[Abstract 1](#_Toc134275420)

[Introduction 2](#_Toc134275421)

[Methods 3](#_Toc134275422)

[Ethics information 3](#_Toc134275423)

[Datasets 3](#_Toc134275424)

[Simulated data 9](#_Toc134275425)

[Analysis Plan 10](#_Toc134275426)

[Data pre-processing 11](#_Toc134275427)

[Calculating the SPE 11](#_Toc134275428)

[Estimating the Reliability 12](#_Toc134275429)

[Data availability 13](#_Toc134275430)

[Code availability 13](#_Toc134275431)

[Results 13](#_Toc134275432)

[Split-Half Reliability 13](#_Toc134275433)

[Intraclass correlation coefficient (ICC) 14](#_Toc134275434)

[Discussion 16](#_Toc134275435)

[Acknowledgements 17](#_Toc134275436)

[Author contributions 17](#_Toc134275437)

[Competing interests 17](#_Toc134275438)

[Supplementary information 18](#_Toc134275439)

[References 19](#_Toc134275440)

**Reliability of Self-Prioritization Effect as Measured by the Perceptual Matching Task: Evidence from Multiple Datasets**

Zheng Liu 1#, Mengzhen Hu2#, Yuanrui Zheng3, Jie Sui4, Hu Chuan-Peng 2\*

1 School of Humanities and Social Sciences, The Chinese University of Hong Kong - Shenzhen, Shenzhen, China

2 School of Psychology, Nanjing Normal University, Nanjing, China

3 School of Education, Kunming City College, Kunming, China

4 School of Psychology, University of Aberdeen, Old Aberdeen, Scotland

# These authors are equally contributed to this study

\* Corresponding authors: Hu Chuan-Peng ([hu.chuan-peng@nnu.edu.cn](mailto:hu.chuan-peng@nnu.edu.cn); hcp4715@hotmail.com)

# **Abstract**

The self-prioritization effect (SPE) refers to the effect that performance on cognitive tasks is better when stimuli are related to the self than when they are not. In the last decade, the self -perceptual matching task (SPMT) has emerged as a mainstream paradigm for studying SPE due to its simplicity and elimination of familiarity effects. As a simple button-pressing task, SPMT yields two outcomes: reaction time and accuracy. Other indices can be derived from reaction times and accuracy, including sensitivity *d* prime under signal-detection theory, the efficiency index through a direct division between reaction times and accuracy, and drift rate (*v*) and starting point (*z*) estimated using drift-diffusion models. All these indices have been used to quantify SPE in the literature. However, the reliability of these SPE indices has never been tested. To fill this gap, we will re-analyze existing data from multiple datasets using intraclass correlation coefficient (ICC) and split-half reliability. Our results will provide a benchmark for future studies.

Keywords: Self-Prioritization Effect (SPE), Self-Perceptual Matching Task (SPMT), Reliability, Multiverse

# **Introduction**

The Self-Prioritization Effect (SPE) refers to the phenomenon whereby performance in cognitive tasks is better when stimuli are related to the self than when they are not. This effect has been widely documented and confirmed since the 1950s. In the early days of cognitive psychology, researchers found that subjects were able to recognize their own names, even when they were mixed with a noisy auditory background and not the target of the task in dichotic listening tasks (Cherry, 1953; Moray, 1959). SPE effect was then reported in memory research by Craik and Tulving (1975), who found that participants were able to recall more words when they were related to the self compared to when they were processed at other levels (e.g., semantic). This SPE effect in memory was then replicated by many others (Conway & Dewhurst, 1995; Rogers et al., 1977; Symons & Johnson, 1997). In the following decades, the SPE has also been found to occur with different stimuli, such own face (Keenan et al., 2000; Kircher et al., 2000; Turk et al., 2002), own voice (Hughes & Harrison, 2013; Payne et al., 2021), own name (Constable, Rajsic, et al., 2019), and newly owned object (Strachan et al., 2020). SPE was found across a variety of cognitive tasks, such as perceptual task (Cunningham & Turk, 2017; Desebrock et al., 2018), decision-making task (Sui & Humphreys, 2013), attentional task (Shapiro et al., 1997), and ownership task (Cunningham et al., 2008).

Although SPE is often argued to be a self-specific effect, it can be challenging to disassociate it from the familiarity effect since most studies use stimuli owned by participants or by others. Sui et al. (2012) proposed a paradigm where participants first associate geometrical shapes (e.g., triangle, square, and circle) with labels of persons (e.g., "You," "friend," and "stranger") and then perform a perceptual matching task in which they decide if the shape-label pairs presented on the screen match the learned association or not (Sui et al., 2012). Because the task requires participants to learn the social meaning of different geometric shapes, it is called the Self-Perceptual Matching Task (SPMT). In this task, Sui et al. (2012) found that shapes associated with the self are performed better, with faster response times, better accuracy, and/or higher sensitivity scores, compared to shapes associated with friends and strangers. Because the self-relatedness is acquired immediately right before they start the perceptual matching task, this paradigm eliminated the effect of familiarity of the stimuli.

Since then, the SPMT has become the mainstream method for investigating the mechanism underlying the SPE. For instance, researchers have explored the importance of personality traits in identity labels (Golubickis et al., 2020), the self-relevant labels that include the past, present, and future self (Golubickis et al., 2017), as well as "good self" and "bad self" labels (Hu et al., 2020), and the group advantage effect of in-group labels (Constable, Elekes, et al., 2019; Constable & Knoblich, 2020; Enock et al., 2018; Enock et al., 2020). Moreover, the SPMT has been applied to various fields. In neuroscience and physiology, researchers investigate which brain regions are activated during self-prioritization effect (Feng et al., 2018; Humphreys & Sui, 2015), and gender differences in self-prioritization effect due to oxytocin (Feng et al., 2020). In clinical research, SPMT has been used to understand atypical self-processing in populations such as those with autism or depression (Gillespie‐Smith et al., 2018; Nijhof & Bird, 2019; Sui & Humphreys, 2017). Cross-cultural studies have shown that individuals from individualistic cultures demonstrate a stronger self-prioritization effect (Jiang et al., 2019), and that the language of the experimental stimuli can affect the strength of the effect (Ivaz et al., 2016). Finally, the SPMT has also been applied to child development, with studies examining developmental changes in self-positivity effects (Maire et al., 2020; Zhou et al., 2019).

While SPMT has become a mainstream method for investigating the mechanism underlying the SPE, little attention has been paid to the exact indices of the effect and their reliability, which require careful examination. (Parsons et al., 2019; Zorowitz & Niv, 2023). This issue is especially important because SPMT is increasingly being used to measure individual differences in psychiatry (Liu et al., 2022) , and social psychology (Enock et al., 2018). To address this gap, we plan to examine the reliability of SPE indices in SPMT by reanalyzing data from multiple sources (see Table 1 for details of the data sources).

