Revisiting and Extending a Response Latency Measure of Inclusion of Other in the Self

Janis H. Zickfeld Thomas W. Schubert

University of Oslo

***In press* at CRSP. This paper is not the copy of record and may not exactly replicate the authoritative document published in Comprehensive Results in Social Psychology.**

Author Note

Janis H. Zickfeld and Thomas W. Schubert, Department of Psychology, University of Oslo. Both authors designed the studies, JHZ conducted the main analyses and wrote the first draft, both authors revised the manuscript.

Correspondence concerning this article should be addressed to Janis H. Zickfeld, Postboks 1094 Blindern, 0317 Oslo, Norway. E-mail: jhzickfeld@gmail.com

Abstract

Although the concept of inclusion of the other in the self (IOS) has been successfully assessed with explicit self-report measures implicit procedures have been neglected in past literature. The present article explores the validity of such an implicit measure by proposing several extensions and adaptions. Based on the *me/not-me* response latency task originally conceptualized by Aron, Aron, Tudor and Nelson (1991) we addressed methodological problems by proposing changes in material, calculation of indices and implemented the task in an online environment. We also addressed earlier problems with statistical power and proposed a more powerful way of statistical analyses using mixed models. The me/not-me task is based on the idea that higher overlap between self and other traits results in faster response times of characterizing such a trait as descriptive of the self. This relationship should be observed for close others but not for non-close others. In a sample of 339 US American adults we experimentally manipulated the nature of the target (close vs. distant) and participants engaged in the adapted me/not-me paradigm. Results indicated that trait match had a stronger negative effect on response times for participants in the close condition. The effect was also stronger for participants rating the target higher on the IOS self-report scale. We also provided convergent validity of the me/not-me procedure with other constructs ostensibly measuring *interpersonal closeness*. Future applications and possible limitations are discussed.

230/250

Keywords: inclusion of the other in the self, implicit measure, interpersonal closeness

Revisiting and Extending a Response Latency Measure of Inclusion of Other in the Self

People are inherently social animals and try to seek and engage in relationships with other individuals. Various conceptualizations of relationships have been proposed (Clark & Lemay, 2010). In one of those, Aron and colleagues (e.g. Aron, Aron, & Smollan, 1992; Aron, Aron, Tudor, & Nelson, 1991) provided a cognitive viewpoint by postulating that the other individual is to some extent included in the self, termed *inclusion of the other in the self* (IOS), thereby creating overlapping cognitive structures with the other person.

Beyond various explicit measures of the 'inclusion of the other in the self' concept, only a few valid implicit measurements exist. One response latency paradigm was provided by Aron and colleagues (1991, Experiment 3). For this task, the self and one or more other targets are rated explicitly on adjectives. In a subsequent response latency task, the same adjectives are then used again to describe the self with a *me/not-me* decision (i.e. *does this trait describe you?*). Based on distinctiveness effects, reaction times should be faster for traits that match for close individuals, compared to those that match for non-close others, or compared to non-matching traits. Hence, including the other in the self is operationalized by the speed-advantage of judging matching characterizations of the self and the other compared to non-matching characteristics.

However, that paradigm has been rarely used in the literature on interpersonal relations. This underrepresentation might be because the original format is a) time consuming, b) requires several steps of statistical processing and c) explicit ratings such as the one item pictorial IOS scale (Aron et al., 1992) have a high degree of face validity and are easier to implement. On the other hand, it is possible that the presented paradigm is not suited to investigate IOS and the few published articles represent type I errors (i.e. no actual effect). Implicit measurements have been argued to provide a more powerful option than explicit measurements by circumventing response biases and social desirability problems (e.g. Fazio

& Olsen, 2003). In the present article, we re-investigate the *me/not-me* procedure, provide a replication of the original Experiment 3 by Aron and colleagues (1991), suggest adaptions in order to maximize the applicability and ecological validity of the procedure, and provide a direct test for its relation with the one-item graphical IOS measure and other concepts of interpersonal closeness. Before turning to the *me*/*not-me* procedure and our suggestions and adaptions, we briefly review the IOS concept, its predictors and outcomes, as well as explicit measures of including the other in the self.

