1.08 -0.22], BF10 = 12.8. In contrast, there was no signicant effect of shape category for the nonmatched pairs, *F*(2, 48) = 0.42, *p* = 0.66, *ω²* < 0.001, BF10 = 0.158.

ANOVA for the *d’* showed a significant effect of moral valence, *F*(2,48) = 4.626, *p* =.0146 *ω²* = 0.019, BF10 = 3.3. Paired *t* test, with boferroni correcttion (*p* < 0.0167), shows that good-person (M=3.13, SD = 0.54) association have advantage over normal person (M=2.88, SD = 0.7), *t*(24) = 2.92, *p* = 0.0076, cohen’s *dz* = 0.583, 95%CI[0.153 1.003], BF10 = 6; while good person and bad person (M=3.03, SD = 0.71) may not differ, *t*(24) = 1.51, *p* = 0.144, cohen’s *dz* = 0.302, 95%CI[-0.102 0.7], BF10 = 0.57; neither did bad persoan and normal person had no differences *t*(24) = -1.6, *p* = 0.123, cohen’s *dz* = 0.32, 95%CI[-0.719 0.086], BF10 = 0.64 (figure B).

## Experiment 5b: EEG study 2

Experiment 5.2 was conducted to study the neural correlates of the interation between self-prioritization and valence-prioritization.

### Method

**Participants.** 23 young healthy adults (XX female, age = 22.04±5.02) participated this experiment. All were right handed and had normal or corrected-to-normal vision. The eeg data of two pariticipants were excluded from analysis because of too much artifacts.

**Design**. The experimental design of this experiment is same as experiment 3a: 2(self vs stranger) × 3 (moral status: good, bad, normal) ×2 (matching judgment: match vs. nonmatch) within subject design; 180 trials for each matched trial, and 90 trials for each nonmatched trials (6\*180 + 6\*90 =1620 trials).

**Stimuli**. Six geometric shapes (triangle, sqaure, circle, trapezoid, dimond, regular pentagon, 118 \* 118 pixels, 4.6º × 4.6º of visual angle) were presented at the center of screen for 50 ms after 500ms of fixation (0.8º × 0.8º of visual angle). The association of the 6 shapes to three aspect of two person (self vs. stranger) was counterbalanced across participants. The words bad person, good person or ordinary person (3.6º × 1.6º) was also displayed at the center fo the screen. Participants had to judge whether the pairings of label and shape matched (e.g., Does the circle represent a bad person?). The experiment was run on a PC using E-prime software (version 2.0). These stimuli were displayed on a 22-in CRT monitor (1024×768 at 100Hz).

As the experiment 5, we used backward masking to avoid over-processing of the moral words, in which a scrabmled picture were presented for 900 ms after the label. Also, to avoid the celling effect on accruacy, shapes were presented on a noisy background based on our pilot studies. The noisy images were made by scrambling a picutre of 3/4gray and ¼ white at resoluation of 2 × 2 pixel.

**Procedure**. In this experiment, we used a two-day paradigm to avoid high error rate at the EEG recording. At day 1, participant underwent a behavioral experiment as experiment 3a, except that the stimuli were presented sequentially.

At day 2, partcipant came to lab again and this time there finish the experiment with EEG recording.

In both day, subjects first learned the associations between labels and shapes and then completed a shape-label matching task (e.g., good person-triangle). In each trial of the matching task, a fixation was first presented for 500 ms, followed by a 50 ms label; then, a scramled picture presented 900 ms. After the backward mask, the shape was presented on a noisy background for 50ms. Participant have to response in 1000ms after the presentation of the shape, and finnally, a feedback screen was presented for 500 ms (see figure 1). The inter-trial interval (ITI) were randomly varied at the range of 1000 ~ 1400 ms.

All the stimuli were presented on a gray background (RGB: 127, 127, 127). E-primed 2.0 was used to present stimuli and collect behavioral results. Data were collected and analyzed when accuracy performance in total reached 60%.

