1   Self-relevance modulates the priorization of the good character in perceptual matching

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

3   [1] Nanjing Normal University, 210024 Nanjing, China

4   [2] Tsinghua University, 100084 Beijing, China

5   [3] University of Aberdeen, Aberdeen, Scotland

6   Author Note

7   Hu Chuan-Peng, School of Psychology, Nanjing Normal University, 210024 Nanjing,

8   China. Kaiping Peng, Department of Psychology, Tsinghua University, 100084 Beijing,

9   China. Jie Sui, School of Psychology, University of Aberdeen, Aberdeen, Scotland. Authors

10 contriubtion: HCP, JS, & KP design the study, HCP collected the data, HCP analyzed the

11 data and drafted the manuscript. All authors read and agreed upon the current version of

12 the manuscripts.

13   Correspondence concerning this article should be addressed to Hu Chuan-Peng,

14 School of Psychology, Nanjing Normal University, Ninghai Road 122, Gulou District,

15 210024 Nanjing, China. E-mail: hcp4715@hotmail.com

Abstract

Moral character is central to social perception and moral psychology, yet how these information is process is unclear. Both negativity effect and positivity effect are possible. Here, we examined these two possibilities using social associataive learning task, where participants first associat social conceptions with neutral stimuli and then perform cognitive tasks (e.g., matching judgment). Across 9 experiments (N = 4XX, trials = XXX), we found a robust positivity effect: shapes associated with good character were prioritized, as compared to shapes associated with neutral or bad characters. Crucially, good character effect was robust when it referred to the self but weak or non-exist when it referred to a stranger. Moreover, even when identity or moral character information became task-irrelevant, participants were still sensitive to good character: the good-self combination facilitate the process while the good-other combination slowed down the process. Together, these results suggested that good character is prioritized in social associative learning task and this effect was due a spontaneous self-referential processing.

Keywords: Perceptual decision-making, Self positivity bias, moral character

Word count: X

³² Self-relevance modulates the priorization of the good character in perceptual matching

³³ Alternative title: The good person is me: Spontaneous self-referential may explain
³⁴ the prioritization of good moral character

Introduction

³⁶ [quotes about moral character]

³⁷ [Morality is central to social life, moral character is the central of morality] **People**
³⁸ **experience a substantial amount of moral events in everyday life (e.g.,**
³⁹ **Hofmann, Wisneski, Brandt, & Skitka, 2014) and judging the moral character**
⁴⁰ **of people is indispensable part of these events**. Whether we are the agent, target, or
⁴¹ a third party of a moral event, we always judge moral behaviors as "right" or "wrong," and
⁴² by doing so, we judge people as "good" or "bad" (Uhlmann, Pizarro, & Diermeier, 2015).
⁴³ Moral character is so important in social life that it is a basic dimension in our social
⁴⁴ evaluation (Goodwin, 2015; Goodwin, Piazza, & Rozin, 2014) and that a substantial part
⁴⁵ of people's conversation are gossiping others' moral character (or, reputation) (e.g.,
⁴⁶ Dunbar, 2004). These moral character information may help us to evaluate our in-group
⁴⁷ members and distinguish out-group members (Ellemers, 2018). Crucially, moral character
⁴⁸ central for self-concept and identity too. A positive moral character is viewed as the core
⁴⁹ feature of identity (e.g., Strohminger, Knobe, & Newman, 2017).

⁵⁰ [Two possibilities about moral character] Given the importance of moral character
⁵¹ and limited cognitive resources to process all the information in a social world, will people
⁵² deferentially process the information related to moral character and prioritize information
⁵³ with certain moral character? One possibility is that 'bad' character are prioritized. This
⁵⁴ possibility is consistent with early studies in impression formation which found that
⁵⁵ negative traits are weighted more in overall impression (N. H. Anderson, 1965; Fiske, 1980;
⁵⁶ Skowronski & Carlston, 1987). This idea also seemed to consistent with the more general

idea that "bad is stronger than good" (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Pratto & John, 1991). A few studies provided evidence for this possibility. For example, E. Anderson, Siegel, Bliss-Moreau, and Barrett (2011) asked participants to associate faces with different behaviors (e.g., negative and neutral behaviors from both social and nonsocial domains) and then perform a binocular rivalry task, where a face and a building were presented to each eye. Participants were required report the content of their visual awareness by pressing buttons. The results revealed that faces associated negative social behaviors dominated participants' visual awareness longer than faces associated with other types of behaviors (but see Stein, Grubb, Bertrand, Suh, & Verosky, 2017). Similarly, Eiserbeck and Abdel Rahman (2020) combined associative learning with attention blink paradigm, where neutral faces were associated with sentences about neutral or negative trust behaviors. They also found that neutral faces associated with negative behavior were processed preferentially.

It is also possible, however, that the negativity effect in impression formation may not be able to generalized to perceptual process, instead, positive moral characters are prioritized (see recent reviews, Pool, Brosch, Delplanque, & Sander, 2016; Unkelbach, Alves, & Koch, 2020). Unkelbach, Alves, and Koch (2020) pointed out that bad is not necessarily stronger than good in all aspects of information processing. Sometimes, good is stronger than bad. For example, when participants are asked to classify words as good or bad, positive trait words are classified faster than negative words (Bargh, Chaiken, Govender, & Pratto, 1992). Similarly, in a lexical decision task, participants judge positive words faster than negative words (Unkelbach et al., 2010). Also, Anisfeld and Lambert (1966) found that positive words are easier to associate with nonsense word-like strings, and this advantage in associative potential also appeared in implicit association test (IAT) (Anselmi, Vianello, & Robusto, 2011). Direct evidence for positivity effect of moral character also exist: Shore and Heerey (2013) found that faces with positive interaction in a trust game were prioritized in pre-attentive process.

⁸⁴   Here, we attempted to distinguish these two possibilities by a social associative

⁸⁵ learning task in which physical features had minimal influences — participants performed a

⁸⁶ perceptual matching task after associated different moral characters (good, neutral, and

⁸⁷ bad) with different geometric shapes. If there is a positivity effect, there should be an

⁸⁸ advantage for shapes associated with good character over shapes associated with neutral or

⁸⁹ bad shapes. If there is a negativity effect, the advantage should be occur on shapes

⁹⁰ associated with bad characters. The first four experiments and two additional follow-up

⁹¹ experiments provided strong evidence for good character effect in the current paradigm.

⁹²   The positivity effect consistent with previous studies where positivity effect of social

⁹³ trait words were found (Anselmi, Vianello, & Robusto, 2011; Bargh, Chaiken, Govender, &

⁹⁴ Pratto, 1992; Unkelbach et al., 2010). However, the effect could not be explained by the

⁹⁵ similarity hypothesis (Unkelbach, Alves, & Koch, 2020) because we only used three stimuli.

⁹⁶ There are two possibility explanations. The first one is the value-based attention account,

⁹⁷ which suggests that stimuli that are valuable to us are prioritized (B. A. Anderson, 2019).

⁹⁸ In our experiments, the good character label "good person" may represent an indirect but

⁹⁹ valuable stimuli because, in social life, a good other is usually more valuable than an bad

¹⁰⁰ other (Abele & Wojciszke, 2007). Another possibility is derived from social categorization

¹⁰¹ theory, which suggested that we automatically categorize others as in-group or out-group

¹⁰² (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). Moral character is an important

¹⁰³ criterion for social categorization (DeScioli, 2016; McHugh, McGann, Igou, & Kinsella,

¹⁰⁴ 2019). However, the above four experiments could not distinguish between these two

¹⁰⁵ possibilities, because "good person" could both be rewarding and be categorized as

¹⁰⁶ in-group member. Given that both rewarding stimuli (e.g., Sui, He, & Humphreys, 2012)

¹⁰⁷ and in-group information (Enock, Hewstone, Lockwood, & Sui, 2020) are prioritized when

¹⁰⁸ using social associative learning paradigm, we further tested these two possibilities in new

¹⁰⁹ experiments.