Inclusion of the Other in the Self (IOS)

Interpersonal closeness has been conceptualized in various ways. While some authors emphasized relational (Fiske, 2004) or normative aspects (Clark & Mills, 1979), Aron and colleagues have taken a rather cognitive viewpoint by suggesting that closeness is defined by overlapping cognitive structures with another individual. The cognitive representation of another individual thereby shares important aspects with the cognitive representation of the self, resulting in an *inclusion of the other in the self* (see Aron, Mashek & Aron, 2004 for a review). Aron, Lewandowski Jr, Mashek, and Aron (2013) highlighted that this self-other merging entails *inter alia* that we shape our own cognitions by our representation of the other individual, which includes similar perspectives, resources or identities.

The conceptualization has been based on initial evidence suggesting that nouns that have been related to a close individual were more easily remembered than nouns attributed to distant others and to about the same degree as nouns related to the self (Aron et al., 1991, Study 2). The authors argued that including the other in the self facilitated recollection and was also based on perspective taking, a mechanism important for IOS (e.g. Galinsky & Moskowitz, 2000). In another experiment, participants allocated an equal amount of money to a close friend, but not to an acquaintance, also if the other target was not aware of the final distribution. Aron and colleagues emphasized that including the other in the self results in

perceiving other resources or perspectives as one's own. They also argued that including the other in the self is different from purely familiarity or similarity mechanisms (Aron et al., 2004), providing evidence that source-confusion increased for targets high in closeness, but not in similarity (Mashek, Aron, & Boncimino, 2003). The IOS concept is part of the *self-expansion model* that proposes a general framework for the mechanisms of seeking, maintaining or dissolving close relationships (e.g. Aron & Aron, 1986).

IOS has been studied extensively during the last three decades in terms of its predictors and outcomes. Self-disclosure in a relationship is argued to be one of the main predictors of inclusion of the other in the self (Aron, Melinat, Aron, Vallone, & Bator, 1997). In one experiment, the authors allocated pairs of participants to a condition in which they engaged in gradually increasing self-disclosure or to a control condition that consisted of small talk only. Dyads receiving the self-disclosure task scored higher on measures assessing IOS afterwards. These findings have also been extended to groups of four people (Slatcher, 2010). Other predictors of IOS have been identified as taking the perspective of the other target (e.g. Galinsky & Moskowitz, 2000) or inducing positive emotions (Waugh & Fredrickson, 2006).

Research has also explored the outcomes of including the other in the self. Various self-serving biases are affected by overlapping cognitive structures with another individual. IOS is associated with a decreased ability to successfully recall and distinguish between information processed about the self or about the close other (Mashek et al., 2003). Similarly, individuals make situational attributions not only for the self but also for close others in contrast to non-close others for negative events (Sande, Goethals, & Radloff, 1988), and this tendency is associated with IOS (Aron & Fraley, 1999). From a linguistic perspective, including the other in the self is also associated with the number of plural pronouns produced

(e.g. we, us) in comparison to personal pronouns (e.g. I, me; Agnew, van Lange, Rusbult, & Langston, 1998).

The *inclusion of the other in the self* concept has been researched extensively to date and applied to several contexts unrelated to interpersonal relationships (Aron et al., 1991; Aron et al., 1992; Aron et al., 2013). Various explicit and implicit instruments assessing the IOS construct have been proposed and utilized. We will first turn to the most prominent one.

Explicit Measures of IOS

The most frequently used measure of inclusion of the other in the self is the one-item graphical measure (also termed IOS; Aron et al., 1992). Based on the metaphor of overlap, the scale is comprised of seven Venn-like diagrams, each showing a pair of circles increasing in overlap. Participants are instructed to pick the one of the seven circle pairs that best represents their relationship with a specified target. In their original article, Aron and colleagues provided indications that the one-item instrument captures the cognitive construct of including a target in the self, but also aspects of *affective* and *behavioral closeness*. They also showed convergent validity in the form of associations with various multiple-item inventories, and applicability across different cultures (Dalsky, Gohm, Noguchi, & Shiomura, 2008; Gächter, Starmer, & Tufano, 2015; Li, 2002). The IOS scale has also been linked to one relational model proposed in relational models theory (RMT, Fiske, 2004), namely communal sharing, which highlights the common and shared essences of individuals and is subjectively experienced as closeness (Schubert, Zickfeld, Seibt, & Fiske, 2016).