To provide a comprehensive assessment of the SPE indices from SPMT, we will include six indices that measure the difference between self and other, using different outcomes of the matching trials. Specifically, these indices comprise of two direct measures based on SPMT, namely reaction times and accuracy, as well as derived measures such as efficiency (Humphreys & Sui, 2015; Stoeber & Eysenck, 2008), *d* prime of Signal Detection Theory (SDT) (Hu et al., 2020; Sui et al., 2012), and drift rate (*v*) and starting point (*z*) from Drift Diffusion Model (DDM) (Golubickis et al., 2017).

Given that there are multiple methods for calculating reliability of cognitive tasks, we will employ Split-Half Reliability and Intraclass Correlation Coefficient (ICC) to calculate the reliability of each SPE index mentioned earlier. This study's findings are exploratory and may yield valuable insights into the reliability and consistency of SPMT. Such insights could potentially enable the use of SPMT in research, clinical settings, and personal performance monitoring in the future.

# **Methods**

## Ethics information

Since our research involves a secondary analysis of pre-existing data obtained from publicly available datasets or archived data from our group, which have used SPMT in recent years, informed consent and confidentiality are not applicable.

## Datasets

In order to assess the reliability of SPMT, we will provide a brief overview of its experimental design, as outlined in Experiment 1 by Sui et al. (2012). The original SPMT used a 2 by 3 within-subject design. The first independent variable, "Matching", has two levels: "Matching" and "Nonmatching" and indicates whether the shape and label match. The second independent variable, "Identity", has three levels: "Self", "Friend", and "Stranger" and represents the identity that the shape corresponds to.

The original SPMT consisted of two phases (see figure 1). In the first phase (learning phase), participants completed a learning task in which they associated three geometric shapes (circle, triangle and square) with three labels (self, friend, and stranger) for approximately 60 seconds. The shape-label associations were balanced across participants. In the second phase (formal experimental phase), participants completed a perceptual matching task. Each trial started with a fixation cross displayed in the center of the screen for 500 ms, followed by a shape-label pairing and fixation cross for 100 ms. the screen then went blank for 1500 ms, or until a response was made. Participants were required to judge whether the presented shape and label matched the learned associations from the learning phase and respond as quickly and accurately as possible by pressing one of two buttons within the allotted timeframe.

Prior to the formal experimental phase, participants completed a training session consisting of 24 practice trials. After the training, participants completed six blocks of 60 trials in the matching task, with two matching types (matching/nonmatching) and three shape associations, for a total of 60 trials per association. Short breaks lasting up to 60 seconds were provided after each block.



**Figure 1.** Procedure of the original SPMT in the Experiment 1 (Sui et al., 2012).

Note that the relation between shape-label pairs is counter-balanced between participants.

We will analyze available datasets that have raw data from empirical studies employed SPMT. These articles are from an on-going meta-analysis (see protocol: <https://osf.io/ygqz9/?view_only=f604a192cac6497b966cc58174e7dc9e>), and all of them shared raw data publicly (Golubickis & Macrae, 2021; Navon & Makovski, 2021; Qian et al., 2020; Schäfer & Frings, 2019; Svensson et al., 2022) and did not deviate from the original experimental paradigm. Additionally, we found five other articles that did not have publicly available data but stated that data were available upon request. (Bukowski et al., 2021; Cheng & Tseng, 2019; Kolvoort et al., 2020; Martínez-Pérez et al., 2020; Xu et al., 2021), one article stated that data were shared on OSF (https://osf.io/pcv3u/) but the repository is empty (Bukowski et al., 2021). We will include the datasets with raw data that are available to us.

Because direct replications are somewhat discouraged by the research culture (Makel et al., 2012), all the datasets we included in our analysis involved some modifications to the original design, such as including additional independent variables, different experimental materials and so on. All datasets are selected based on two criteria: (1) the experimental design did not deviate from the original SPMT (Sui et al., 2012); (2) the trial-level data is available so that we can estimate at least one reliability index. The dataset is described below.

Table 1. Dataset Information

| Paper | Exp. | Independent Variable | | | | Sample  Size | # of Trials per Condition | Self-Prioritization Effect Indices | | | | | | Reliability | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| IV 1 | IV 2 | IV 3 | IV 4 | RT | ACC | d | Eff | v | z | ICC | SHR |
| Hu et al. (2023) | 1 | Matching | Identity | Emotion  Control, Neutral,  Happy, Sad | Session | 34 | 60 | √ | √ | √ | √ | √ | √ | √ | √ |
| Constable and Knoblich (2020) | 1 | Matching | Identity | Switch Identity  Partner, Stranger | Phase | 92 | 40 | √ | √ | √ | √ | √ | √ |  | √ |
| Constable et al. (2021) | 2 | Matching | Identity  Self; Stranger |  |  | 51 | 24 | √ | √ | √ | √ | √ | √ |  | √ |
| Qian et al. (2020) | 1 | Matching | Identity Self; Stranger; Celebrity | Mood (Session) |  | 24 | 24 | √ | √ | √ | √ | √ | √ |  | √ |
| 2 | Matching | Identity Self; Celebrity | Cue  With, Without |  | 25 | 50 | √ | √ | √ | √ | √ | √ |  | √ |
| Schäfer and Frings (2019) | 1 | Matching | Identity Self; Mother; Acquaintance |  |  | 103 | 24 | √ | √ | √ | √ | √ | √ |  | √ |
| Golubickis and Macrae (2021) | 1 | Matching | Identity | Presentation Mixed; Blocked |  | 30 | 30 | √ | √ | √ | √ | √ | √ |  | √ |
| Navon and Makovski (2021) | 1 | Matching | Identity |  |  | 13 | 60 | √ | √ | √ | √ | √ | √ |  | √ |
| 3 | Matching | Identity  Self; Father; Stranger |  |  | 27 | 60 | √ | √ | √ | √ | √ | √ |  | √ |
| 4 | Matching | Identity |  |  | 26 | 60 | √ | √ | √ | √ | √ | √ |  | √ |
| Svensson et al. (2022) | 1 | Matching | Identity Self; Friend |  |  | 20 | 50 | √ | √ | √ | √ | √ | √ |  | √ |
| 2 | Matching | Identity Self; Friend | Frequency  self > friend |  | 24 | 100 | √ | √ | √ | √ | √ | √ |  | √ |
| 3 | Matching | Identity Self; Friend | Frequency  self < friend |  | 25 | 100 | √ | √ | √ | √ | √ | √ |  | √ |
| Cheng and Tseng (2019) | 1 | Matching | Identity | Go/No-go |  | 22 | 75 | √ | √ | √ | √ | √ | √ |  | √ |
| 2 | Matching | Identity | Go/No-go |  | 26 | 75 | √ | √ | √ | √ | √ | √ |  | √ |
| 3 | Matching | Identity | Go/No-go |  | 22 | 75 | √ | √ | √ | √ | √ | √ |  | √ |
| Bukowski et al. (2021) | 1 | Matching | Identity | Imitation |  | 91 | 60 | √ | √ | √ | √ | √ | √ |  | √ |
| 2 | Matching | Identity | Imitation |  | 109 | 60 | √ | √ | √ | √ | √ | √ |  | √ |
| Kolvoort et al. (2020) | 1 | Matching | Identity | Delay  0, 40, 120, 700 |  | 31 | 25 | √ | √ | √ | √ | √ | √ |  | √ |
| Martínez-Pérez et al. (2020) | 1 | Matching | Identity | Stimulation |  | 90 | 40 | √ | √ | √ | √ | √ | √ |  | √ |
| Xu et al. (2021) | 1 | Matching | Identity | Feedback | Sex | 105 | 60 | √ | √ | √ | √ | √ | √ |  | √ |
| Woźniak et al. (2018) | 1 | Matching | Identity | Facial Gender  Mele; Female |  | 18 | 56 | √ | √ | √ | √ | √ | √ |  | √ |
| 2 | Matching | Identity | Facial Gender  Mele; Female |  | 18 | 60 | √ | √ | √ | √ | √ | √ |  | √ |
| Liu et al. (2023) | 1 | Matching | Identity  Self; Stranger |  |  | 298 | 16 | √ | √ | √ | √ | √ | √ |  | √ |