¹¹⁰   To distinguish the value-based account and the social categorization explanations, we

¹¹¹ introduced the identity (self- vs. other-referential) of moral character as an addition

¹¹² independent variable in exp 3a, 3b, and 6b. Now moral valence is orthogonal to the

¹¹³ identity. In this case, the identity of moral character information become salient and

¹¹⁴ participants are less likely to spontaneously categorize a good-other as an extension of self,

¹¹⁵ but the value of good-person still exists. If the positivity effect was driven by social

¹¹⁶ categorization theory, then participants prioritize good-self but not good-other. If the

¹¹⁷ value-based attention theory is true, then, both good-self and good-other are prioritized, or

¹¹⁸ maybe good-other are even more prioritized.

¹¹⁹   Although the introduction of self- and other-referential processing provided evident

¹²⁰ that value-based account can not explain the good-character effect, it might introduce the

¹²¹ good-self effect, i.e., the good-self is prioritized over all the other stimuli. This effect, if

¹²² true, may suggest underlying mechanisms other than social-categorization. For example,

¹²³ the moral true self account. Moral true self view suggested that moral self if the true self

¹²⁴ (Strohminger, Knobe, & Newman, 2017). Therefore, even good-self can be viewed as

¹²⁵ categorized to in-group, it can also be viewed as the core of the self and it is the anchor of

¹²⁶ all the other effects.

¹²⁷   To test the moral true self view and the social-categorization account, we designed

¹²⁸ two complementary experiments. In experiment 4a, participants only learned the

¹²⁹ association between self and other, the words "good-person," "neutral person," and "bad

¹³⁰ person" were presented as task-irrelevant stimuli, while in experiment 4b, participants

¹³¹ learned the associations between "good-person," "neutral-person," and "bad-person," and

¹³² the "self" and "other" were presented as task-irrelevant stimuli. These two experiments can

¹³³ be used to distinguish the moral-self view and social categorization" account. If moral-self

¹³⁴ view is true, then, in both experiments, good-self will show advantage over all other

¹³⁵ stimuli, and there will be no other effects. More specifically, in experiment 4a, where only

¹³⁶ the self-referential processing is task-relevant, there will be advantage for good as

¹³⁷ task-irrelevant condition than when bad or neutral character as task-irrelevant for the self

conditions, while there is no other effects; in experiment 4b, in the good condition, there will be an advantage for self as task-irrelevant condition over other as task-irrelevant condition, and no other effects. If social categorization is true, then, the prioritization effect will depends on whether the stimuli can be categorized as the same group of good-self. More specifically, in experiment 4a, there will be good effect in self conditions, this prediction is the same as the moral self-view; it predicts a reverse good effect in other condition because good and other a conflict in terms of social-categorization, this prediction is different from the "good-self" anchor account; however, for experiment 4b, it predicts no identity effect in the good-person condition because both self and other are in the good group.

[Good self in self-reported data] As an exploration, we also collected participants' self-reported psychological distance between self and good-person, bad-person, and neutral-person, moral identity, moral self-image, and self-esteem. All these data are available (see Liu et al., 2020). We explored the correlation between self-reported distance and these questionnaires as well as the questionnaires and behavioral data. However, given that the correlation between self-reported score and behavioral data has low correlation (Dang, King, & Inzlicht, 2020), we didn't expect a high correlation between these self-reported measures and the behavioral data.

Disclosures

We reported all the measurements, analyses, and results in all the experiments in the current study. Participants whose overall accuracy lower than 60% were excluded from analysis. Also, the accurate responses with less than 200ms reaction times were excluded from the analysis. These excluded data can be found in the shared raw data files.

All the experiments reported were not pre-registered. Most experiments (1a ~ 4b, except experiment 3b) reported in the current study were first finished between 2013 to

2016 in Tsinghua University, Beijing, China. Participants in these experiments were recruited in the local community. To increase the sample size of experiments to 50 or more (Simmons, Nelson, & Simonsohn, 2013), we recruited additional participants in Wenzhou University, Wenzhou, China in 2017 for experiment 1a, 1b, 4a, and 4b. Experiment 3b was finished in Wenzhou University in 2017. To have a better estimation of the effect size, we included the data from unreported data in our three-level models (experiment 5, 6a, 6b) (See Table S1 for overview of these experiments).

All participant received informed consent and compensated for their time. These experiments were approved by the ethic board in the Department of Psychology, Tsinghua University.

General methods

Design and Procedure

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

The experiment 1a, 1b, 1c, 2, 5, and 6a shared a 2 (matching: match vs. nonmatch) by 3 (moral character: good vs. neutral vs. bad person) within-subject design. Experiment

188 1a was the first one of the whole series studies and found the prioritization of stimuli

189 associated with good-person. To confirm that it is the moral character that caused the

190 effect, we further conducted experiment 1b, 1c, and 2. More specifically, experiment 1b

191 used different Chinese words as labels to test whether the effect only occurred with certain

192 words. Experiment 1c manipulated the moral valence indirectly: participants first learned

193 to associate different moral behaviors with different Chinese names, after remembered the

194 association, they then performed the perceptual matching task by associating names with

195 different shapes. Experiment 2 further tested whether the way we presented the stimuli

196 influence the effect of valence, by sequentially presenting labels and shapes. Note that part

197 of participants of experiment 2 were from experiment 1a because we originally planned a

198 cross task comparison. Experiment 5 was designed to compare the effect size of moral

199 character and other importance social evaluative dimensions (aesthetics and emotion).

200 Different social evaluative dimensions were implemented in different blocks, the moral

201 character blocks shared the design of experiment 1a. Experiment 6a, which shared the

202 same design as experiment 2, was an EEG experiment which aimed at exploring the neural

203 correlates of the effect. But we will focus on the behavioral results of experiment 6a in the

204 current manuscript.

205 For experiment 3a, 3b, and 6b, we included self-reference as another within-subject

206 variable in the experimental design. For example, the experiment 3a directly extend the

207 design of experiment 1a into a 2 (matching: match vs. nonmatch) by 2 (reference: self

208 vs. other) by 3 (moral character: good vs. neutral vs. bad) within-subject design. Thus in

209 experiment 3a, there were six conditions (good-self, neutral-self, bad-self, good-other,

210 neutral-other, and bad-other) and six shapes (triangle, square, circle, diamond, pentagon,

211 and trapezoids). The experiment 6b was an EEG experiment based on experiment 3a but

212 presented the label and shape sequentially. Because of the relatively high working memory

213 load (six label-shape pairs), experiment 6b were conducted in two days: the first day

214 participants finished perceptual matching task as a practice, and the second day, they

215 finished the task again while the EEG signals were recorded. We only focus on the first

216 day's data here. Experiment 3b was designed to separate the self-referential trials and

217 other-referential trials. That is, participants finished two different types of block: in the

218 self-referential blocks, they only responded to good-self, neutral-self, and bad-self, with half

219 match trials and half nonmatch trials; in the other-reference blocks, they only responded to

220 good-other, neutral-other, and bad-other.

221   Experiment 4a and 4b were design to explore the mechanism underlying the

222 prioritization of good-self. In 4a, we only used two labels (self vs. other) and two shapes

223 (circle, square). To manipulate the moral character, we added the moral-related words

224 within the shape and instructed participants to ignore the words in the shape during the

225 task. In 4b, we reversed the role of self-reference and moral character in the task:

226 participant learned three labels (good-person, neutral-person, and bad-person) and three

227 shapes (circle, square, and triangle), and the words related to identity, "self" or "other,"

228 were presented in the shapes. As in 4a, participants were told to ignore the words inside

229 the shape during the task.