One of the instrument's main advantages is the straightforward metaphor, which circumvents possible language problems. Another advantage is its shortness, making administration feasible and efficient on paper-and-pencil tests and over the Internet (Hodges, Sharp, Gibson, & Tipsord, 2013; Le, Moss, & Mashek, 2007). The IOS scale has also been

adapted to in-groups (Schubert & Otten, 2002), nature (e.g. Schultz, 2002) or God (e.g. Hodges et al., 2013).

Despite its broad applicability and efficiency, the IOS scale is still a self-report measure and may suffer from typical intentional or unintentional report biases such as answers based on social desirability or errors concerning memory (Fazio & Olson, 2003). Therefore, Aron and colleagues suggested that implicit or nonobvious measures of IOS should be utilized in addition to the nonverbal IOS self-report scale. One of these implicit measures is the *me/not me* response latency task originally proposed by Aron and colleagues (1991) and used as evidence for their IOS model. Such an implicit measure might have the advantage of capturing more automatic processing concerning the inclusion of the other in the self (Aron & Fraley, 1999). We will now review the original task and its application in the literature with regard to interpersonal relationships.

The Me/Not-Me Procedure

The Original Me/Not Me Task

In order to show that close individuals are to some extent cognitively included in the self, Aron and colleagues (1991) hypothesized that trait adjectives that are shared by both the self and the other are judged faster as descriptive for the self than descriptions that are true only of the self, the other, or a third non-close other. This line of reasoning is based on two effects: the *descriptiveness effect*, the tendency to recognize traits faster that are descriptive than non-descriptive (Kuiper, 1981), and the *distinctiveness effect,* the tendency to recognize traits more slowly that are only descriptive of the self or only of others, compared to traits where self and other are the same (Ross, Mueller, & De La Torre, 1986). Aron and colleagues argued that a trait that is only self-descriptive but not descriptive of close others results in confusion and thereby slower reaction times. For example, if Ernie considers himself and his close friend Bert as *timid,* he will be faster to recognize himself as *timid*

compared to traits he thinks are only true of him (orange) or Bert (yellow). Due to the fact that the cognitive structures of him and Bert on the domain of timidness overlap to some extent Ernie will be faster to characterize himself as descriptive of this trait than if no overlap is present.

The authors tested this proposition in a straightforward reaction time paradigm. Married participants first rated 90 trait adjectives adapted from Anderson (1968) for themselves, their spouse and a well-known *neutral* non-close other (i.e. Bill Cosby (sic!)) on a 7-point scale. The trait lists included 30 positive, 30 neutral and 30 negative traits and were presented in random order. Afterwards, participants completed a computerized reaction time task. They were again presented with the 90 adjectives and instructed to rate as fast as possible whether the presented trait was descriptive of them or not (i.e. *me* vs. *not me*).

For the analyses, Aron and colleagues (1991) then collapsed ratings on the initial trait rating into *descriptive* (scale points 5-7) and *non-descriptive* (1-3), thereby omitting all traits rated on the midpoint of the scale. Then four cells were created: (1) Ratings that were descriptive of the self and the close person but not the non-close other (TrueTrueFalse, TTF); (2) Ratings that were non-descriptive of the self and the close person but descriptive of the non-close other (FFT); (3) Ratings that were descriptive of the self, non-descriptive of the close other and descriptive of the non-close target (TFT); and finally (4) Ratings that were non-descriptive of the self, descriptive of the close target and non-descriptive of the non-close other (FTF). Traits that were descriptive of all three targets or of nobody (TTT and FFF) were omitted.

The original prediction by Aron and colleagues was that traits that match for the self and their spouse (TTF and FFT) should be judged faster than non-matching traits (TFT and FTF). In an initial 2x2 repeated measures ANOVA they analyzed two different factors, self-descriptiveness (true/false) and trait match with the close other (match/non-match). Self-

descriptive traits were identified faster than non-descriptive traits. Aron et al. found no interaction effect of descriptiveness of self with trait match and therefore collapsed the design into a *t*-test, testing the difference in response time for TTF/FFT vs. TFT/FTF. This resulted in a small significant difference.[1]

Aron and colleagues (1991) also presented the same analyses on the reaction times of *errors* (a trait rated as self-descriptive in the first but not second task and vice versa) and obtained similar results. Notably, they also reported a follow-up experiment in which they failed to replicate the effect regarding the difference between reactions times for TTF/FFT versus TFT/FTF. In addition, based on the small effect and the rather small sample ($N = 10$) one should interpret Aron et al.'s findings with caution.