Note. ICC: Intraclass Correlation Coefficient, SHR: split-half reliability

**Dataset 0**: Hu et al. (2023). This dataset is obtained from our lab but never published before. In this experiment, 34 participants (age: mean + SD, XX female) were recruited. The experiment design has 4 independent variables: 2 (Matching: matching, nonmatching) × 3 (Identity: self, friend, stranger) × 4 (Emotion: control, neutral, happy, sad) × 6 (sessions: 1-6). In each session, participants completed 60 trials for each experimental condition, with a one-week interval between each session. In each session, in addition to the matching and identity variables consistent with the original SPMT experiment, the emotional expression conveyed by the shape was controlled. Unlike Sui et al. (2012), this study included an additional independent variable: emotions (control, neutral, happy, and sad). Only the data from the "control emotion" condition will be included in the current analysis.

**Dataset 1**: Phase 1 of Constable and Knoblich (2020). In this study, 92 participants completed 40 trials per experimental condition. The experiment consisted of 4 independent variables: 2 (Matching: matching, nonmatching) ×3 (Identity: self; mother; acquaintance) ×3 (Switch Identity: Partner, Stranger) ×2 (Phase: 1, 2). In phase 1, participants completed the original SPMT, while in phase 2, they were required to respond to different shape-label pairing rules. In the "Switch Identity: Partner" group, shapes originally matched with the self were now matched with the partner, while in the " Switch Identity: Stranger" group, shapes originally matched with the self were now matched with the stranger. Only data from “Phase 1” condition were included in the analysis.

**Dataset 2**: Experiment 2 of Constable et al. (2021). In this study, 51 participants completed 24 trials in each experimental condition. The experiment consisted of 2 independent variables: 2 (Matching: matching, nonmatching) ×2 (Identity: self, stranger). It's worth noting that in this study, identities are not represented by different shapes, but rather by the brightness or darkness of the shapes.

**Dataset 3**: Qian et al. (2020). In Experiment 1 of this study, 24 participants completed 24 trials in each experimental condition. The experiment consisted of 3 independent variables: 2 (Matching: matching, nonmatching) ×3 (Identity: self; stranger; celebrity) × 3 (Session: happiness, anxiety, serenity, depression). Unlike Sui et al. (2012), the experiment was repeated four times, with participants being in different moods in each session. The authors found that different moods can lead to varying degrees of SPE, making it unclear whether the difference in SPE across the four sessions is due to time or mood factors. Consequently, we will not calculate ICC on this data.

In Experiment 2 of this study, 25 participants completed 50 trials in each experimental condition. The experiment consisted of 3 independent variables: 2 (Matching: matching vs. nonmatching) ×2 (Identity: self; celebrity) × 2 (Cue: with; without). To avoid the influence of “cue” on SPE, only data from the “without cue” condition was included in the analysis.

**Dataset 4**: Schäfer and Frings (2019). In this study, 103 participants completed 24 trials in each experimental condition. The experiment consisted of 2 independent variables: 2 (Matching: matching, nonmatching) ×3 (Identity: self; mother; acquaintance). The only difference between this experiment and the Sui et al. (2012) is that the labels used for the “Identity” were changed from "friend" and "stranger" to "mother" and "acquaintance".

**Dataset 5**: Experiment 1 of Golubickis and Macrae (2021). In this study, 30 participants completed 30 trials in each experimental condition. The experiment consisted of 3 independent variables: 2 (Matching: matching, nonmatching) ×3 (Identity: self, friend, stranger) × 2 (Presentation: mixed; blocked). In the mixed-presentation blocks, shapes were displayed in a randomized order and were equally likely to Sui et al. (2012). To avoid the influence of presentation on SPE, we will only analyze the data from the mixed condition.

**Dataset 6**: Experiment 1, 3 and 4 of Navon and Makovski (2021). The Experiment 1, 2, and 4 all consisted of 2 independent variables: 2 (Matching: matching, nonmatching) × 3 (Identity: self, friend, stranger). Only the “Identity” in Experiment 3 is different, where one level is not 'friend' but 'father'. In addition, the labels of this study are all in Hebrew.

**Dataset 7**: Experiment 1 of Svensson et al. (2022). In this study, 20 participants completed 50 trials in each experimental condition. The experiment consisted of 2 independent variables: 2 (Matching: matching, nonmatching) ×2 (Identity: self, friend).

**Dataset 8**\*[[1]](#footnote-1): Experiment 1, 2, and 3 of Cheng and Tseng (2019). The Experiment 1, 2, and 3 all consisted of 3 independent variables: 2 (Matching: matching, nonmatching) × 2 (Identity: self, friend) × 2 (Go/No-go: green, red). The shape and label stimuli were presented in different colors, with green indicating a "go" trial that required a response from the participant, while red indicated a "no-go" trial that did not require a response from the participant but was responded to by the partner. In Experiment 1, the partner was actually present, while in Experiments 2 and 3, the partner was not physically present. There were 22 participants in Experiment 1, 26 participants in Experiment 2 and 22 participants in Experiment 3. And each experiment had 100 trials per condition.

**Dataset 9\***: Experiment 1 and 2 of Bukowski et al. (2021). The Experiment 1 and 2 both consisted of 3 independent variables: 2 (Matching: matching, nonmatching) × 2 (Identity: self, friend, stranger) × 4 (Imitation: imitation, imitation-inhibition, inhibition-control, be-imitated). There were 91 participants in Experiment 1 and 109 participants in Experiment 2. And each experiment had 60 trials per condition. The variables of “Imitation” is not of interest to us, and therefore we will conduct our analysis without taking this independent variable into consideration.

**Dataset 10**: Kolvoort et al. (2020). In this study, 31 participants completed 25 trials in each experimental condition. The experiment consisted of 3 independent variables: 2 (Matching: matching, nonmatching) × 2 (Identity: self, friend, stranger) × 3 (Delay: 0, 40ms, 120ms, 700ms). To avoid the influence of “Delay” on SPE, only data from “no delay” condition was included in the analysis.