230   E-prime 2.0 was used for presenting stimuli and collecting behavioral responses. For

231 participants recruited in Tsinghua University, they finished the experiment individually in

232 a dim-lighted chamber, stimuli were presented on 22-inch CRT monitors and their head

233 were fixed by a chin-rest brace. The distance between participants' eyes and the screen was

234 about 60 cm. The visual angle of geometric shapes was about $3.7° \times 3.7°$, the fixation cross

235 is of $0.8° \times 0.8°$ visual angle at the center of the screen. The words were of $3.6° \times 1.6°$ visual

236 angle. The distance between the center of the shape or the word and the fixation cross was

237 $3.5°$ of visual angle. For participants recruited in Wenzhou University, they finished the

238 experiment in a group consisted of $3 \sim 12$ participants in a dim-lighted testing room.

239 Participants were required to finished the whole experiment independently. Also, they were

240 instructed to start the experiment at the same time, so that the distraction between

241 participants were minimized. The stimuli were presented on 19-inch CRT monitor. The

242 visual angles are could not be exactly controlled because participants' chin were not fixed.

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

244 questionnaire after they finish the behavioral tasks. All the questionnaire data are open

245 (see, dataset 4 in Liu et al., 2020). See Table S1 for a summary information about all the

246 experiments.

Data analysis

248   We used the `tidyverse` of r (see script `Load_save_data.r`) to preprocess the data.

249 Results of each experiment were then analyzed using Bayesian hierarchical models.

250   We used the Bayesian hierarchical model (BHM, or Bayesian generalized linear mixed

251 models, Bayesian multilevel models) to model the reaction time and accuracy data,

252 because BHM provided three advantages over the classic NHST approach (repeated

253 measure ANOVA or t-tests): first, BHM estimate the posterior distributions of parameters

254 for statistical inference, therefore provided uncertainty in estimation (Rouder & Lu, 2005).

255 Second, BHM, where generalized linear mixed models could be easily implemented, can use

256 distributions that fit the distribution of real data instead of using normal distribution for

257 all data. Using appropriate distributions for the data will avoid misleading results and

258 provide better fitting of the data. For example, Reaction times are not normally

259 distributed but right skewed, and the linear assumption in ANOVAs is not satisfied

260 (Rousselet & Wilcox, 2019). Third, BHM provided an unified framework to analyze data

261 from different levels and different sources, avoid the information loss when we need to

262 combine data from different levels.

263   We used the `r` package `BRMs` (Bürkner, 2017), which used Stan (Carpenter et al.,

264 2017) for the BHM analyses. We estimated the over-all effect across experiments with

265 similar experimental design, instead of using a two-step approach where we first estimate

266 parameters, e.g., d' for each participant, and then use a random effect model meta-analysis

267 to synthesize the effect (Goh, Hall, & Rosenthal, 2016).

268   **Accuracy.**   We followed practice of previous studies (Hu, Lan, Macrae, & Sui, 2020;

269 Sui, He, & Humphreys, 2012) and used signal detection theory approach to analyze the

270 accuracy data. More specifically, the match trials are treated as signal and the non-match

271 trials are noise. As we mentioned above, we estimated the sensitivity and criterion of SDT

272 by BHM (Rouder & Lu, 2005). Because the BHM can model different level's data using a

273 single unified model, we used a three-level HBM to model the moral character effect, which

274 include five experiments: 1a, 1b, 1c, 2, 5, and 6a. Similarly, we modeled experiments with

275 both self-referential and moral character with a three-level HBM model, which includes 3a,

276 3b, and 6b. For experiment 4a and 4b, we used two-level models for each separately.

277 However, we could compare the posterior of parameters directly because we have full

278 posterior distribution of parameters.

279   We used the Bernoulli distribution to model the accuracy data. For a single

280 participant, we assume that the accuracy of ith trial is Bernoulli distributed (binomial with

281 1 trial), with probability p_i that $y_i = 1$.

$$y_i \sim Bernoulli(p_i)$$

282 and the probability of choosing "match" p_i at the ith trial is a function of the trial type:

$$\Phi(p_i) = \beta_0 + \beta_1 IsMatch_i$$

283 therefore, the outcomes y_i are 0 if the participant responded "nonmatch" on the ith trial, 1

284 if they responded "match." We then write the generalized linear model on the probits

285 (z-scores; Φ, "Phi") of ps. Φ is the cumulative normal density function and maps z scores

286 to probabilities. In this way, the intercept of the model (β_0) is the standardized false alarm

287 rate (probability of saying 1 when predictor is 0), which we take as our criterion c. The

288 slope of the model (β_1) is the increased probability of responding "match" when the trial

289 type is "match," in z-scores, which is another expression of d'. Therefore, $c = $ -zHR $=$

290 $-\beta_0$, and $d' = \beta_1$.

291   In our experimental design, there are three conditions for both match and non-match

292 trials, we can estimate the d' and c separately for each condition. In this case, the criterion

293 c is modeled as the main effect of valence, and the d' can be modeled as the interaction

294 between valence and match:

$$\Phi(p_i) = 0 + \beta_0 Valence_i + \beta_1 IsMatch_i * Valence_i$$

295   In each experiment, we had multiple participants. We can estimate the group-level

296 parameters by extending the above model into a two-level model, where we can estimate

297 parameters on individual level (varying effect) and the group level parameter

298 simultaneously (fixed effect). The probability that the jth subject responded "match"

299 $(y_{ij} = 1)$ at the ith trial p_{ij}. In the same vein, we have

$$y_{ij} \sim Bernoulli(p_{ij})$$

300 The the generalized linear model can be re-written to include two levels:

$$\Phi(p_{ij}) = 0 + \beta_{0j} Valence_{ij} + \beta_{1j} IsMatch_{ij} * Valence_{ij}$$

301 We again can write the generalized linear model on the probits (z-scores; Φ, "Phi") of ps.

302   The subjective-specific intercepts $(\beta_0 = -zFAR)$ and slopes $(\beta_1 = d')$ are describe

303 by multivariate normal with means and a covariance matrix for the parameters.

$$\begin{bmatrix} \beta_{0j} \\ \beta_{1j} \end{bmatrix} \sim N(\begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}, \sum)$$

304   For experiments that had 2 (matching: match vs. non-match) by 3 (moral character:

305 good vs. neutral vs. bad), i.e., experiment 1a, 1b, 1c, 2, 5, and 6a, the formula for accuracy

306 in BRMs is as follow:

```
saymatch ~ 0 + Valence + Valence:ismatch + (0 + Valence +
Valence:ismatch | Subject), family = bernoulli(link="probit")
```

For experiments that had two by two by three design, we used the follow formula for the BGLM:

```
saymatch ~ 0 + ID:Valence + ID:Valence:ismatch + (0 + ID:Valence +
ID:Valence:ismatch | Subject), family = bernoulli(link="probit")
```

In the same vein, we can estimate the posterior of parameters across different experiments. We can use a nested hierarchical model to model all the experiment with similar design:

$$y_{ijk} \sim Bernoulli(p_{ijk})$$

the generalized linear model is then

$$\Phi(p_{ijk}) = 0 + \beta_{0jk}Valence_{ijk} + \beta_{1j}IsMatch_{ijk} * Valence_{ijk}$$

The outcomes y_{ijk} are 0 if participant j in experiment k responded "nonmatch" on trial i, 1 if they responded "match."

$$\begin{bmatrix} \beta_{0jk} \\ \beta_{1jk} \end{bmatrix} \sim N(\begin{bmatrix} \theta_{0k} \\ \theta_{1k} \end{bmatrix}, \sum)$$

and the experiment level parameter mu_{0k} and mu_{1k} is from a higher order distribution:

$$\begin{bmatrix} \theta_{0k} \\ \theta_{1k} \end{bmatrix} \sim N(\begin{bmatrix} \mu_0 \\ \mu_1 \end{bmatrix}, \sum)$$

in which mu_0 and mu_1 means the population level parameter.