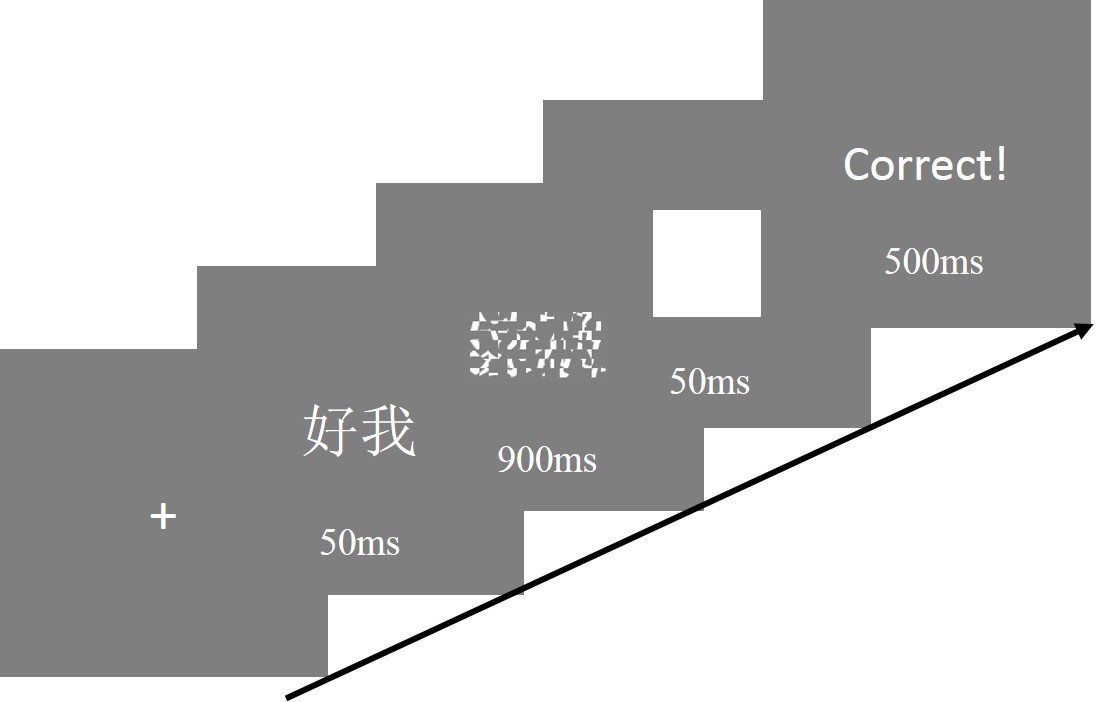


Figure 1. Examples of stimuli and procedures in the experiment

### Results

Only the behavioral results were reported here.

**Day1.** ANOVA for the reaction times showed a significant interaction of shape category and matching judgment, *F*(2,48) = 21.85, *p* = 0.000, *ηp²*=0.48. The main effect of moral valence also significant, *F*(2,48) = 30.4, *p* = .000, *ηp²*=559. The main effect of match, *F*(1,24) = 64, *p* = .000, *ηp²*= 0.727. Test of matching trials were followed. The data showed a significant effect of shape category for matched pairs, *F*(2, 48) = 35.6, *p* =.000, *ηp²*= .597. paired-wise *t*-tests, with Boferroni corrected at p < 0.0167, showed that bad-person (M = 0.93, SD = 0.09) association is slower than good- person *p* = 0.000, cohen’s *dz* = , and normal person, *p* = 0.008, cohen’s *dz*= ; normal person association is also slower than good person, *p* = 0.000, cohen’s dz =. In contrast, there was no signicant effect of shape category for the nonmatched pairs ( *F* (2, 48) = 0.42, *p* = 0.66 ).

ANOVA for the *d’* showed a significant effect of moral valence, *F*(2,42) = 11.98, *p* < 0.001 *ω²*= 0.0398, BF10 = 143.7; The main effect of self-relevance is not sig. *F*(1, 21) = 1.21, *p* = 0.284, *ω²*= 0.0006, BF10 = 2.64e-5.; the interaction was sig., *F*(2,42) = 12.88, *p* < 0.001 *ω²*= 0.041, BF10 = 708. Paired *t* test showed that good-self (M= 2.71, SD = 0.93) association have advantage over normal self (1.98 ± 0.71), *t*(21) = 5.98, *p* < 0.0001, Cohen’s *dz* = 1.28, 95CI [0.7 1.83], BF10 = 3412.7, and bad self (2.07 ± 0.72), *t*(21) = 6.55, *p* < 0.0001, Cohen’s *dz* = 1.4, 95%CI [0.8 1.98], BF10 =10932.6. Bad self and normal self did not show differences *t*(21) = -1.06, *p* = 0.3, Cohen’s *dz* = -0.23, 95%CI [-0.65 0.20], BF10 =0.37 (figure B).

**Day2.** ANOVA for the reaction times showed a significant interaction of shape category and matching judgment,

## Experiment 6: Specificity of Moral Self?

Experiment 6 was conducted to study the specificity of moral self.

### Method

**Participants.** 43 participants recruited from Tsinghua University university community (21 females; 18 to 29 years of age, mean age = 22.93 ±2.56). All participants were right-handed, and all had normal or corrected-to-normal vision. Informed consent was obtained from all participants prior to the experiment according to procedures approved by the local ethics committee. The data from 5 participants were excluded from analysis, one (male) didn’t finished the experiment, and the other four accuracy was less than 60%. The results for the remaining 38 subjects, with mean age of 22.875 ± 2.59, were included in data analyses.

**Design**. A 4 × 3 × 2 within-subject design was used. The first independent variable was stimuli categories (morality, atttractiveness of people, attractiveness of scene, and emotional words); the second independent variables is valence (positive, neutral and negative); the third variable was the matching between shape and label (match vs. mismatch for the association). The task was to learn the association between each geometric shape and the self/other label.