**Dataset 11\***[[2]](#footnote-2): Martínez-Pérez et al. (2020). In this study, 90 participants completed 40 trials in each experimental condition. The experiment consisted of 3 independent variables: 2 (Matching: matching, nonmatching) × 2 (Identity: self, friend, stranger) × 5 (Stimulation: DLPFC-A, DLPFC-C, Sham, VMPFC-A, VMPFC-C). The variable of “Stimulation” is not of interest to us, and therefore we will conduct our analysis without taking this independent variable into consideration.

**Dataset 12**: Xu et al. (2021). In this study, 105 participants completed 60 trials in each experimental condition. The experiment consisted of 4 independent variables: 2 (Matching: matching, nonmatching) × 2 (Identity: self, friend, stranger) × 2 (Feedback: acceptance, rejection) × 2 (sex: men, women). The variables of “Feedback” and “Sex” are not of interest to us, and therefore we will conduct our analysis without taking these independent variables into consideration.

**Dataset 13**: Experiment 1 and 2 of Woźniak et al. (2018). The Experiment 1 and 2 both consisted of 3 independent variables: 2 (Matching: matching, nonmatching) × 2 (Identity: self, friend, stranger) × 2 (Facial Gender: male; female). There were 18 participants in Experiment 1 and 18 participants in Experiment 2. The Experiment 1 had 56 trials per condition and The Experiment 2 had 60 trials per condition. Labels are matched with different faces. However, “Facial Gender” is not of interest to us, and therefore we will conduct our analysis without taking this independent variable into consideration.

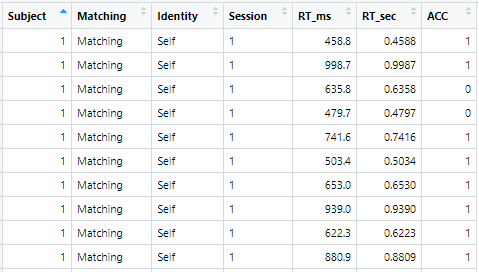
**Dataset 14**:Liu et al. (2023). In this study, 298 participants completed 16 trials in each experimental condition. The experiment consisted of 2 independent variables: 2 (Matching: matching, nonmatching) × 2 (Identity: self, stranger). Labels are matched with different faces.

The experimental design of all the included datasets did not significantly deviate from the original task. However, not all studies had repeated measures like our simulated data. If a publicly available data did not have repeated SPMT within a certain time interval, we will not calculate its ICC, but only calculate split-half reliability.

## Simulated data

The simulated data was generated for checking the analytical script and avoid peeking the real data and inducing potential biases. This section is for preregistration only and will be removed in the real data analysis. Instead, we generated a fake dataset with the same format as the primary data. We used an open dataset from a previous study examining the self-prioritization effect as a reference to create our pilot data.

We utilized Bootstrap methods, drawing samples from Hu et al. (2020) open dataset (accessible at <https://osf.io/mhdsn/>) with replacement (allowing the same sample to be repeated in the pilot data). The pilot data includes 6 sessions of data from 34 participants, with each participant having 24 practice trials and 360 experimental trials (6 different types of shape-label associations: 2 (matching: matching, nonmatching) × 3(Identity: self, friend, stranger), 60 trials per association) per session. Figure 1 shows the first 6 rows of the pilot data.



**Figure 2.** The first six rows of the simulated data

## Analysis Plan

All analyzes will be performed in R 4.2.2 (R Core Team, 2023)[[3]](#footnote-3).

As mentioned earlier and shown in Table 1, we have currently collected 13 publicly available datasets, one dataset from our own laboratory, and one dataset our collaborators (Liu et al., 2023). Note that the number of datasets may increase as we will continue searching for publicly available experimental data that utilize SPMT.



**Figure 3.** Roadmap of the current study.

Note. SPE: self-prioritization effect; *d* prime is the sensitivity index under the signal detection theory; drift rate and starting point are parameters derived from drift-diffusion model; ICC: Intraclass Correlation Coefficient, SHR: split-half reliability.

### Data pre-processing

We will pre-process the secondary data using the following criteria (Note: we did not pre-process the data before preregistration):

1. Participant exclusion criteria
2. Participants who had wrong trial numbers because of procedure errors will be excluded from the analysis.
3. Participants with an overall accuracy < 0.5 will be excluded from the analysis.
4. Participants with any of the conditions with zero accuracy will be excluded from the analysis.
5. Behavioural data exclusion criteria
6. Trials with no response or wrong key press will be excluded from the analysis.
7. Responses with reaction times less than 200ms or greater than 1500ms will be excluded from the analysis.
8. The practice trials will be excluded from the formal analysis.
9. The data under conditions other than the “control condition” will not be used in the current study.
10. Organize data structures
11. Standardize the labels in the variable 'Matching' to ‘Matching’ and ‘Nonmatching’
12. Standardize the labels in the variable 'Identity' to 'Self', ‘Friend’ and 'Stranger’
13. Convert the unit of reaction time from milliseconds to seconds or from seconds to milliseconds, and name them as RT\_ms and RT\_sec respectively, while keeping both variables.
14. Arrange the variables in the order of Subject, Session (if applicable), Matching, Identity, RT\_ms, RT\_sec, and ACC.

### Calculating the SPE

For each dataset, we will calculate six indices for each experimental condition: reaction times, accuracy, *d* prime, efficiency, drift rate (*v*), and starting point (*z*). Reaction time and accuracy can be obtained directly from the datasets, while *d* prime and efficiency will be calculated based on reaction time and accuracy using a simple formula (see Table 3). The drift rate (*v*) and starting point (*z*) of the drift-diffusion model will be estimated by the "fit\_ezddm" function in the "hausekeep" package (Lin et al., 2020), which wrapped the function from EZ-DDM (Wagenmakers et al., 2007).

Based on different targets, we calculate SPEs separately for “Self” vs “Friend”, and for “Self” vs “Stranger”.

Table 2. Indices in SPMT and corresponding calculation of indices and SPE

| **Indices** | **Indices Calculation** | **SPE Calculation Based on Indices** | **Source** |
| --- | --- | --- | --- |
| Mean  Reaction Times  (RT) |  |  | Sui et al. (2012) |
| Accuracy (ACC) |  |  | Sui et al. (2012) |
| *d* prime |  |  | Sui et al. (2012) |
| Efficiency |  |  | Humphreys and Sui (2015); Stoeber and Eysenck (2008) |
| Drift rate (*v*) | DDM：parameters will be identified through model selection |  | Golubickis et al. (2017) |
| Starting Point (*z*) |  | Golubickis et al. (2017) |

Note. DDM: Drift Diffusion Model.

### Estimating the Reliability

**Split-half reliability.** We will calculate the split-half reliability of the six indices using four types of split-half reliability measures: odd-even, front-back, permutation, and Monte Carlo (Kahveci et al., 2022; Pronk et al., 2022). The odd-even split divides trials into odd and even numbered sequences, while the front-back split divides the first and second halves of trials. The permutation split shuffles the trial order and randomly assigns each half to a group. The Monte Carlo split-half is similar to the permutation split-half, but it repeats the process thousands of times to calculate the average and 95% confidence interval of the split-half reliability. This study will primarily use Monte Carlo split-half to determine the split-half reliability of SPMT for its robustness (Pronk et al., 2022). The results of the other three split-half methods will be presented in the supplementary materials.