Reaction times. For the reaction time, we used the log normal distribution (https://lindeloev.github.io/shiny-rt/#34_(shifted)_log-normal) to model the data. This

324   means that we need to estimate the posterior of two parameters: μ, σ. μ is the mean of the

325   `logNormal` distribution, and σ is the disperse of the distribution. Although the log normal

326   distribution can be extended to shifted log normal distribution, with one more parameter:

327   shift, which is the earliest possible response, we found that the additional parameter didnt'

328   improved the model fitting and therefore used the logNormal in our final analysis.

329   The reaction time of the jth subject on ith trial is a linear function of trial type:

$$y_{ij} = \beta_{0j} + \beta_{1j} * IsMatch_{ij} * Valence_{ij}$$

330   while the log of the reaction time is log-normal distributed:

$$log(y_{ij}) \sim N(\mu_j, \sigma_j)$$

331   y_{ij} is the RT of the ith trial of the jth participants.

$$\mu_j \sim N(\mu, \sigma)$$

$$\sigma_j \sim Cauchy()$$

332   Formula used for modeling the data as follow:

333   RT_sec ~ Valence*ismatch + (Valence*ismatch | Subject), family =

334   lognormal()

335   or

336   RT_sec ~ ID*Valence*ismatch + (ID*Valence*ismatch | Subject), family =

337   lognormal()

338   we expanded the RT model three-level model in which participants and experiments

339   are two group level variable and participants were nested in the experiments.

$$log(y_{ijk}) \sim N(\mu_{jk}, \sigma_{jk})$$

y_{ijk} is the RT of the ith trial of the jth participants in the kth experiment.

$$\mu_{jk} \sim N(\mu_k, \sigma_k)$$

$$\sigma_{jk} \sim Cauchy()$$

$$\mu_k \sim N(\mu, \sigma)$$

$$\theta_k \sim Cauchy()$$

Effect of moral character.   We estimated the effect size of d' and RT from experiment 1a, 1b, 1c, 2, 5, and 6a for the effect of moral character. We reported fixed effect of three-level BHM that included all experiments that tested the valence effect.

Interaction between moral character and self-referential process.   We also estimated the interaction between moral character and self-referential process, which included results from experiment 3a, 3b, and 6b. Using three-level models, we tested two possible explanations for the prioritization of good character: value-based or social categorization based prioritization.

Implicit interaction between valence and self-relevance.   In the third part, we focused on experiment 4a and 4b, which were designed to examine two more nuanced explanation concerning the good-self. The design of experiment 4a and 4b are complementary. Together, they can test whether participants are more sensitive to the moral character of the Self (4a), or the identity of the good character (4b).

For the questionnaire part, we are most interested in the self-rated distance between different person and self-evaluation related questionnaires: self-esteem, moral-self identity, and moral self-image. Other questionnaires (e.g., personality) were not planned to

correlated with behavioral data were not included. Note that all questionnaire data were
reported in (Liu et al., 2020).

<div align="center">Results</div>

Perceptual processing moral character related information

In this part, we report results from five experiments that tested whether an
associative learning task, including 192 participants. Note that for both experiment 1a and
1b, there were two independent samples with different equipment, trials numbers and
testing situations. Therefore, we modeled them as independent samples. These five
experiments revealed a robust effect of moral character on perceptual matching task.

For the d prime, we found robust effect of moral character. Shapes associated with
good character ("good person," "kind person" or a name associated with morally good
behavioral history) has higher sensitivity (median = 2.49, 95% HDI = [2.19 2.75]) than
shapes associated with neutral character (median = 2.18, 95% HDI = [1.90 2.48]),
$median_{diff}$ = 0.31, 95% HDI [0.02 0.63] , but we did not find differences between shapes
associated with bad character (median = 2.23, 95% HDI = [1.94 2.53]) and neutral
character, $median_{diff}$ = 0.05, 95% HDI [-0.29 0.37].

For the reaction times, we also found robust effect of moral character for both match
trials (see figure 1 C) and nonmatch trials (**see supplementary materials**). For match
trials, shapes associated with good character has faster responses (median = 578.64 ms,
95% HDI = [508.15 661.14]) than shapes associated with neutral character (median =
623.45 ms, 95% HDI = [547.98 708.24]), $median_{diff}$ = -44.05, 95% HDI [-59.96 -30.43].
We also found that the responses to shapes associated with bad character (median =
640.41 ms, 95% HDI = [559.94 719.63]) were slower as compared to the neutral character,
$median_{diff}$ = 17.04, 95% HDI [4.02 29.92]. See Figure 1.

For the nonmatch trials, we also found the advantage of good character: Shapes

385 associated with good character (median = 653.21 ms, 95% HDI = [574.65 739.57]) are

386 faster than shapes associated with neutral (median = 671.14 ms, 95% HDI = [591.71

387 760.09]), $median_{diff}$ = -17.65 ms, 95% HDI [-23.85 -10.36]. Similarly, the shapes

388 associated with bad character (median = 676.35 ms, 95% HDI = [599.13 767.76]) was

389 responded slower than shapes associated with neutral character, $median_{diff}$ = 17.04 ms,

390 95% HDI [4.02 29.92], but the effect size was smaller, (**see supplementary materials**).

391 **Self-referential process modulate prioritization of good character**

392   In this part, we report results from three experiments (3a, 3b, and 6b) that aimed at

393 testing whether the moral valence effect found in the previous experiments is modulated by

394 self-referential processes. These three experiments included data from 108 participants.

395   Because we have found that a facilitation effect of good character and slow-down

396 effect of bad character in the first part, in this part, we will focus on the whether such

397 effect interact with self-referential factor. In others words, we not only reported differences

398 between good/bad character with neutral character for self-referential and other-referential

399 separately, but also compare the differences between the difference.

400   For the d prime, we found that an interaction between moral character effect and

401 self-referential, the self- and other-referential difference was greater than zero for good

402 vs. neutral character differences ($median_{diff}$ = 0.51; 95% HDI = [-1.48 2.61]) but not for

403 bad vs. neutral differences ($median_{diff}$ = -0.02; 95% HDI = [-1.85 2.17]). Further analyses

404 revealed that the good vs. neutral character effect only appeared for self-referential

405 conditions but not other-referential conditions. The estimated d prime for good-self was

406 greater than neutral-self ($median_{diff}$ = 0.56; 95% HDI = [-1.05 2.15]), d prime for

407 good-self was also greater than good-other condition ($median_{diff}$ = ; 95% HDI = [ ]). The

408 differences between bad-self and neutral-self, good-other and neutral-other, bad-other and

409 neutral-other are all centered around zero (see Figure 2, B, D).

410   For the RTs part, we also found the interaction between moral character and

411 self-referential, the self- and other-referential differences was below zero for the good

412 vs. neutral differences ($median_{diff}$ = -105.39; 95% HDI = [-533.16 281.69]) but not for the

413 bad vs. neutral differences ($median_{diff}$ = -9.46; 95% HDI = [-290.72 251.38]). Further

414 analyses revealed a robust good-self prioritization effect as compared to neutral-self

415 ($median_{diff}$ = -47.58; 95% HDI = [-202.88 16.83]) and good-other ($median_{diff}$ = -57.14;

416 95% HDI = [-991.89 621.29]) conditions. Also, we found that both good character and bad

417 character were responded slower than neutral character when it was other-referential. See

418 Figure 2.

Binding the good and self

420   In this part, we reported two studies in which the moral valence or the self-referential

421 processing is not task-relevant. We are interested in testing whether the task-relevance

422 modulated the effect observed in previous experiment.