**Stimuli**. 4 sets of shapes were included (three circle, three rectangle, three kind of triangle, and three kinds of quadrangle), each set of shape were paired with one category of words, counter-balanced across subjects. Besides the behavioral experiments, participants also finished questionnaires XXXXXXXX.

**Procedure.** Participants finish 4 session of experiment, and each include one experiment as in experiment 1. And the order of each category was randomnized for each participant. Each session started with a practice, and proceed to formal experiment when reached over 60% accuracy. Each session included 6 blocks of trial, each with 120 trials.

### Results

Correct responses shorter than 200ms were excluded from the analysis, eliminating fewer than 5% of the trials overall.

***d’*.** An ANOVA for the *d’* results, with stimuli categories (morality, attractiveness of person, attractiveness of scene, and emotion) and valence (positive, neutral, and negative) as within-subjects factors, showed a significant interaction between stimuli categories and valence: Greenhouse-Geisser corrected *F*(4.61, 170.7) = 4.726, *p* < 0.001, *ω2* = 0.089. The main effect of stimuli categories was not significant. *F*(3, 111) = 0.694, *p* = 0.558, *ω2 <* 0.001, and the main effect of valence was significant, *F*(2, 74) = 22.834, *p* < 0.001, *ω2* = 0.362.

Therefore, we conducted four ANOVA for each stimulus separately. As for the morality condition condition, results showed a significant effect of morality on *d’*: (F (2,74) = 2.402, *p* = 0.098, *ω2* = 0.035. Simple effect analysis for *d’* : morally good (M = 2.2 SD = 0.923) than for morally bad (M = 2.03, SD = 0.902) (*t*(37) = 1.481, *p(turkey)* = 0.306. Morally good vs morally neutral (1.96 + 0.993), t(37) = 2.14, p(tukey) = 0.89. No difference between bad person and normal person(*t*(37) = -0.659, *p* = 0.788, Cohen’s *dz*= ) .

As for the personal attractiveness condition, results showed a significant effect of valence on *d’*: (F (2,74) = 16.9, *p* < 0.001, *ω2* = 0.292. Simple effect analysis showed that the *d’* was larger for attactiveness person (M = 2.528, SD = 0.943) than for ugly person (M = 1.924, SD = 1.04) (*t*(37) = -5.056, *p(turkey)* < 0.001, Cohen’s *dz*= ) , and normal person (M = 1.929, SD = 1.04) (*t*(37) = -5.014, *p(turkey)* < 0.001, Cohen’s *dz*=). But the normal person condition is not different from ugly person condition (*t*(37) = 0.043, *p* = 0.999, Cohen’s *dz*=).

As for the scene attractiveness condition, results showed a significant effect of valence on *d’*: (F (2,74) = 14.1, *p* < 0.001, *ω2* = 0.254. Simple effect analysis showed that the *d’* was larger for attactiveness scence (M = 2.444, SD = 0.864) than for ugly scene (M = 1.978, SD = 0.869) (*t*(37) = 3.556, *p(turkey)* = 0.002, Cohen’s *dz*= ) , and normal scence (M = 1.764, SD = 1.07) (*t*(37) = 5.118, *p(turkey)* < 0.001, Cohen’s *dz*=). But the normal scene condition is not different from ugly scene condition (*t*(37) = -1.632, *p* = 0.239, Cohen’s *dz*=).

As for the emotion condition, results showed a significant effect of valence on *d’*: (F (2,74) = 4.726, *p* = 0.012, *ω2* = 0.088. Simple effect analysis showed that the *d’* was larger for positive emotion (M = 2.28, SD = 0.952) than for negative emotion (M = 1.965, SD = 0.869) (*t*(37) = 2.544 , *p(turkey)* = 0.034, Cohen’s *dz*= ). Also, the neutral emotion condition (M = 2.308, SD = 0.769) was larger than negative emotion, (*t*(37) = 2.767, *p(turkey)* = 0.019, Cohen’s *dz*=). But the positive emotion condition is not different from neutral condition (*t*(37) = -0.224, *p* = 0.973, Cohen’s *dz*=).

We also tested the effect of stimuli categorization by comparing *d’* values for each valence. The results showed that, **for the positive valence**, the effect of stimuli was not significant, F(3,111) = 2.139, p = 0.099, *ω2* = 0.029; **for the neutral valence**, the effect of stimuli was significant, F(3,111) = 5.662, *p* < 0.001, *ω2* = 0.108; **for the negative condition**, the effect of stimuli was not significant, F(3,111) = 0.226, *p* = 0.878, *ω2* < 0.001. We further tested the effect of stimuli under neutral condition by carried out an simple effect analysis, which showed that the neutral emotion condition has greater d prime than neutral moral condition (t(37) = -2.557, p(tukey) = 0.057), normal scene (t (37) = -4.005, p(tukey) < 0.001, and neutral attractive perosn, t(37) = -2.786, p(tukey) = 0.031. there was no other significant difference between other conditions, p(tukey)s > 0.472.