First, the data will be stratified according to Session (if applicable), Matching, and Identity. If the data is not stratified, directly splitting it in half will result in an uneven distribution of trials for each experimental condition in the two halves, which can lead to an overestimation or underestimation of split-half reliability. Therefore, once the data is stratified, we split it into two halves. For example, when using Monte Carlo Split-Half, we randomly split the data into two halves. Then we repeat this process 1000 times. This will result in 1000 pairs of two halves of the data. Next, we use these 1000 pairs of data to calculate 1000 Pearson correlation coefficients, and then obtain the average and 95% confidence interval of the Monte Carlo split reliability. First-second split, odd-even split, and permutated split are similar to Monte Carlo method, but they only perform one split, so only one split-half reliability is obtained without an interval estimate of the split-half reliability.

**Test-Retest Reliability (ICC).** For dataset with more than one experiment sessions, we will also calculate the Intraclass Correlation Coefficient (ICC) to evaluate the test-retest reliability of these six indices. Specifically, we will use the "psych" package to calculate ICC for these indices (Revelle, 2017). ICC is a well-established measure of reliability in test-retest, intra-rater, and inter-rater studies (Fisher, 1992). Compared to Pearson correlation coefficient, ICC considers both the degree of correlation and agreement between multiple measurements, making it a more comprehensive measure of test-retest reliability (Koo & Li, 2016).

ICC2 measures the proportion of the total variance that is due to between-subject variance. For the calculation of ICC2 estimates and their 95% confidence intervals, the formula is:

Note. MSR = mean square for rows; MSE = mean square for error; MSC = mean square for columns; n = number of subjects.

ICC2k measures the proportion of the total variance that is due to within-subject variance. For the calculation of ICC2k estimates and their 95% confidence intervals, the formula is:

Note. MSR = mean square for rows; MSE = mean square for error; MSC = mean square for columns; n = number of subjects; k = number of raters/measurements.

A value less than 0.6 means poor reliability, a value between 0.6 and 0.8 indicates substantial reliability, and a value greater than 0.8 means excellent reliability (Cicchetti & Sparrow, 1981; Kupper & Hafner, 1989). If SPMT is reliable for measuring individual differences, then the ICC2 is large and the ICC2k is small.

# **Data availability**

The de-identified raw data from our lab (Dataset 0) is available at: (https://doi.org/10.57760/sciencedb.08117). The simulated data is accessible on GitHub (https://github.com/Chuan-Peng-Lab/ReliabilitySPE).

# **Code availability**

Code used to simulate and analyze the data is made accessible at https://github.com/Chuan-Peng-Lab/ReliabilitySPE.

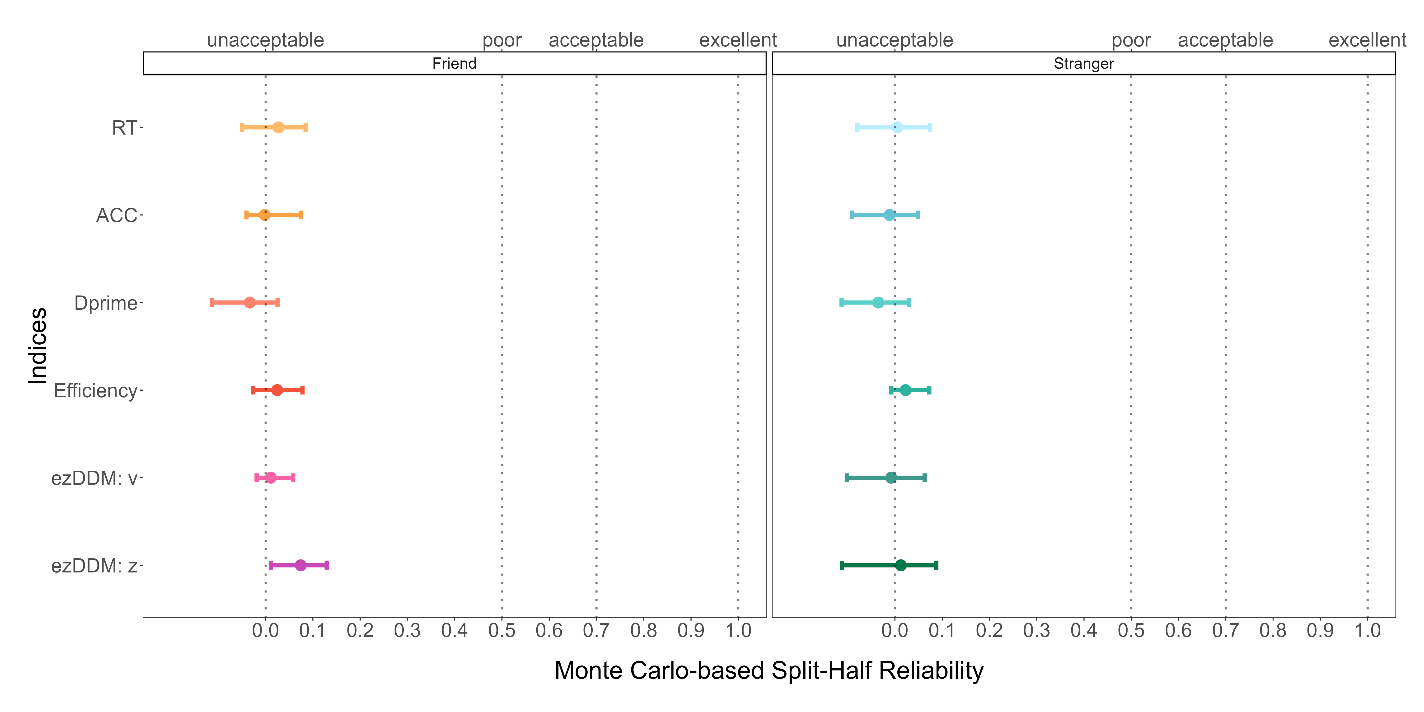
# **Results**

The results reported below are based on simulated data, will be updated with real data in the final report.

## 

## Split-Half Reliability

First, we stratified the data based on three variables: Session (if applicable), Matching, and Identity, and then split the stratified data into two halves using four methods. Next, we calculated the SPE for each of the six indices for each half of the data. Finally, we calculated the split-half reliability for each of the six SPEs. As shown in Figure 4, when using the Monte Carlo split-half, the split-half reliability of the six indices obtained is very high, with the highest value of Efficiency, which means that it is the most stable of the six SPE indexing calculations for split-half reliability. The results from the other three split-half methods were like the Monte Carlo method, which will be presented in the supplementary material.