Further Me/Not Me Publications

Two more articles using the task in an interpersonal context have been published. Aron and Fraley (1999) used nearly the exact same procedure as Aron and colleagues (1991). However, this time the original trait adjectives were only rated for the self and the spouse, and *descriptive* ratings were now coded as ranging from 4-7, thereby including midpoint ratings in contrast to the original version. In contrast to the original article, Aron and Fraley calculated a combined index of reaction time and error effect. The effect was calculated not by taking the difference, but by regressing the average latency (and error count) for mismatching traits onto latencies (and errors) for matching traits, and saving the residual (i.e., difference scores were replaced by residuals of a linear regression). They found a significant effect with judgments of non-matching traits being slower and including more errors, similar

[1] Aron and colleagues (1991) indicate the effect with $p < .05$. Using the available information from the original article ($df = 9$, $t = 2.21$), our calculation gives an actual p-value of .027 for a one-tailed test.

to findings in the original experiment. Aron and Fraley (1999) did not include a non-close target as in the original version. Instead, they related the reaction time data to the graphical measure of inclusion in the self. The overall reaction time index was associated with ratings on the self-report IOS scale. Hence, explicitly judging increased overlap with a target is also linked to faster reaction times in the me/not me task.

Smith, Coats, and Walling (1999) did again include ratings for the romantic partner or spouse and a non-close other (Bill Cosby as in Aron et al., 1991). They also added ratings for an in-group (their student fraternity) and an out-group (people not belonging to the student fraternity). However, the non-close individual was only sampled as a filler according to the authors and was not included in any analyses. Smith and colleagues operationalized *descriptive* ratings again as ranging from 5-7 as in the original experiment, thereby treating ratings of 4 as missing. One important difference was that participants were not asked to rate themselves but their partner on the reaction task. Apart from that, Smith and colleagues followed the general procedure and materials of the original version.

In contrast to the two studies by Aron and colleagues (1991, 1999) Smith and colleagues (1999) utilized a regression model. While Aron et al. (1991) used the self-descriptiveness judgment and their trait match index as the main variables, Smith and colleagues focused on the descriptiveness judgment and the original explicit self-ratings. In contrast to Aron they thereby predicted an interaction effect for these two variables. Descriptiveness judgments that match with explicit ratings should be judged faster than ratings exhibiting no match with the response in the reaction time task. As Aron and Fraley, Smith and colleagues tested the predictive value of the me/not me reaction task for closeness with the IOS scale. They found a significant three-way interaction of trait match with descriptiveness and IOS measure ratings with differences between matches and mismatches increasing when the IOS increased.

A review of the three publications utilizing the me/not me reaction tasks reveals several issues and differences. First, the procedure varies, sometimes including a non-close target for comparison reason, and sometimes a self-report of closeness. Second, the operationalization of trait match is inconsistent. Sometimes midpoint ratings are not included in the final analyses, which leads to excluding several cases. Third, no two studies have ever been analyzed in the same way. [2] Furthermore, recent statistical advances suggest that reaction time data such as these should be analyzed by taking into account all level-1 and level-2 effects (Hoffman & Rovine, 2007; Judd, Westfall, & Kenny, 2012; M. P. Sullivan & Venter, 2005). In the light of the presented issues and inconsistencies, we now provide possible suggestions and adaptions of the me/not-me task for interpersonal contexts that make use of available technology and suitable statistical techniques.

The Extended Me/Not-Me Task

We suggest the following changes to the paradigm:

1.) The original paradigm was mostly characterized by lab experiments because of the computerized reaction time task. With the steady increase of online survey programs, we suggest an online version of the me/not-me task. This feature helps to reach diverse samples with considerable sizes in a short amount of time (Barnhoorn, Haasnoot, Bocanegra, & van Steenbergen, 2014; Crump, McDonnell, & Gureckis, 2013; Schubert, Murteira, Collins, & Lopes, 2013).