**Figure 1. the d prime and reaction times for each conditions**

**RTs**. A 4\*3\*2 repeated ANOVA was carried out for reaction times. The main effect of matchness was significant: F(1,37) = 210.9, p < 0.001, *ω2* = 0.843; The main effect of stimuli categories was not singificant, F(1.78,66) = 0.357, p = 0.784, *ω2* = ; the main effect of valence was significant, F(2,74) = 54.16, p < 0.001, *ω2* = 0.58. The interaction between matchness and stimuli categories is not significant, F(3, 111) = 1.16, p = 0.328, *ω2* = 0.004. The interaction between matchness and valence is significant: F(2, 74) = 26.11, p < 0.001, *ω2* = 0.395. The interaction between cateogories and valence is significant: F(4.2, 156.2) = 7.292, p < 0.001, *ω2* = 0.142. The interaction between matchness, categorization and valence is not sigificant: F(3.37, 124) = 0.287, p = 0.943, *ω2* < 0.0001.

We analyzed the matched and mismatched trials as previous studies. For **match trials**, the main effect of Categories: *F*(1.9, 75.257) = 0.463, *p* = 0. 709, *ω2* < 0.0001; The main effect of Valence: *F*(2, 74) = 51.901, *p* < 0.001, *ω2* = 0.569; the interaction between Categories and Valence: F(3.44, 127.3) = 3.298, *p* = 0.018, *ω2* = 0.057.

We then tessted the effect of valence for each stimuli type:

For moral stimuli: there was a significant effect of valence, F(2,74) = 7.51, p = 0.001, *ω2* = 0.145; Simple effect analysis showed that morally positve condition (656 + 97.3) has faster response than morally negative conditions (692 + 90.8): t(37) = -3.875, p < 0.001, Cohen’s d = while there is no differences between morally positive and morally neutral (673 + 97.3), t(37) = -1.956, p(tukey) =0.13, or between morally negative and morally neutral, t(37) = -1.919, p(tukey) = 0.14.

For attractiveness of person: main effect of valence was sig. F(2,74) = 40.6, p < 0.001, *ω2* = 0.507, beauty perosn (638 + 56.8) faster than neutral (691 + 68), t(37) = -6.77, p(tukey) < 0.001, and ugly person (705 + 73.5), t(37) = -8.54, p(tukey) < 0.001, but no difference between neg and neutral, t(37) = -1.769, p = 0.187.

For attractiveness of scene: main effect of valence: F(2, 74) = 15.8, p < 0.001, *ω2* = 0.278; beautiful scence ( 651 + 68) faster than neutral (696 + 77), t(37) = -4.88, p(tukey) < 0.001, and ugly (696.4 +71.8), t(37) = -4.852, p(tukey) < 0.001. But no difference between neutral and ugly, t(37) = 0.032, p(tukey) = 0.999.

For emotion: main effect of valence, F(1.5, 55.6) = 9.325, p < 0.001, *ω2* = 0.178. Positive emotional words (654 + 87) was faster than negative words (695 + 70), t(1,37) = -4.084, p(tukey) < 0.001; neutral condiotn (662 +58.6) is also faster than negative words, t(37) = -3.257, p(tukey) = 0.005. No difference between positive and neutral conditions, t(37) = -0.827, p(tukey) = 0.687.

The effect of stimuli under different valence: For positive valence: No main effect of stimuli category, F(2.2,82.1) = 0.601, p = 0.568, *ω2* < 0.0001;

For Neutral valence: Main effect of sitmuli categoriz, F(2.2, 80.5) = 3.7, p = 0.021, *ω2* = 0.07; Neutral emotion is faster than neutral scene, t(37) = 3.045, p(tukey) = 0.015.

For Negative valence: No main effect of stimuli, F(2.33, 86.3) = 0.429, p = 0.683, *ω2* < 0.0001.

For **the mismatch trials**, ANOVA did not show significant effect of stimuli categories, *F*(1.78, 65.9) = 0.35, *p* = 0.789, *ω2* < 0.001. There was significant main effect of valence, *F*(2,74) = 11.91, *p* < .001, *ω2* = 0.221, 95% CI []), and a significant interaction (*F*(4.37, 161.9) = 5.52, *p* < 0.001, *ω2* = 0.106, 95% CI []).

For moral condition, no main effect for mismatched trials: F(2,74) = 1.651, p = 0.199, *ω2* = 0.017

For personal attractiveness, main effect for mismathced trials: F(2,74) = 20.42, p < 0.001, *ω2* = 0.335; beatiful person (709 + 55.8) is faster than neutral (741 + 49), t(37) = -6.17, p(tukey) < 0.001, and ugly person ( 732 + 59), t(37) = -4.5, p(tukey) < 0.001. No difference between neutral and ugly person, t(37) = .166, p(tukey) = 0.227.

For scene attractiveness, main effect for mismatched trials: F(2, 74) = 6.8, p = 0.002,

*ω2* = 0.131. beautiful scene (722 + 63.9) is faster than neutral scene (745 + 55.5), t(37) = -3.64, p(tukey) = 0.001; no difference between beautiful scene and ugly scence ( 731 + 60), t (37) = -1.29, p(tukey) = 0.404; no differences between neutral and ugly scence, t(37) = 2.347, p(tukey) = 0.056.