****

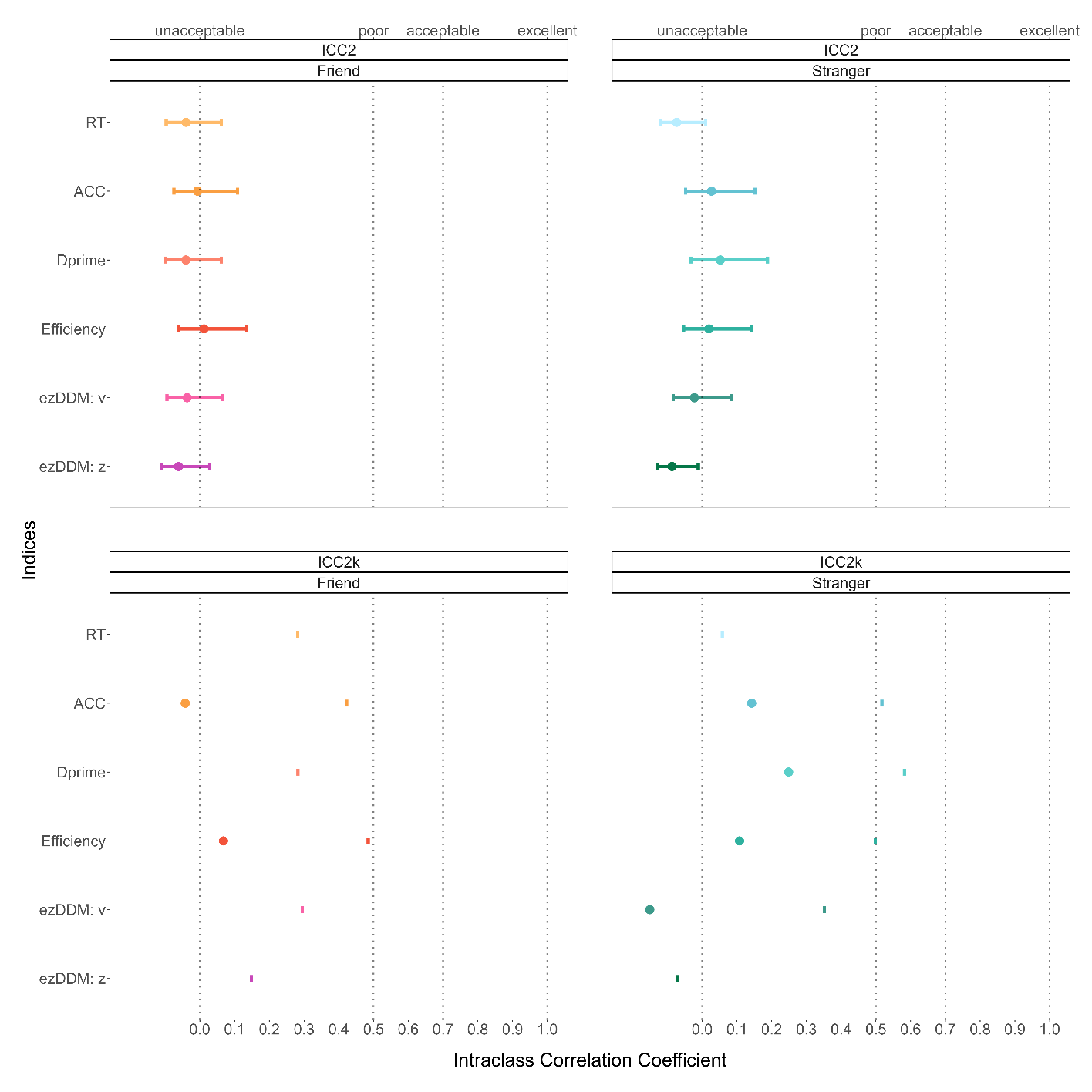
**Figure 4. Split-Half Reliability.**

*Note:* the results here are from simulation, this figure will be replaced with plots from real data.

## Intraclass correlation coefficient (ICC)

Intraclass correlation coefficient (ICC) is a measure of the consistency or reliability of measurements made by different raters (observers) or repeated measurements made by the same rater (observer). We will calculate ICC only if the study involves repeated measurements of SPMT. Essentially, it tells us how much of the variation in the data can be attributed to differences between raters or repeated measurements, and how much can be attributed to differences within the subjects being measured. In simple terms, it provides an idea of the proportion of total variation in the data that is due to the true differences between subjects, versus due to measurement error or random fluctuations.

The present study aimed to investigate the stability of six indices, including reaction time (RT), accuracy (ACC), *d* prime, Efficiency, drift rate (*v*) and starting point (*z*) in the diffusion decision model (DDM), across time sessions. We utilized the Intraclass Correlation Coefficients (ICC) to evaluate the proportion of variation in SPMT that could be attributed to within-subject repeatability over time and between-subject differences. Specifically, we are most interested in ICC2 and ICC2k, where ICC2 represents the ratio of between-subject variance to total variance, and ICC2k represents the ratio of within-subject variance to total variance. Thus, if SPMT is reliable for measuring individual differences, then the ICC2 is large and the ICC2k is small. As shown in Figure 5, the ICC2 values of the six indices are relatively large and ICC2k values are relatively small, supporting our hypothesis.

****

**Figure 5. Intraclass correlation coefficient****.**

*Note:* this plot is from simulated data, will be replaced with real data in our formal analyses

# **Discussion**

Do **not** include a **Discussion** section.

# **Acknowledgements**

The present research is support by.

# **Author contributions**

HCP contributed to the conception and supervision of the study. JS contributed to fund raising, HCP contributed to data collection. ZL, ZYR and HMZ will perform the data pre-processing, analysis and visualize the results. In addition, ZL, JS, HMZ and HCP will contribute to discussing the results and the drafting of the final manuscript. All authors will critically revise the manuscript.

# **Competing interests**

The authors declare no competing interests.

# **Supplementary information**

Table S1 Split-Half Reliability of Other Split Methods based on simulated data

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Indices of SPE | Method of SH | SHR |  | Indices of SPE | Method of SH | SHR |
| RT | First-Second | .01 |  | Efficiency | First-Second | .07 |
| RT | Odd-Even | -.05 |  | Efficiency | Odd-Even | -.04 |
| RT | Permuted | .01 |  | Efficiency | Permuted | .05 |
| ACC | First-Second | .02 |  | DDM: v | First-Second | .04 |
| ACC | Odd-Even | -.05 |  | DDM: v | Odd-Even | -.05 |
| ACC | Permuted | .07 |  | DDM: v | Permuted | .10 |
| Dprime | First-Second | .01 |  | DDM: z | First-Second | .07 |
| Dprime | Odd-Even | -.08 |  | DDM: z | Odd-Even | .01 |
| Dprime | Permuted | -.02 |  | DDM: z | Permuted | .13 |

Note. Target: Friend, SH: split-half, SHR: split-half reliability, SPE: self-prioritization effect

Table S2 Split-Half Reliability of Other Split Methods based on simulated data

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Indices of SPE | Method of SH | SHR |  | Indices of SPE | Method of SH | SHR |
| RT | First-Second | .01 |  | Efficiency | First-Second | .07 |
| RT | Odd-Even | -.01 |  | Efficiency | Odd-Even | -.03 |
| RT | Permuted | .08 |  | Efficiency | Permuted | -.01 |
| ACC | First-Second | .02 |  | DDM: v | First-Second | .03 |
| ACC | Odd-Even | -.03 |  | DDM: v | Odd-Even | -.01 |
| ACC | Permuted | -.10 |  | DDM: v | Permuted | -.11 |
| Dprime | First-Second | .03 |  | DDM: z | First-Second | .01 |
| Dprime | Odd-Even | -.07 |  | DDM: z | Odd-Even | .03 |
| Dprime | Permuted | -.07 |  | DDM: z | Permuted | -.14 |