423   In experiment 4a, where self- and other-referential were task-relevant and moral

424 character are task-irrelevant. We found self-related conditions were performed better than

425 other-related conditions, on both d prime and reaction times. This pattern is consistent

426 with previous studies (e.g., Sui, He, and Humphreys (2012)).

427   More importantly, we found evidence, albeit weak, that task-irrelevant moral

428 character also played an role. For shapes associated with self, d' was greater when shapes

429 had a good character inside the shape (median = 2.83, 95% HDI [2.63 3.01]) than shapes

430 that have neutral character (median = 2.74, 95% HDI [2.58 2.95], BF = 4.4) or bad

431 character (median = 2.76, 95% HDI [2.56 2.95], 3.1), but we did not found difference

432 between shapes with bad character and neutral character inside for the self-referential

433 shapes. For shapes associated with other, the results of d' revealed a reversed pattern to

434 the self-referential condition: d prime was smaller when shapes had a good character inside

435 (median = 1.87, 95% HDI [1.71 2.04]) than had neutral (median = 1.96, 95% HDI [1.80

436 2.14]) or bad character (median = 1.98, 95% HDI [1.79 2.17]) inside. See Figure 3.

437   The same pattern was found for RTs. For self-referential condition, when good

438 character was presented as a task-irrelevant stimuli, the responds (median = 641, 95% HDI

439 [623 662]) were faster than when neutral character (median = 649, 95% HDI [631 668]) or

440 bad character (median = 648, 95% HDI [628 667]) were inside. This effect was reversed for

441 other-referential condition: shapes associated with other with good character inside

442 (median = 733, 95% HDI [711 754]) were slower than with neutral character (median =

443 721, 95% HDI [702 741]) or bad character (median = 718, 95% HDI [696 740]) inside.

444   In experiment 4b, moral character was the task-relevant factor, and we found that

445 there were main effect of moral character: shapes associated with good character were

446 performed better than other-related conditions, on both d' and reaction times.

447   Most importantly, we found evidence that task-irrelevant self-referential process also

448 played an role. For shapes associated with good person, the d prime was greater when

449 shapes had an "self" inside than with "other" inside ($mean_{diff} = 0.14$, 95% credible

450 intervals [-0.02, 0.31], BF = 12.07, p = 0.92), but this effect did not happen when the

451 target shape where associated with "neutral" ($mean_{diff} = 0.04$, 95% CI [-.11, .18]) or

452 "bad" person ($mean_{diff} = -.05$, 95% CI[-.18, .09]).

453   The same trend appeared for the RT data. For shapes associated with good person,

454 with a "self" inside the shape reduced the reaction times as compared with when a "other"

455 inside the shape ($mean_{diff} = -55$ ms, 95%CI[-75, -35]), but this effect did not occur when

456 the shapes were associated neutral ($mean_{diff} = 10$, 95% CI [1, 20]) or bad ($mean_{diff} = 5$,

457 95%CI [-16, 27]) person. See Figure 3.

Self-reported personal distance

⁴⁵⁹   We explored the self-reported psychological distance between different person.

⁴⁶⁰ Participants were presented a pair of two-person each time, and moved a slide to represent

⁴⁶¹ the distance between the pair of two persons. We found that, on average, participants rated

⁴⁶² self is closest to a neutral person, and then a good person. These two are not different from

⁴⁶³ each other. However, both are closer than the distance between good person and neutral

⁴⁶⁴ person. On average, participants rated themselves has furthest distance to bad person.

⁴⁶⁵   Correlation analysis showed that most psychological distance ratings were positively

⁴⁶⁶ correlated to each other, but the self-bad and self-good are negatively correlated.

⁴⁶⁷   [use the network view to visualize the distance]

⁴⁶⁸   See Figure 4 and Figure 5.

Discussion

⁴⁷⁰   We human inevitably view other people and ourselves in a moral lens. Yes how this

⁴⁷¹ moral lens will change our information processes is unknown. Across nine experiments, we

⁴⁷² studied the processing of moral character using a social associative learning task, we

⁴⁷³ examined the effect of moral character on a matching task and explored the mechanisms

⁴⁷⁴ underlying the effect. We found robust evidence that good character are prioritized in the

⁴⁷⁵ matching task, regardless of the label used for moral characters or the way stimuli are

⁴⁷⁶ presented. We documented that this positivity effect was driven by a self-referential

⁴⁷⁷ processing: prioritization only occur when moral characters are self-referential but not

⁴⁷⁸ other-referential. The prioritization effect occur when self and good character are

⁴⁷⁹ combined, whether task-relevant or not. When good character were other-referential, even

⁴⁸⁰ implicitly, the information process might be slowed down. Together, our findings highlight

⁴⁸¹ the importance of the self-referential in perceiving positive moral character. These findings

⁴⁸² contributed to a growing literature on the social nature of perception (Freeman, Stolier, &

483 Brooks, 2020; Xiao, Coppin, & Bavel, 2016) by supporting the idea that people can

484 prioritized not just physically salient, or affective stimuli, but also socially salient stimuli,

485 i.e., instantly acquired moral information.

486   First, we examined the perceptual process of moral character to understand how

487 moral information are processed. Prior research has demonstrated that bad moral behavior

488 is stronger in impression formation [] and bad moral character are attended quicker than

489 neutral moral character []. The empirical studies on the moral character often focus on

490 self-reported data rather than behavioral response. In this paper, we examined the

491 perceptual processes of moral character. In doing so, we shifted the focus from

492 consequences of information process of moral character to the information process itself,

493 thus broaden the scope of the existing research.

494   Second, perceptual processing is the upstream of our information that can help us

495 understand priority of different information. We thus contributed to the research on moral

496 character by demonstrating that information related to good character in general, and

497 good moral self in specific, is prioritized. Specifically, we found that instantly learned

498 moral character information can change the information process of neutral, non-social

499 information. Presumably this prioritization occurs because good moral character and

500 moral self is central to one's social life. Research has found that moral self is essential for

501 one's identity and people has stronger self-enhancement effect in moral domain than in

502 other social domain. This positivity effect is opposite to negativity effect in impression

503 formation, suggesting that impression formation and perceptual-matching may involved

504 different information process mechanisms. This positivity effect, though surprising at first,

505 is well supported by previous studies. Positivity effect had been found in associative

506 learning [], lexical decision-making [], and IAT []. A common feature of these paradigms is

507 that decision-making occurs at relative later stage of the information processing in

508 perception, instead of early sensory processing stage. In the current paradigm, participants

509 made a matching judgment, which was only possible after participants formed a perception

of both the shape and the label and retrieve the association between them. ample evidence

supported the idea that positive stimuli have advantage at the later stage of perception

(Pool, Brosch, Delplanque, & Sander, 2016). The task used in the current study may

explain why the result are different from previous studies such as E. Anderson, Siegel,

Bliss-Moreau, and Barrett (2011) and Eiserbeck and Abdel Rahman (2020), where the

early processing stage were targeted by attention blink paradigm.

The absence of negativiity effect may also caused by the fact that the bad character

here is an abstract concept that may not bring concrete threatening to the participants,

therefore it is not as strong as previous studies used emotional stimuli that has higher

arousal. Besides, recent study found that when the moral violation is not life-threatening,

the impression of bad character is volatile in the social context [].

Third, knowing that good character and moral-self is prioritized is not sufficient; we

need to know why this prioritization occurs. Our results indicate that the good character

prioritization is driven by spontaneous self-referential processing. Also, these results

revealed that either a general-self based social categorization or moral self as anchor view

alone can explain the results. Instead, we proposed that moral-self based social

categorization can better account for the results, especially the results where either identity

or moral character information were task-irrelevant. These results echo prior research on

moral-self view, suggesting that moral-self as true self is not only at self-report level but

also at perceptual level. Further, our results showed that we not only regard moral self as

the true-self, but also seek to categorize information based on moral-self: when good-other

creates an ambiguous situation, the responses was slowed down in perceptual processing.