For emotion, no main effect of valence for mismatched trials, F(2,74) = 0.468, p = 0.628, *ω2*  < 0.0001.

## Cross-experiment analysis

To get a better estimation of the effect in the current study, we combined the data of our 11 experiments and from 2 experiment from another study ([Hu, Lan, Macrae, & Sui, in Prep](#_ENREF_13)).

First, we conducted repeated measure ANOVAs for the psychological distance across all experiment, to check the validity of psychological distance. We predicted that the distance between self and good person should be the shortest, while self and bad-person would show longest distance.

Second, we conducted random effect model meta-analysis of the effect size of *d*’ and RTs across our 10 experiments. Mini meta-analysis has been shown to be an effective way to get more precise and stable estimation of effect ([Goh, Hall, & Rosenthal, 2016](#_ENREF_11)). We conducted a mini meta-analysis to estimate the effects observed in current study.

Third, we conducted a correlation analysis for behavioral data and score data, i.e., between psychological distance and the bias in perceptual matching task.

### Methods

**Software**. Meta-analysis was carried out in R 3.2. As for the meta-analysis of the effect size of *d’* and RTs, we used “metafor” package ([Viechtbauer, 2010](#_ENREF_30)).

**Data analysis.** Before analyzing, the psychological distance data were normalized by dividing the sum of psychological distance in six conditions. The RTs data were also normalized (see data analysis section of experiment 1).

### Results

**Personal distance.**

A repeated measure ANOVA for normalized psychological distance, with psychological distance (6 different distance) as within subject factor and experiments (5 experiments) as between subject factor, showed a significant main effect of psychological distance, *F*(3.8885, 937.126) = 262.69, *p* < 0.001, *ω*² = 0.5195; while the main effect of experiment was not significant, *F*(5, 236) < 0.001, *p* = 1, *ω*² < 0.001, the interaction between psychological distance and experiments were not significant, *F*(19.35, 913) = 1.936, *p* = 0.009, *ω*² = 0.0039. Given the insignificant interaction, we combine the data from all experiments.

Post hoc analysis showed that only self-bad and good-bad distance showed no difference between: *t*(241) = 0.0096, *p* = 0.9208, Cohen’s *d* = 0.0064, 95% CI[-0.1196 0.1324], BF10 = . The self-good and self-neutral were not signficant after correct: *t*(241) = 2.7686, *p* = 0.0061, Cohen’s *d* = 0.178, 95% CI[0.0508 0.3048]. The distance from shortest to longest as figure below showed: self-neutral = self-good < good-neutral < bad-neutral <good-bad = self-bad

The results show that while the psychological distance between good, neutral and bad personal are gradiant, the distance between self and bad, self and good, self and neutral are rather binary. These binary results aresimilar to the reaction and *d’*.



***d’* and RT**

The meta-analytic results (table 10) showed that the response under the good person association condition were significantly better than that of the neutral person association condition, both in *d’* (Cohen’s *d* = 0.27, 95% CI[0.008 0.45]) and RTs (Cohen’s *d* = -0.40, 95% CI [-0.60 -0.20]). Also, the performance under good-person association condition was also significantly better than that under bad-person association condition, both in *d*’(Cohen’s *d* = 0.25, 95% CI [0.14 0.37]) and RTs (Cohen’s *d* = -0.47, 95% CI [-0.66 -0.29]). But the difference between the performance under neutral-person association condition and bad association condition is not significant, in *d’* (Cohen’s *d* = 0, 95% CI [-0.14 0.14]) and RT (Cohen’s *d* = -0.08, 95% CI [-0.19 0.04]).

We also conducted meta-analyses separately for the moral valence effect for self- and other-referential condition (experimental 3a, 3b, 4a, 4b, and two exepriment from ([Hu et al., in Prep](#_ENREF_13))).

Table X. Effect size of mini meta-analysis (Cohen’s *d*, 95% CI)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Indecies | Conditions | Moral-Neutral | Moral-Immoral | Neutral-Immoral |
| RT | Self | -0.51 [-0.82, -0.20] | -0.59 [-.092, -0.26] | -0.08 [-0.21, 0.05] |
| Other | -0.09 [-0.25, 0.06] | -0.23 [-0.47, 0.01] | -0.13 [-0.29, 0.02] |
| Overall | -0.40 [-0.60, -0.20] | -0.47 [-0.66, -0.29] | 0.08 [-0.19, 0.04] |
| *d* prime | Self | 0.34 [0.09, 0.58] | 0.46 [0.19, 0.74] | 0.13 [ -0.02 0.27] |
| Other | -0.03 [-0.22, 0.16] | -0.08 [-0.18, 0.02] | 0.03 [-0.07, 0.12] |
| Overall | 0.27 [0.008 0.45] | 0.25 [0.14 0.37] | 0 [-0.14 0.14] |

**Correlation**

we correlate the normalized personal distance (self-good, self-neutral, self-bad, good-bad, good-neutral, neutral-bad) with effect of moral valence in reaction time and *d’* (the differenrces between different moral valence condition). There are 209 participants have both the behavioral data and the personal distance data. The results showed there was no significant relationship between the effect of moral valence and personal distance (*ps* > 0.0891).