2.) We suggest shortening the original number of 90 adjectives to 60, again including a balanced amount of neutral, positive and negative and eliminating potential emotionally loaded trait words to speed up the procedure, and allow it to be used

[2] The me/not-me task has been successfully and extensively utilized and applied in an intergroup context (Craemer, 2014; Otten & Epstude, 2006; Smith & Henry, 1996; Tropp & Wright, 2001; van Veelen, Otten, Cadinu, & Hansen, 2015). However, analytical procedures vary in this body of work as well.

as a measurement technique in emotion research. Building on our first suggestion an online procedure can reach larger sample sizes more feasibly, thereby compensating the removal of 30 adjectives.

3.) The dichotomous coding of trait match between the self and the target into different scale points loses a lot of information (e.g. trait match of the self with a rating of 5 and the other with a rating of 7 is identical to a rating of 7 for both), and creates the problem of how to deal with the midpoint. We therefore suggest expressing trait match as an interaction effect of two continuous predictors (i.e., recoding both self and other ratings on the 7-point scales as ranging from -1 to 1, then calculating their product and dividing that by 2, resulting in a continuous and fractional trait match score from -.5 to .5, with high values indicating high trait match).

4.) Following Sullivan and Venter (2005), we use hierarchical models (mixed models in SPSS). They allow estimating both the within and the between variance simultaneously. Thereby both Level 1 and Level 2 effects are calculated. Level 1 expresses effects within participants, while Level 2 expresses effects between participants.

In order to test and validate our extended measure, we conduct one experiment. In this study we experimentally manipulate the closeness of the target by either instructing participants to complete the me/not-me procedure with regard to a close other or a distant acquaintance. We also explore convergent validity of the *me/not-me* procedure by examining its association with other measures supposedly assessing interpersonal closeness. We make four predictions; in the following, we use the term *trait match* to denote the interaction term computed from the explicit ratings of the self and the other as detailed above, and the term IOS to denote the explicit graphical self-reported scale as introduced earlier.

H1. Main descriptiveness effect; descriptive adjectives will be rated faster than non-descriptive ones (Level 1)

H2. When the trait match is higher, reaction times will be lower (Level 1)

H3. A two-way interaction of trait match and manipulation: The negative influence of trait match on reaction times (H2) will be stronger in the close other condition (Level 1 x Level 2 interaction).

H4. A two-way interaction of trait match and the self-report IOS rating: The negative influence of trait match on reaction times (H2) will be stronger for others that are rated as closer on the IOS scale (Level 1 x Level 2 interaction).

It should be noted that the present study not only presents an extended validation of the original *me/not-me* procedure, but also a basic test of Aron and colleagues concept of inclusion of the other in the self. Hypotheses 3 and 4 are the core hypotheses coming out of the model developed by Aron and colleagues: Closeness is assumed to be associated with a stronger overlap of the representations of self and other. If the traits match, this should result in a stronger facilitation in the condition where a close other is described, or for those others that are rated as close on the graphical scale. With our design, we provide an experimental test of Aron's model (stronger influence of trait match for close others, H3) and a correlative test that links the me/not me effect to the graphical IOS scale.

Pilot Studies

In order to test the general effectiveness and stability of our online procedure we conducted one pilot study (Pilot Study 3) with a small sample. In addition, we ran two pilot investigations (Pilot Studies 1-2) of the match effect (H2), which can be found in the Appendix. In contrast to the main study these pilot studies were not pre-registered.

Method

Participants. Participants were sampled on MTurk requesting only U.S. Americans. Participants were removed if more than 50% of their cases were filtered out because reaction times were lower than 300 ms or exceeded 5000 ms (cf. Smith et al., 1999). In addition, only right-handed participants were retained because key assignment was not counterbalanced. The sample consisted of 55 participants, of which 42 (20 female, 22 male) were retained. Age ranged from 18 to 65 years (M = 34.07, SD = 10.83).

Materials and Design. Participants were randomly allocated to a close (n = 22) or distant (n = 20) condition. We adapted the 90 adjectives from Aron and colleagues (1991) and shortened them to 60. A balanced number of positive, neutral and negative traits was retained (see Appendix 3). During the reaction time task, participants were instructed to judge the adjectives as fast as possible as either descriptive (yes) by pressing *k* or as non-descriptive of themselves (no) by pressing *a*. Each adjective was preceded by a fixation cross and appeared on the screen for a maximum of 5000 ms. Each adjective was judged twice in order to compute profile correlations and test-retest reliabilities among the ratings, resulting in 120 trials.