Note. Target: Friend, SH: split-half, SHR: split-half reliability, SPE: self-prioritization effect

# **References**

Bukowski, H., et al. (2021). Socio-cognitive training impacts emotional and perceptual self-salience but not self-other distinction. *Acta psychologica, 216*, 103297. <https://doi.org/10.1016/j.actpsy.2021.103297>

Cheng, M., & Tseng, C.-h. (2019). Saliency at first sight: instant identity referential advantage toward a newly met partner. *Cognitive Research: Principles and Implications, 4*(1), 1-18. <https://doi.org/10.1186/s41235-019-0186-z>

Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America, 25*(5), 975-979. <https://doi.org/10.1121/1.1907229>

Cicchetti, D.V., & Sparrow, S.A. (1981). Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *Am J Ment Defic, 86*(2), 127-137. <https://psycnet.apa.org/record/1982-00095-001>

Constable, M.D., et al. (2021). Affective compatibility with the self modulates the self-prioritisation effect. *Cognition and Emotion, 35*(2), 291-304. <https://doi.org/10.1080/02699931.2020.1839383>

Constable, M.D., et al. (2019). Relevant for us? We-prioritization in cognitive processing. *Journal of Experimental Psychology: Human Perception and Performance, 45*(12). <https://doi.org/10.1037/xhp0000691>

Constable, M.D., & Knoblich, G. (2020). Sticking together? Re-binding previous other-associated stimuli interferes with self-verification but not partner-verification. *Acta psychologica, 210*, 103167. <https://doi.org/10.1016/j.actpsy.2020.103167>

Constable, M.D., et al. (2019). It is not in the details: Self-related shapes are rapidly classified but their features are not better remembered. *Memory & Cognition, 47*, 1145-1157. <https://doi.org/10.3758/s13421-019-00924-6>

Conway, M.A., & Dewhurst, S.A. (1995). The self and recollective experience. *Applied Cognitive Psychology, 9*(1), 1-19. <https://doi.org/10.1002/acp.2350090102>

Craik, F.I.M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268-294. <https://doi.org/10.1037/0096-3445.104.3.268>

Cunningham, S.J., & Turk, D.J. (2017, Jun). Editorial: A review of self-processing biases in cognition. *Quarterly journal of experimental psychology, 70*(6), 987-995. <https://doi.org/10.1080/17470218.2016.1276609>

Cunningham, S.J., et al. (2008). Yours or mine? Ownership and memory. *Consciousness and cognition, 17*(1), 312-318. <https://doi.org/10.1016/j.concog.2007.04.003>

Desebrock, C., et al. (2018). Self-reference in action: Arm-movement responses are enhanced in perceptual matching. *Acta psychologica, 190*, 258-266. <https://doi.org/10.1016/j.actpsy.2018.08.009>

Enock, F., et al. (2018). Self and team prioritisation effects in perceptual matching: Evidence for a shared representation. *Acta psychologica, 182*, 107-118. <https://doi.org/10.1016/j.actpsy.2017.11.011>

Enock, F.E., et al. (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>

Feng, C., et al. (2018). Neural representations of the multidimensional self in the cortical midline structures. *NeuroImage, 183*, 291-299. <https://doi.org/10.1016/j.neuroimage.2018.08.018>

Feng, C., et al. (2020). Effect of intranasal oxytocin administration on self-other distinction: Modulations by psychological distance and gender. *Psychoneuroendocrinology, 120*, 104804. <https://doi.org/10.1016/j.psyneuen.2020.104804>

Fisher, R.A. (1992). Statistical methods for research workers. *Springer New York*. <https://doi.org/10.1007/978-1-4612-4380-9_6>

Gillespie‐Smith, K., et al. (2018). The I in autism: Severity and social functioning in autism are related to self‐processing. *British journal of developmental psychology, 36*(1), 127-141. <https://doi.org/10.1111/bjdp.12219>

Golubickis, M., et al. (2020). Parts of me: Identity-relevance moderates self-prioritization. *Consciousness and cognition, 77*, 102848. <https://doi.org/10.1016/j.concog.2019.102848>

Golubickis, M., et al. (2017). Self-prioritization and perceptual matching: The effects of temporal construal. *Mem Cognit, 45*(7), 1223-1239. <https://doi.org/10.3758/s13421-017-0722-3>

Golubickis, M., & Macrae, C.N. (2021). Judging me and you: Task design modulates self-prioritization. *Acta psychologica, 218*, 103350. <https://doi.org/10.1016/j.actpsy.2021.103350>

Hu, C.-P., et al. (2020). Good Me Bad Me: Prioritization of the Good-Self During Perceptual Decision-Making. *Collabra. Psychology, 6*(1), 20. <https://doi.org/10.1525/collabra.301>

Hu, C.-P., et al. (2023, 2023-05-06). Data for Training Effect of Self Prioritization[DS/OL]. V1. *Science Data Bank*. <https://doi.org/10.57760/sciencedb.08117>.

Hughes, S.M., & Harrison, M.A. (2013). I like my voice better: Self-enhancement bias in perceptions of voice attractiveness. *Perception, 42*(9), 941-949. <https://doi.org/10.1068/p7526>

Humphreys, G.W., & Sui, J. (2015). The salient self: Social saliency effects based on self-bias. *Journal of cognitive psychology, 27*(2), 129-140. <https://doi.org/10.1080/20445911.2014.996156>

Ivaz, L., et al. (2016). The emotional impact of being myself: Emotions and foreign-language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(3), 489. <https://doi.org/10.1037/xlm0000179>

Jiang, M., et al. (2019). Cultural Orientation of Self-Bias in Perceptual Matching. *Front Psychol, 10*, 1469. <https://doi.org/10.3389/fpsyg.2019.01469>

Kahveci, S., et al. (2022). Reliability of reaction time tasks: how should it be computed? <https://doi.org/10.31234/osf.io/ta59r>

Keenan, J.P., et al. (2000). Self-recognition and the right prefrontal cortex. *Trends in cognitive sciences, 4*(9), 338-344. <https://doi.org/10.1016/S1364-6613> (00)01521-7

Kircher, T.T., et al. (2000). Towards a functional neuroanatomy of self processing: effects of faces and words. *Cognitive Brain Research, 10*(1-2), 133-144. <https://doi.org/10.1016/S0926-6410(00)00036-7>

Kolvoort, I.R., et al. (2020). Temporal integration as “common currency” of brain and self‐scale‐free activity in resting‐state EEG correlates with temporal delay effects on self‐relatedness. *Human brain mapping, 41*(15), 4355-4374. <https://doi.org/10.1002/hbm.25129>

Koo, T.K., & Li, M.Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine, 15*(2), 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>

Kupper, L.L., & Hafner, K.b. (1989). On Assessing Interrater Agreement for Multiple Attribute Responses. *Biometrics, 45*(3), 957-967. <https://doi.org/10.2307/2531695>

Lin, H., et al. (2020). Strong effort manipulations reduce response caution: A preregistered reinvention of the ego-depletion paradigm. *Psychological science, 31*(5), 531-547. <https://doi.org/10.1177/0956797620904990>

Liu, T., et al. (2023). To see or not to see: The parallel processing of self-relevance and facial expressions. *Manuscript submitted for publication.* .