**References**

Anderson, E., Siegel, E. H., Bliss-Moreau, E., & Barrett, L. F. (2011). The Visual Impact of Gossip. *Science, 332*(6036), 1446-1448. doi:10.1126/science.1201574

Aquino, K., & Reed II, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83*(6), 1423-1440. doi:10.1037/0022-3514.83.6.1423

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

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31*(3), 187-276. doi:10.1016/0010-0277(89)90023-1

Cosmides, L., & Tooby, J. (2000). The cognitive neuroscience of social reasoning. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2 ed., pp. 1259-1272). Cambridge, MA: MIT Press.

Enock, F., Sui, J., Hewstone, M., & Humphreys, G. W. (2018). Self and team prioritisation effects in perceptual matching: Evidence for a shared representation. *Acta Psychologica, 182*, 107-118. doi:10.1016/j.actpsy.2017.11.011

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

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi:10.3758/BF03193146

Fouragnan, E., Chierchia, G., Greiner, S., Neveu, R., Avesani, P., & Coricelli, G. (2013). Reputational Priors Magnify Striatal Responses to Violations of Trust. *The Journal of Neuroscience, 33*(8), 3602-3611. doi:10.1523/jneurosci.3086-12.2013

Frederick, S. (2005). Cognitive Reflection and Decision Making. *The Journal of Economic Perspectives, 19*(4), 25-42. doi:10.2307/4134953

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. doi:10.1111/spc3.12267

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*.

Hu, C.-P., Lan, Y., Macrae, C. N., & Sui, J. (in Prep). The moral me first: Prioritized response to the moral self in perceptual decision-making.

Jeffreys, H. (1961). *Theory of probability* (3 ed.). Oxford, UK: Oxford University Press.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research, 2*(1999), 102-138.

Jordan, J., Leliveld, M. C., & Tenbrunsel, A. (2015). The Moral Self-Image Scale: Measuring and Understanding the Malleability of the Moral Self. *Frontiers in Psychology, 6*. doi:10.3389/fpsyg.2015.01878

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. doi:10.3389/fpsyg.2013.00863

Liu, Q., Jian, J., Hu, C.-P., & Peng, K. (2015). *Revision of Chinese version of belief in free will and determinism.* Paper presented at the The 18th National Academic Congress of Psychology, China, Tianing, PR. China.

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. doi:10.18637/jss.v088.i02

Ostrove, J. M., Adler, N. E., Kuppermann, M., & Washington, A. E. (2000). Objective and subjective assessments of socioeconomic status and their relationship to self-rated health in an ethnically diverse sample of pregnant women. *Health Psychology, 19*(6), 613-618. doi:10.1037/0278-6133.19.6.613

Paulhus, D. L., & Carey, J. M. (2010). The FAD–Plus: Measuring Lay Beliefs Regarding Free Will and Related Constructs. *Journal of Personality Assessment, 93*(1), 96-104. doi:10.1080/00223891.2010.528483

Pool, E., Brosch, T., Delplanque, S., & Sander, D. (2016). Attentional Bias for Positive Emotional Stimuli: A Meta-Analytic Investigation. *Psychological Bulletin, 142*(1), 79-106. doi:10.1037/bul0000026

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Schönbrodt, F. D., Maier, M., Heene, M., & Zehetleitner, M. (2015). Voluntary commitment to research transparency. Retrieved from <http://www.researchtransparency.org/>

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47*(5), 609-612. doi:10.1016/j.jrp.2013.05.009

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. *SSRN eLibrary*. Retrieved from <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2160588>

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2013). Life after P-Hacking. Retrieved from <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2205186>

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. doi:10.1037/a0029792

Sui, J., Sun, Y., Peng, K., & Humphreys, G. (2014). The automatic and the expected self: separating self- and familiarity biases effects by manipulating stimulus probability. *Attention, Perception, & Psychophysics*, 1-9. doi:10.3758/s13414-014-0631-5

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48.

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. doi:10.3758%2Fs13423-017-1323-7

Wu, M. S., Schmitt, M., Zhou, C., Nartova-Bochaver, S., Astanina, N., Khachatryan, N., & Han, B. (2014). Examining Self-Advantage in the Suffering of Others: Cross-Cultural Differences in Beneficiary and Observer Justice Sensitivity Among Chinese, Germans, and Russians. *Social Justice Research, 27*(2), 231-242. doi:10.1007/s11211-014-0212-8

Wu, M. S., Yan, X., Zhou, C., Chen, Y., Li, J., Zhu, Z., . . . Han, B. (2011). General belief in a just world and resilience: Evidence from a collectivistic culture. *European Journal of Personality, 25*(6), 431-442. doi:10.1002/per.807

师保国, & 申继亮. (2007). 家庭社会经济地位、智力和内部动机与创造性的关系. *心理发展与教育*(1), 30-34.