The reaction time task was implemented using the QRTEngine, a Java Script extension for Qualtrics (Barnhoorn et al., 2014). Research has indicated that Java Script is an adequate instrument for assessing reaction times (de Leeuw & Motz, 2015; Reimers & Stewart, 2014).

Procedure. After being informed about the study purposes and accepting the informed consent, participants were asked to indicate a ‘close friend’ or a ‘distant acquaintance’. Participants were instructed to provide only first names or initials in order to ensure anonymity. For each participant, one of the targets was chosen randomly and used throughout the rest of the survey. Participants were instructed to rate how ‘descriptive each adjective is’ of them and the target person (i.e. close friend or distant acquaintance) on the 60

different adjectives. Ratings were completed on a 7-point scale ranging from ‘not at all’ to ‘very much’. The sets asking for ratings of the self and the target person, as well as the adjectives within each set, were presented in random order. Then, participants completed the IOS measure (Aron et al., 1992) asking which of the seven circles increasing in overlap best represented the relationship with regard to the main target.

After providing demographic information, participants were given instructions about the reaction time task. Correct understanding of instructions was probed with a question and after answering correctly the reaction time task started. Participants were also instructed to close all other browser activities and minimize other background processes.

Results

Data preparation. Reaction times lower than 300 ms or exceeding 5000 ms were removed (cf. Smith et al., 1999). A trait match index was calculated in the following way: Explicit adjective ratings for the friend and the self were recoded to ranging from -1 to 1. For all adjectives, the product of the self-rating and the target rating was computed, which then again ranged from -1 to 1. In order to make coefficients comparable among several predictors, we then rescaled the trait match index to range from -.5 to .5 by dividing it by 2. We also calculated the original index utilized by Aron and colleagues (1991) by coding ratings for the self and the target as matching that are rated as 5 to 7 or 1 to 3. Thereby, traits with ratings of 4 were lost. The descriptiveness me/not me judgment was contrast coded as -.5 (not me) and .5 (me). The self-report IOS measure was *z*-standardized. The reaction times were highly non-normal and transformed with a log-transformation. Graphical inspection revealed improved normality. In the text below, we present coefficients and their 95% confidence intervals calculated in analyses of the original untransformed latencies to ease interpretation. The respective analyses and inferential test statistics of the transformed latencies are presented in Table 1.

Main Analyses. Utilizing a Welch's *t*-test we tested for differences in IOS ratings across the two conditions. Results indicated a significant difference between the conditions, $t(33.96) = -8.16$, $p < .001$, Hedges $g = 2.56$, with IOS ratings being higher for the close friend ($M = 5.38$, $SD = 1.07$) in contrast to the distant acquaintance ($M = 2.33$, $SD = 1.24$).

The (log-transformed) reaction time (Table 1) was used as the dependent variable of a mixed model (analyzed in SPSS 23), treating individual response latencies as unit of analysis. The condition (close/distant) and the descriptiveness judgment (me/not-me) were included as factors and the trait match index as continuous predictor, together with their two-way interaction terms. Participant and adjective were entered as two cross-classifying random factors. Intercepts were allowed to vary randomly across participants and adjectives. The slope of the overlap index was allowed to vary randomly according to participants.

The model indicated an intercept of $B = 971$ [868, 1075]. Me-decisions were $B = -77$ [-114, -40] ms faster than not-me decisions (H1). The trait match index predicted reaction time negatively, $B = -228$ [-361, -95]: Higher matching with the target was associated with faster reaction times (H2). Condition did not interact significantly with the trait match index, $B = -119$ [-278, 41], though graphical inspection indicated higher matching resulting in faster reaction times in the close condition, but not in the distant one (H3, Supplementary Material Figure 3-4). Descriptiveness interacted significantly with condition, $B = -.81$ [-.132, -.29]. The interactions between descriptiveness and trait match were not significant (Table 1). Intercepts differed between participants (Wald $Z = 4.28$) and between adjectives (Wald $Z = 2.15$). The trait match index slope did not vary between participants (Wald $Z = 1.64$).