Fourth, we find that behavioral data and self-reported data doesn't congruent. When

asked to rate the distance between self and good person, the distance is similar to the

distance between self and neutral person. However, the distance between self and bad

person is the longest, even longer than the distance between good and bad. These results

might be caused by the social desirability effect that often occurs in self-reporting. However, we didn't not find strong evidence for the correlation between behavioral results and the self-reported person distance.

[Memory or perception.] One would argue that the effect here may represent a memory effect instead of perpetual effect *per se.* (1) how to define perception is debated, while some researchers included memory components in perception, others do not. Here, we are more on the broader view of perception. (2), the memory effect view predict that the effect will be eliminated after participants became familiar with the association. We did supplementary analysis where we divided the whole experiment into three different stage: early, middle, and later, and then compared the results pattern of early and later stages. These results revealed null effect of training. These additional analysis suggested that memory effect may exist, but in a sense that they reflected a long-term, stable pattern of different valenced moral character, instead of a short-term, associative learning induced effect.

[free association from small world of words]

References

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93*(5), 751–763. https://doi.org/10.1037/0022-3514.93.5.751

Anderson, B. A. (2019). Neurobiology of value-driven attention. *Current Opinion in Psychology, 29*, 27–33. https://doi.org/10.1016/j.copsyc.2018.11.004

Anderson, E., Siegel, E. H., Bliss-Moreau, E., & Barrett, L. F. (2011). The visual impact of gossip. *Science, 332*(6036), 1446–1448. https://doi.org/10.1126/science.1201574

Anderson, N. H. (1965). Averaging versus adding as a stimulus-combination rule in

561   impression formation. *Journal of Experimental Psychology, 70*(4), 394–400.

562   https://doi.org/10.1037/h0022280

563   Anisfeld, M., & Lambert, W. E. (1966). When are pleasant words learned faster

564   than unpleasant words? *Journal of Verbal Learning and Verbal Behavior, 5*(2),

565   132–141. https://doi.org/10.1016/S0022-5371(66)80006-3

566   Anselmi, P., Vianello, M., & Robusto, E. (2011). Positive associations primacy in

567   the IAT. *Experimental Psychology.* Retrieved from

568   https://econtent.hogrefe.com/doi/abs/10.1027/1618-3169/a000106

569   Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the

570   automatic attitude activation effect. *Journal of Personality and Social*

571   *Psychology, 62*(6), 893–912. https://doi.org/10.1037/0022-3514.62.6.893

572   Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is

573   stronger than good. *Review of General Psychology, 5*(4), 323–370.

574   https://doi.org/10.1037/1089-2680.5.4.323

575   Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is

576   stronger than good. *Review of General Psychology, 5*(4), 323–370.

577   https://doi.org/10.1037/1089-2680.5.4.323

578   Bürkner, P.-C. (2017). Brms: An r package for bayesian multilevel models using

579   stan [Journal Article]. *Journal of Statistical Software; Vol 1, Issue 1 (2017).*

580   Retrieved from

581   https://www.jstatsoft.org/v080/i01%20http://dx.doi.org/10.18637/jss.v080.i01

582   Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M.,

583   … Riddell, A. (2017). Stan: A probabilistic programming language [Journal

584   Article]. *Journal of Statistical Software, 76*(1).

585   https://doi.org/10.18637/jss.v076.i01

Dang, J., King, K. M., & Inzlicht, M. (2020). Why are self-report and behavioral measures weakly correlated? *Trends in Cognitive Sciences*, *24*(4), 267–269. https://doi.org/10.1016/j.tics.2020.01.007

DeScioli, P. (2016). The side-taking hypothesis for moral judgment. *Current Opinion in Psychology*, *7*, 23–27. https://doi.org/10.1016/j.copsyc.2015.07.002

Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of General Psychology*, *8*(2), 100–110. https://doi.org/10.1037/1089-2680.8.2.100

Eiserbeck, A., & Abdel Rahman, R. (2020). Visual consciousness of faces in the attentional blink: Knowledge-based effects of trustworthiness dominate over appearance-based impressions. *Consciousness and Cognition*, *83*, 102977. https://doi.org/10.1016/j.concog.2020.102977

Ellemers, N. (2018). Morality and social identity. In M. van Zomeren & J. F. Dovidio (Eds.), *The oxford handbook of the human essence* (pp. 147–158). New York, NY, US: Oxford University Press.

Enock, F. E., Hewstone, M. R. C., Lockwood, P. L., & Sui, J. (2020). Overlap in processing advantages for minimal ingroups and the self. *Scientific Reports*, *10*(1), 18933. https://doi.org/10.1038/s41598-020-76001-9

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, *38*(6), 889–906. https://doi.org/10.1037/0022-3514.38.6.889

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, *38*(6), 889–906. https://doi.org/10.1037/0022-3514.38.6.889

Freeman, J. B., Stolier, R. M., & Brooks, J. A. (2020). Chapter five - dynamic interactive theory as a domain-general account of social perception. In B.

Gawronski (Ed.), *Advances in experimental social psychology* (Vol. 61, pp. 237–287). Academic Press. https://doi.org/10.1016/bs.aesp.2019.09.005

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 [Journal Article]. *Social and Personality Psychology Compass*, *10*(10), 535–549. https://doi.org/10.1111/spc3.12267

Goodwin, G. P. (2015). Moral character in person perception. *Current Directions in Psychological Science*, *24*(1), 38–44. https://doi.org/10.1177/0963721414550709

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*(1), 148–168. https://doi.org/10.1037/a0034726

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science*, *345*(6202), 1340–1343. https://doi.org/10.1126/science.1251560

Hu, C.-P., Lan, Y., Macrae, C. N., & Sui, J. (2020). Good me bad me: Does valence influence self-prioritization during perceptual decision-making? [Journal Article]. *Collabra: Psychology*, *6*(1), 20. https://doi.org/10.1525/collabra.301

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

McHugh, C., McGann, M., Igou, E. R., & Kinsella, E. (2019). *Moral judgment as categorization (MJAC)*. PsyArXiv. https://doi.org/10.31234/osf.io/72dzp

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. https://doi.org/10.1037/bul0000026

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, *61*(3), 380–391. https://doi.org/10.1037//0022-3514.61.3.380

Rouder, J. N., & Lu, J. (2005). An introduction to bayesian hierarchical models with an application in the theory of signal detection [Journal Article]. *Psychonomic Bulletin & Review*, *12*(4), 573–604. https://doi.org/10.3758/bf03196750

Rousselet, G. A., & Wilcox, R. R. (2019). Reaction times and other skewed distributions: Problems with the mean and the median [Preprint]. *Meta-Psychology.* https://doi.org/10.1101/383935

Shore, D. M., & Heerey, E. A. (2013). Do social utility judgments influence attentional processing? *Cognition*, *129*(1), 114–122. https://doi.org/10.1016/j.cognition.2013.06.011

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

Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology*, *52*(4), 689–699. https://doi.org/10.1037/0022-3514.52.4.689

Stein, T., Grubb, C., Bertrand, M., Suh, S. M., & Verosky, S. C. (2017). No impact of affective person knowledge on visual awareness: Evidence from binocular rivalry and continuous flash suppression. *Emotion*, *17*(8), 1199–1207. https://doi.org/10.1037/emo0000305

Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological

concept distinct from the self: *Perspectives on Psychological Science.*

https://doi.org/10.1177/1745691616689495

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

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory.* Cambridge, MA, US: Basil Blackwell.