白福宝. (2013). *不公正感体验与公正感特质对风险水平的影响.* (博士), 南京师范大学. Available from Cnki

谭永红, 丛中, & 鲁晓华. (2007). 中文厌恶感量表的初步编制及信度、效度检验. *中国心理卫生杂志, 121*(10), 696-699.

**Tables and Figures**

**Table 1**. The performance of accuracy (mean and standard deviation in brackets) as function of the moral association (good, neutral, or bad) and judgment (match vs. mismatch) in experiment 1a.

|  |  |  |
| --- | --- | --- |
| Condition | Moral association | Accuracy |
| Match | Bad person | 0.77(0.15) |
|  | Neutral person | 0.82(0.15) |
|  | Good Person | 0.87(0.13) |
| Mismatch | Bad person | 0.80(0.14) |
|  | Neutral person | 0.78(0.17) |
|  | Good Person | 0.77(0.13) |

**Table 2**. The performance of accuracy (mean and standard deviation in brackets) as function of the moral association (good, neutral, or bad) and judgment (match vs. mismatch) in Experiment 1b.

|  |  |  |
| --- | --- | --- |
| Condition | Moral association | Accuracy |
| Match | Bad person | 0.77(0.14) |
|  | Neutral person | 0.75(0.14) |
|  | Good Person | 0.87(0.08) |
| Mismatch | Bad person | 0.82(0.1) |
|  | Neutral person | 0.78(0.12) |
|  | Good Person | 0.76(0.12) |

**Table 3**. The performance of accuracy (mean and standard deviation in brackets) as function of the moral association (good, neutral, or bad) and judgment (match vs. mismatch) in Experiment 2.

|  |  |  |
| --- | --- | --- |
| Condition | Moral association | Accuracy |
| Matched | Bad person | 0.87(0.09) |
|  | Neutral person | 0.83(0.14) |
|  | Good Person | 0.93(0.09) |
| Mis-match | Bad person | 0.86(0.12) |
|  | Neutral person | 0.86(0.11) |
|  | Good Person | 0.86(0.09) |

**Table 4**. The performance of accuracy (mean and standard deviation in brackets) as function of the moral association (good, neutral, or bad) and judgment (match vs. mismatch) in Experiment 3.

|  |  |  |  |
| --- | --- | --- | --- |
| Condition | Identity | Moral valence | Accuracy |
| Matched | Self | Bad | 0.76(0.13) |
|  |  | Neutral | 0.76(0.15) |
|  |  | Good | 0.88(0.13) |
|  | Other | Bad | 0.82(0.13) |
|  |  | Neutral | 0.81(0.13) |
|  |  | Good Person | 0.84(0.14) |
| Mis-match | Self | Bad | 0.76(0.11) |
|  |  | Neutral | 0.75(0.12) |
|  |  | Good | 0.73(0.14) |
|  | Other | Bad | 0.79(0.12) |
|  |  | Neutral | 0.77(0.14) |
|  |  | Good | 0.77(0.12) |

**Table 5**. The performance of accuracy (mean and standard deviation in brackets) as function of the moral association (good, neutral, or bad) and judgment (match vs. mismatch) in Experiment 4a.

|  |  |  |  |
| --- | --- | --- | --- |
| Condition | Identity | Moral valence | Accuracy |
| Matched | Self | Bad | 0.93(0.06) |
|  |  | Neutral | 0.94(0.07) |
|  |  | Good | 0.93(0.05) |
|  | Other | Bad | 0.80(0.14) |
|  |  | Neutral | 0.77(0.11) |
|  |  | Good Person | 0.78(0.13) |
| Mis-match | Self | Bad | 0.81(0.11) |
|  |  | Neutral | 0.81(0.13) |
|  |  | Good | 0.81(0.11) |
|  | Other | Bad | 0.85(0.11) |
|  |  | Neutral | 0.85(0.10) |
|  |  | Good | 0.85(0.13) |

**Table 6**. The performance of accuracy (mean and standard deviation in brackets) as function of the moral association (good, neutral, or bad) and judgment (match vs. mismatch) in Experiment 4b.

|  |  |  |  |
| --- | --- | --- | --- |
| Condition | Identity | Moral valence | Accuracy |
| Matched | Self | Bad | 0.84(0.11) |
|  |  | Neutral | 0.84(0.10) |
|  |  | Good | 0.90(0.10) |
|  | Other | Bad | 0.84(0.12) |
|  |  | Neutral | 0.83(0.11) |
|  |  | Good Person | 0.87(0.10) |
| Mis-match | Self | Bad | 0.84(0.13) |
|  |  | Neutral | 0.86(0.11) |
|  |  | Good | 0.83(0.12) |
|  | Other | Bad | 0.86(0.11) |
|  |  | Neutral | 0.86(0.10) |
|  |  | Good | 0.85(0.10) |

**Figure Legends**

**Figure 1.** (A) d prime in pilot study as a function of moral association (good, neutral, or bad). (B) The mean performance of reaction times for match pairs in pilot study as a function of moral association (good, neutral, or bad) and match judgment (match vs. mismatch). Error bars represents one standard error.