The model was repeated with the IOS ratings as a continuous predictor instead of a factor coding for condition. The intercept was $B = 984$ [902, 1066]. Me-decisions were $B = -46$ [-74, -17] ms faster than not-me decisions (H1). Again, the trait match index predicted reaction time negatively, $B = -172$ [-290, -55] (H2). The trait match index showed an

interaction with the IOS ratings, B = -88 [-168, -8], with higher matching resulting in lower reaction times for high IOS values, but not for low IOS ratings (H4, Figure 1). Intercepts differed between participants (Wald Z = 4.12), but not between adjectives (Wald Z = 1.70). The trait match index slope did not vary between participants (Wald Z = 1.44).

Additional Analyses. Because our trait match index is the product of the original explicit self and target ratings on the 60 adjectives and can therefore be regarded as their interaction, we ran a model including the self *and* the target ratings (i.e., both "main effects"), as well as the trait match index as predictors with the reaction time scores as the dependent variable. In this model the trait match index still showed a significant effect.

One might wonder whether the relationship of the trait match index to reaction times is driven by trait match ratings on the midpoint (0). These values represent judgments that are not reflective of the self or the other, which could lead to confusion and thereby also to slower response times. Together with a higher frequency of positive match values, this could result by itself in a negative correlation. We therefore repeated our main analysis with the condition, the trait match index and the descriptiveness judgment as predictors but without cases where the trait match index had a value of 0. Results indicated even stronger effects for the trait match index and its interaction with condition, ruling out this alternative interpretation.

We also repeated the main analysis with the original factor coding for matching as utilized by Aron and colleagues (1991) instead of our trait match index, and found similar yet weaker effects. The Aron index did in contrast to the trait match index not interact with condition or the IOS ratings.

Discussion

A test with a relatively small sample (total N = 42) provided first insight into the usefulness and effectiveness of our adapted paradigm. We confirmed that me ratings were

faster than not me ratings, that matching traits were judged faster, and that this effect of matching was stronger for others rated as close on the graphical scale. The pilot study did not show a significant effect of the manipulation, i.e., close others did not lead to a stronger match effect than distant others.

Based on the present results we conclude that our version of the me/not-me task is suited for an online confirmatory study and we will therefore investigate our setup in the following study.

Main Study: Basic Validation of the Me/Not-Me Reaction Task

In order to provide further validation of our version of the me/not-me task we conducted a study utilizing a similar procedure as the original experiment by Aron and colleagues (1991). We implemented the major differences and extensions presented earlier and provide a stronger test by including an adequately powered sample. The general structure followed the aforementioned pilot study including the same manipulation for the closeness of the target, as well as an explicit measure of inclusion of the other in the self (the one item graphical IOS scale) in order to link its ratings to the reaction task as done in previous studies (see Smith et al., 1999).

Method

Participants. Participants were recruited via Amazon MTurk and paid about $2.50 for completion of the task. We only requested U.S. American citizens and right-handed individuals. Participants were excluded on the basis of not fulfilling the nationality or right-handedness requirements. Individual cases (latencies) were excluded based on reaction times lower than 300ms or exceeding 5000ms (cf. Smith et al., 1999). If more than 50% of the reactions of a participant were excluded, all data from that participant was removed. Participants were also removed if more than 85% of their responses in the speeded task were yes or no answers, or if their speeded responses differed from their dichotomized explicit

ratings in more than 50% of cases. These exclusion criteria also functioned indirectly as an attention check. All of these procedures were pre-registered.

Utilizing the MLPowSim software (Brown, Lahi & Parker, 2009) we conducted a model simulation and subsequent power analysis based on the obtained pilot data (see Appendix 6). Setting $\alpha = .05$ we obtained a recommendation between 300 and 350 participants at a $\beta - 1 = .90$ level. Because we could not target the effect of interest directly, namely the interaction between the trait match score and condition, we oversampled, aiming for a total $N = 400$, or $n = 200$ in each condition. In total 401 participants were sampled. After applying all exclusion criteria, the final sample consisted of 339 individuals (168 females) ranging from 20 to 71 years of age ($M_{age} = 36.48$, $SD_{age} = 10.40$). Of the final sample, $n = 175$ participants were in the close condition, and $n = 164$ were in the distant condition.