Liu, Y.S., et al. (2022). Depression screening using a non-verbal self-association task: A machine-learning based pilot study. *Journal of Affective Disorders, 310*, 87-95. <https://doi.org/10.1016/j.jad.2022.04.122>

Maire, H., et al. (2020). A Developmental Study of the Self‐Prioritization Effect in Children Between 6 and 10 Years of Age. *Child development, 91*(3), 694-704. <https://doi.org/10.1111/cdev.13352>

Makel, M.C., et al. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science, 7*(6), 537-542. <https://doi.org/10.1177/1745691612460688>

Martínez-Pérez, V., et al. (2020). Examining the dorsolateral and ventromedial prefrontal cortex involvement in the self-attention network: A randomized, sham-controlled, parallel group, double-blind, and multichannel HD-tDCS study. *Frontiers in Neuroscience, 14*, 683. <https://doi.org/10.3389/fnins.2020.00683>

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly journal of experimental psychology, 11*(1), 56-60. <https://doi.org/10.1080/17470215908416289>

Navon, M., & Makovski, T. (2021). Are Self-related Items Unique? the Self-prioritization Effect Revisited. <https://doi.org/10.31234/osf.io/9dzm4>

Nijhof, A.D., & Bird, G. (2019). Self‐processing in individuals with autism spectrum disorder. *Autism research, 12*(11), 1580-1584. <https://doi.org/10.1002/aur.2200>

Parsons, S., et al. (2019). Psychological Science Needs a Standard Practice of Reporting the Reliability of Cognitive-Behavioral Measurements. *Advances in methods and practices in psychological science, 2*(4), 378-395. <https://doi.org/10.1177/2515245919879695>

Payne, B., et al. (2021). Perceptual prioritization of self‐associated voices. *British Journal of Psychology, 112*(3), 585-610. <https://doi.org/10.1111/bjop.12479>

Pronk, T., et al. (2022). Methods to split cognitive task data for estimating split-half reliability: A comprehensive review and systematic assessment. *Psychonomic Bulletin & Review, 29*(1), 44-54. <https://doi.org/10.3758/s13423-021-01948-3>

Qian, H., et al. (2020). Prioritised self-referential processing is modulated by emotional arousal. *Quarterly journal of experimental psychology, 73*(5), 688-697. <https://doi.org/10.1177/1747021819892158>

R Core Team. (2023). R: A Language and Environment for Statistical Computing. <https://www.R-project.org/>

Revelle, W.R. (2017). psych: Procedures for personality and psychological research. <https://CRAN.R-project.org/package=psych>

Rogers, T.B., et al. (1977, Sep). Self-reference and the encoding of personal information. *J Pers Soc Psychol, 35*(9), 677-688. <https://doi.org/10.1037//0022-3514.35.9.677>

Schäfer, S., & Frings, C. (2019). Understanding self-prioritisation: The prioritisation of self-relevant stimuli and its relation to the individual self-esteem. *Journal of cognitive psychology, 31*(8), 813-824. <https://doi.org/10.1080/20445911.2019.1686393>

Shapiro, K.L., et al. (1997). Personal names and the attentional blink: a visual "cocktail party" effect. *J Exp Psychol Hum Percept Perform, 23*(2), 504-514. <https://doi.org/10.1037//0096-1523.23.2.504>

Stoeber, J., & Eysenck, M.W. (2008). Perfectionism and efficiency: Accuracy, response bias, and invested time in proof-reading performance. *Journal of research in personality, 42*(6), 1673-1678. <https://doi.org/10.1016/j.jrp.2008.08.001>

Strachan, J.W., et al. (2020). It goes with the territory: Ownership across spatial boundaries. *Journal of Experimental Psychology: Human Perception and Performance, 46*(8), 789. <https://doi.org/10.1037/xhp0000742>

Sui, J., et al. (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. <https://doi.org/10.1037/a0029792>

Sui, J., & Humphreys, G.W. (2013, Nov). Self-referential processing is distinct from semantic elaboration: Evidence from long-term memory effects in a patient with amnesia and semantic impairments. *Neuropsychologia, 51*(13), 2663-2673. <https://doi.org/10.1016/j.neuropsychologia.2013.07.025>

Sui, J., & Humphreys, G.W. (2017). The self survives extinction: Self-association biases attention in patients with visual extinction. *Cortex, 95*, 248-256. <https://doi.org/10.1016/j.cortex.2017.08.006>

Svensson, S.L., et al. (2022). More or less of me and you: self-relevance augments the effects of item probability on stimulus prioritization. *Psychological Research, 86*(4), 1145-1164. <https://doi.org/10.1007/s00426-021-01562-x>

Symons, C.S., & Johnson, B.T. (1997, May). The self-reference effect in memory: a meta-analysis. *Psychological Bulletin, 121*(3), 371-394. <https://doi.org/10.1037/0033-2909.121.3.371>

Turk, D.J., et al. (2002). Mike or me? Self-recognition in a split-brain patient. *Nature neuroscience, 5*(9), 841-842. <https://doi.org/10.1038/nn907>

Wagenmakers, E.-J., et al. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review, 14*(1), 3-22. <https://doi.org/10.3758/BF03194023>

Woźniak, M., et al. (2018). Prioritization of arbitrary faces associated to self: An EEG study. *PloS one, 13*(1), e0190679. <https://doi.org/10.1371/journal.pone.0190679>

Xu, Y., et al. (2021). Romantic feedbacks influence self-relevant processing: the moderating effects of sex difference and facial attractiveness. *Current Psychology*, 1-13. <https://doi.org/10.1007/s12144-021-02114-7>

Zhou, A., et al. (2019). Self-referential processing can modulate visual spatial attention deficits in children with dyslexia. *Frontiers in Psychology, 10*, 2270. <https://doi.org/10.3389/fpsyg.2019.02270>

Zorowitz, S., & Niv, Y. (2023). Improving the reliability of cognitive task measures: A narrative review. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. <https://doi.org/10.1016/j.bpsc.2023.02.004>

1. Note: The datasets marked with an asterisk (\*) indicate that we have contacted the original authors but has not yet got the dataset. [↑](#footnote-ref-1)
2. After contacting the author, we obtained the mean data for each subject under each experimental condition. However, the trial-level data is still unavailable upon we preregister this protocol, we may update this part if the authors provide us trial-level data. [↑](#footnote-ref-2)
3. We may use the latest R version after preregistration and update the R version and packages on Github. [↑](#footnote-ref-3)