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment: *Perspectives on Psychological Science.* https://doi.org/10.1177/1745691614556679

Unkelbach, C., Alves, H., & Koch, A. (2020). Chapter three - negativity bias, positivity bias, and valence asymmetries: Explaining the differential processing of positive and negative information. In B. Gawronski (Ed.), *Advances in experimental social psychology* (Vol. 62, pp. 115–187). Academic Press. https://doi.org/10.1016/bs.aesp.2020.04.005

Unkelbach, C., Hippel, W. von, Forgas, J. P., Robinson, M. D., Shakarchi, R. J., & Hawkins, C. (2010). Good things come easy: Subjective exposure frequency and the faster processing of positive information. *Social Cognition*, *28*(4), 538–555. https://doi.org/10.1521/soco.2010.28.4.538

Xiao, Y. J., Coppin, G., & Bavel, J. J. V. (2016). Perceiving the world through group-colored glasses: A perceptual model of intergroup relations. *Psychological Inquiry*, *27*(4), 255–274. https://doi.org/10.1080/1047840X.2016.1199221
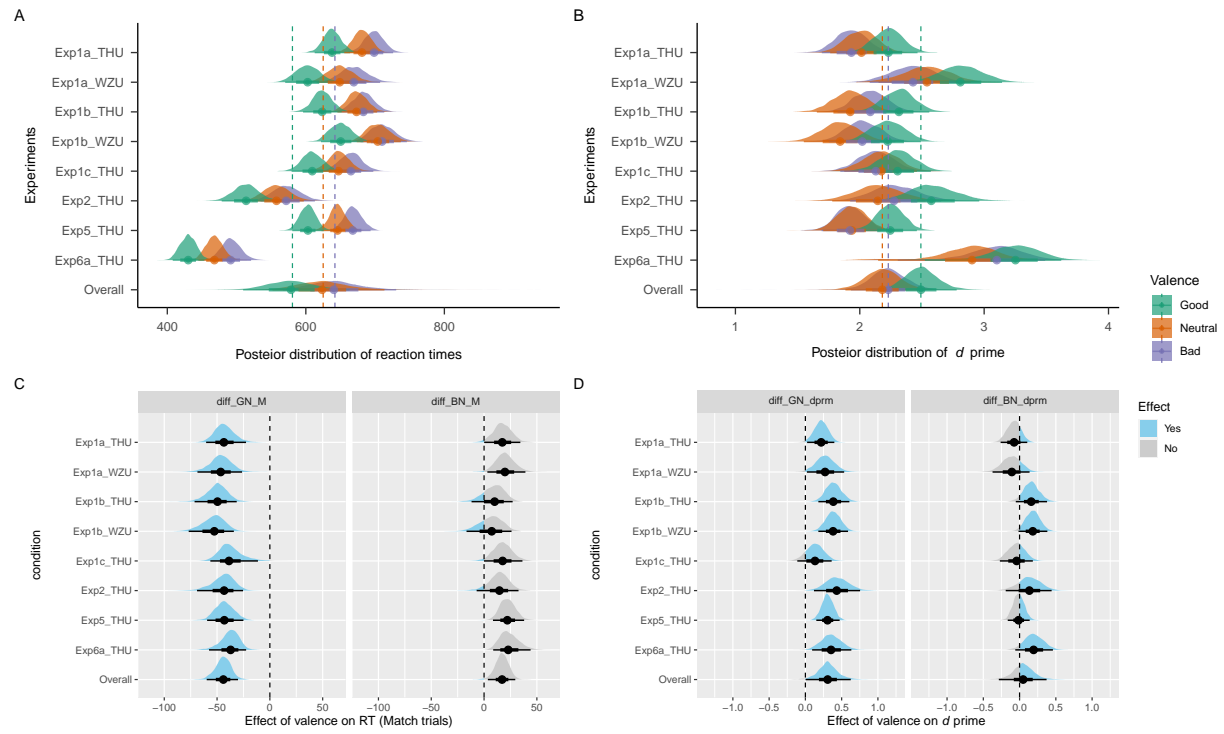
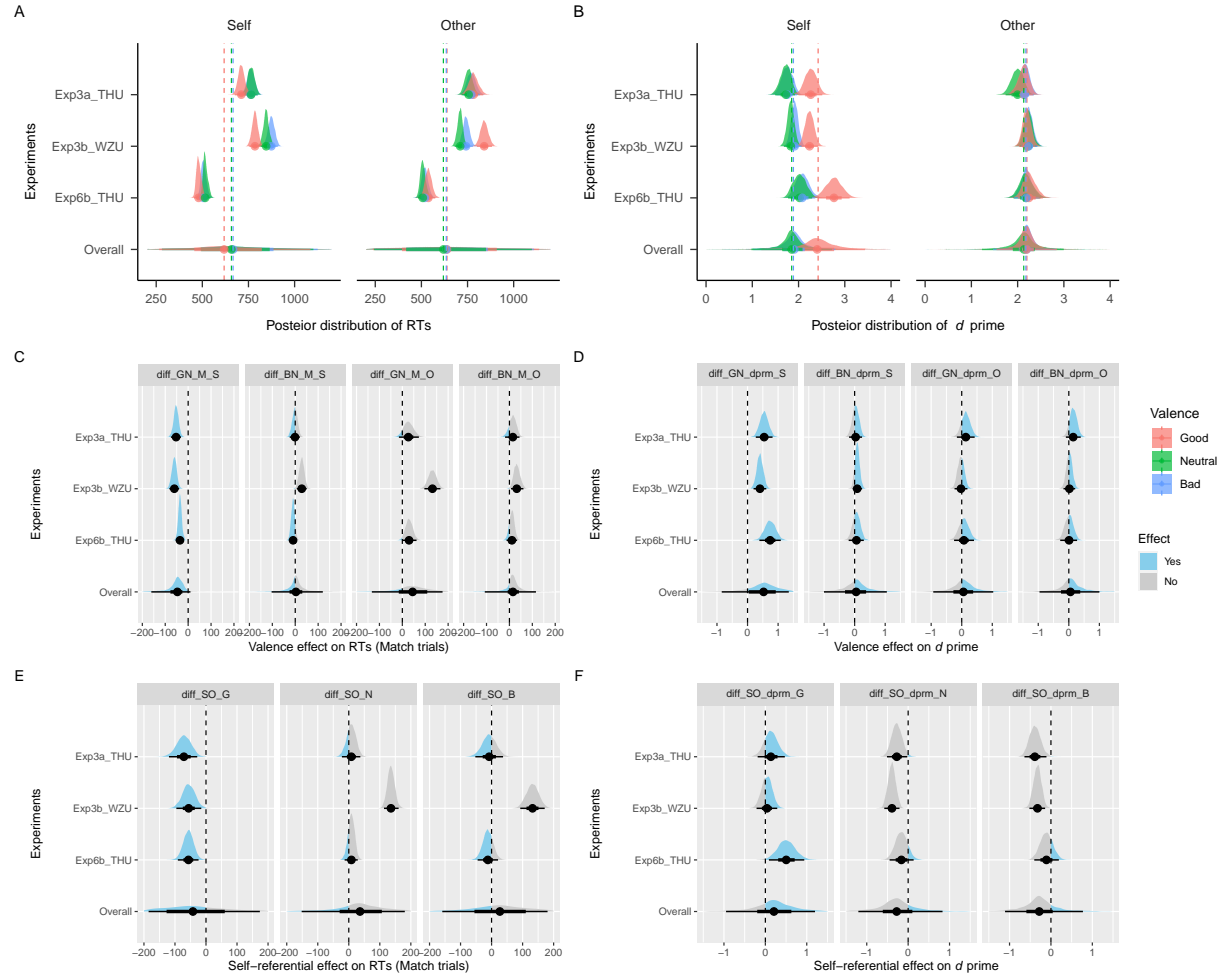*Figure 1.* Effect of moral valence on RT and d'