**Figure 2.** (A) d prime in Experiment 1 as a function of moral association (good, neutral, or bad). (B) The mean performance of reaction times for match pairs in Experiment 1 as a function of moral association (good, neutral, or bad) and match judgment (match vs. mismatch). Error bars represents one standard error.

**Figure 3.** (A) d prime in Experiment 2 as a function of moral association (good, neutral, or bad). (B) The mean performance of reaction times for match pairs in Experiment 2 as a function of moral association (good, neutral, or bad) and match judgment (match vs. mismatch). Error bars represents one standard error.

**Figure 4.** (A) d prime in Experiment 3 as a function of moral association (good, neutral, or bad). (B) The mean performance of reaction times for match pairs in Experiment 3 as a function of moral association (good, neutral, or bad) and match judgment (match vs. mismatch). Error bars represents one standard error.

**Figure 5.** (A) d prime in Experiment 4 as a function of moral association (good, neutral, or bad). (B) The mean performance of reaction times for match pairs in Experiment 4 as a function of moral association (good, neutral, or bad) and match judgment (match vs. mismatch). Error bars represents one standard error.

**Figure 6.** (A) d prime in Experiment 5 as a function of moral association (good, neutral, or bad). (B) The mean performance of reaction times for match pairs in Experiment 5 as a function of moral association (good, neutral, or bad) and match judgment (match vs. mismatch). Error bars represents one standard error.

**Figure 7.** Mean psychological distance across experiments.

**Figure 8.** Correlation between RT bias (bad-person minus neutral person) and psychological distance (self- good and good-neutral) across experiments.

Appendix

**Table 1.** Mean Criterion as function of Moral valence Across experiments

|  |  |  |  |
| --- | --- | --- | --- |
| Experiment | Bad person | Good Person | Normal |
| Experiment 1 | 0.081 | -0.250 | -0.082 |
| Experiment 2 | 0.015 | -0.247 | 0.074 |

Note. Criterion = – 0.5 × (Z[Hit]+Z[False Alarm]).

**Table 10.** meta-analysis of the effect size of d’ and RTs under self condition cross 4 experiments.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | *d’* | | | RT | | |
| experiments | Bad-Neutral | Good-Neutral | Good-Bad | Bad-Neutral | Good-Neutral | Good-Bad |
| 1 | -0.012 | 0.388 | 0.408 | 0.079 | -0.532 | -0.624 |
| 2 | 0.231 | 0.5997 | 0.391 | 0.036 | -0.539 | -0.6167 |
| 3 | -0.0179 | 0.767 | 0.794 | -0.058 | -0.890 | -0.8788 |
| 4 | 0.01 | 0.049 | 0.0367 | 0.089 | -0.264 | -0.3469 |
| 5 | -0.088 | 0.3407 | 0.4158 | 0.0535 | -0.644 | -0.6786 |
| estimates | 0.025 | 0.293 | 0.366 | 0.0490 | -0.542 | -0.6076 |
| 95% CI | [-0.085 0.135] | [0.091 0.495] | [0.169 0.563] | [-0.05 0.148] | [-0.75 -0.34] | [-0.80 -0.41] |
| *p* | 0.656 | 0.0045 | 0.0003 | 0.333 | 0.000 | 0.0000 |

\*\*. Correlation is significant at the 0.01 level (2-tailed).

\*. Correlation is significant at the 0.05 level (2-tailed).

(Normal Good): The Normal condition – Good condition of matched self trials;

(Bad Normal): The Bad condition – Normal condition of matched self trials.

**Table 11.** meta-analysis of the effect size of d’ and RTs under self condition cross experiment 3 and 4.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | d’ | | |  | RTs | |
| expt | Bad-Neutral | Good-Neutral | Good-Bad | Bad-Neutral | Good-Bad | Good-Bad |
| 3 | -.017 | 0.768 | 0.794 | -0.06 | -0.894 | -0.89 |
| 4 | .01 | 0.049 | 0.037 | 0.089 | -0.264 | -0.347 |
| meta | -0.0026 | 0.39 | 0.387 | 0.034 | -0.57 | -0.5945 |
| 95% CI | [-.20 .196] | [-.31 1.09] | [-.354 1.127] | [-.109 .177] | [-1.19 0.05] | [-1.125 -0.064] |
| p | 0.9798 | 0.2778 | 0.3058 | 0.6406 | 0.0707 | 0.028 |

\*\*. Correlation is significant at the 0.01 level (2-tailed).

\*. Correlation is significant at the 0.05 level (2-tailed).

(Normal Good) : The Normal condition – Good condition of matched self trials;

(Bad Normal): The Bad condition – Normal condition of matched self trials.