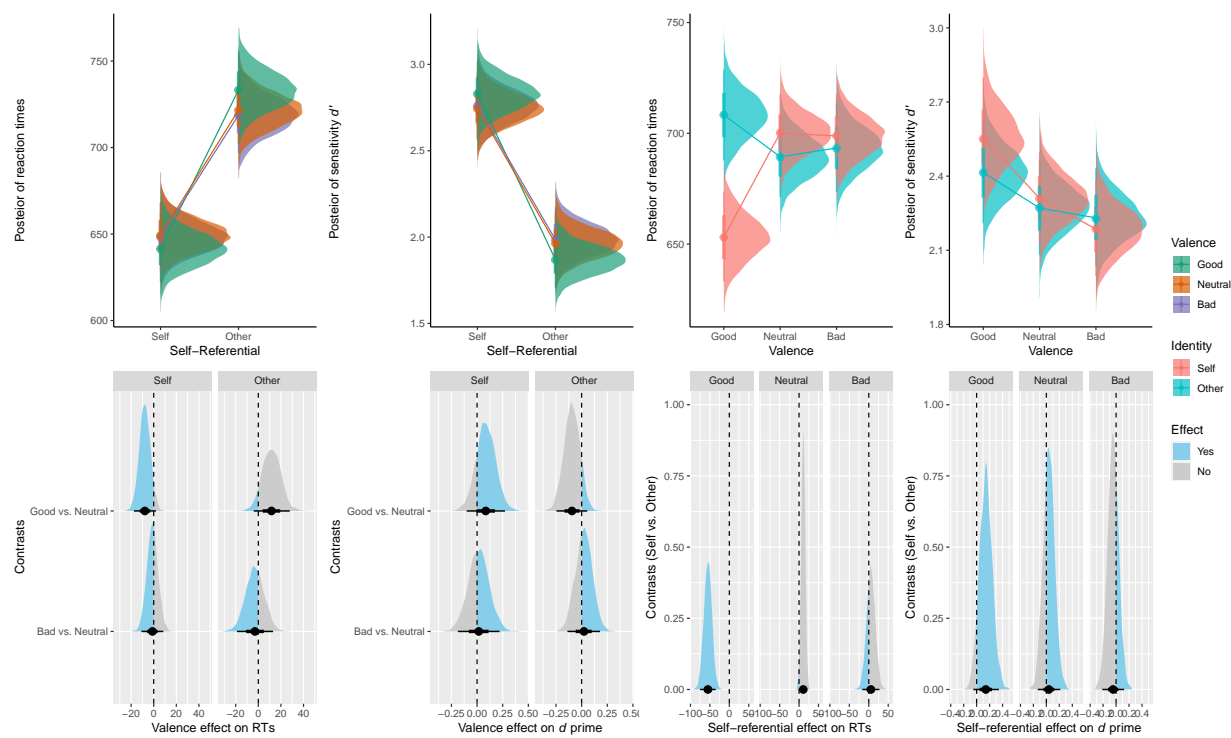*Figure 2.* Interaction between moral valence and self-referential

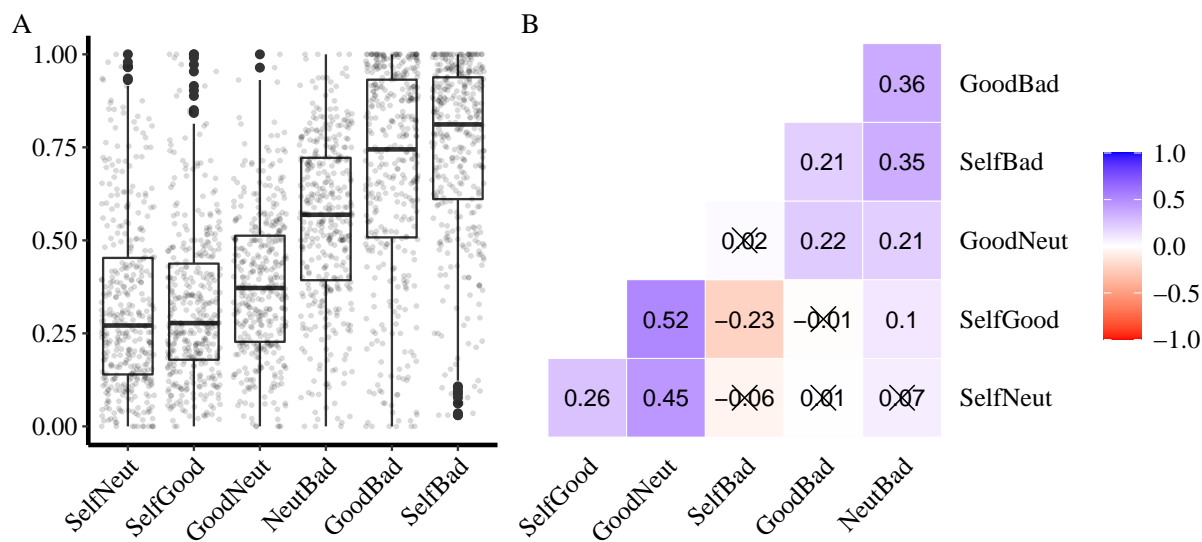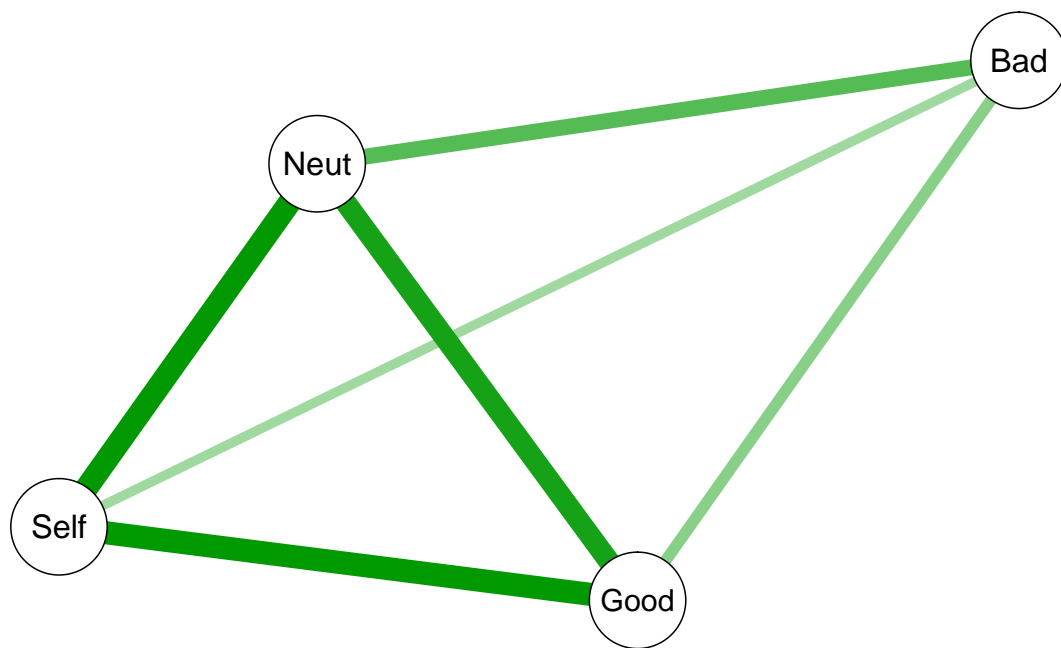*Figure 3.* exp4: Results of Bayesian GLM analysis.

Figure 4. Self-rated personal distance

Figure 5. Self-rated personal distance (Network view)