Materials and Design. The general procedure as well as materials and designs were adapted from the pilot study unless stated differently. The task was programmed in jsPsych because of the retirement of the QRTEngine (de Leeuw, 2015). After completing the reaction time task, participants answered various scales assessing *interpersonal closeness*, including the communal sharing (CS) subscale of the MORQ ($\alpha = .93$, Haslam & Fiske, 1999), the Communal Strength scale ($\alpha = .95$, Mills et al., 2004) and the personal acquaintance measure ($\alpha = .88$, PAM, Starzyk et al., 2006).

The communal sharing subscale consists of eight items such as “The two of you are a unit: you belong together” measured on a 7-point scale anchored at “Not at all true for this relationship” and “Very true for this relationship”. The scale is generally targeted at relational aspects and how these are socially cognized (Fiske, 2004). The Communal Strength scale is thought to assess the motivation to respond to another person’s needs (Mills et al., 2004) and consists of ten items such as “How readily can you put the needs of *x* out of your mind?”. The scale is completed on a 10-point scale anchored at “nothing at all” and “extremely

large". Finally, the PAM includes 18 items (e.g., "I have known *x* for many years") measuring a number of different relational aspects such as duration, frequency, goals, physical intimacy, self-disclosure, and familiarity with the target's social network (Starzyk et al., 2006). Answers are provided on a 5-point scale anchored at "Definitely false or strongly disagree" and "Definitely true or strongly agree". All measures were completed with regard to the main target.

In addition, participants completed the perspective taking subscale ($\alpha = .90$) of the Interpersonal Reactivity Index (IRI, Davis, 1980) consisting of seven items (e.g., "I try to look at everybody's side of disagreement before I make a decision") completed on a 5-point scale anchored at "Does not describe me well" and "Describes me well". Finally, one item was presented asking "How similar is [target] to you?" on a 7-point scale ranging from *not at all* to *very much*.

Results

Data preparation. General data preparation and the calculation of indices followed the same steps as outlined in the pilot study and pre-registered.

General Modeling Strategy. The original reaction time and the log-transformed reaction time variable were used as the dependent variables during all of the analyses unless indicated otherwise. Because each participant had more than one response (in total 120), data were analyzed using mixed models (MIXED command in SPSS 23). All models included intercepts that were allowed to vary randomly across participants and adjective; the trait match index slope was set to vary randomly across participants. We consider results as statistically significant if *p* values are lower than .05.

Main analyses. As pre-registered, the same main analyses were performed as in the pilot study. Recall that coefficients and their 95% confidence intervals of reaction times are

reported for untransformed latencies to ease interpretation, but that reported significance tests are computed on log-transformed latencies.

First, a Welch's *t*-test was conducted in order to test for differences in IOS ratings across the two conditions. As expected, ratings in the close condition ($M = 4.77$, $SD = 1.46$) were significantly higher than ratings in the distant condition ($M = 2.59$, $SD = 1.41$), $t(337) = 14.01$, $p < .001$, Cohen's $d = 1.52$.

Second, a mixed model was run including condition and the descriptiveness rating as factors and the trait match index as continuous predictor, as well as all two-way interactions. The model indicated an intercept of $B = 1026$ [984, 1068] ms. Me-decisions were $B = -39$ [-54, -24] ms faster than not-me decisions (H1). The trait match index predicted reaction time negatively, $B = -110$ [-161, -60]: Higher matching with the target on a trial was associated with faster reaction times (H2). Condition interacted significantly with the trait match index, $B = -74$ [-138, -9] albeit it was a small effect (Table 2). The negative effect of the trait match index on response times was stronger in the close condition (H3, Figure 2, see also Supplementary Figures 5-6). The interaction between descriptiveness and trait match was not significant, though the interaction between descriptiveness and condition indicated a small effect (Table 2). Intercepts differed between participants (Wald $Z = 12.54$) and between adjectives (Wald $Z = 4.83$). The trait match index slope did also vary between participants (Wald $Z = 4.63$).

The model was repeated with the IOS ratings as a continuous predictor instead of a factor coding for condition (Table 2). The intercept was $B = 1035$ [1005, 1066] ms. Me-decisions were $B = -30$ [-42, -18] ms faster than not-me decisions (H1). Again, the trait match index predicted reaction time negatively, $B = -150$ [-187, -113] (H2). The trait match index showed an interaction with the IOS ratings, $B = -51$ [-85, -16], with the negative effect of trait match on response times being stronger for high IOS ratings (H4, Figure 3). The other

interactions and the IOS main effect were not significant. Intercepts differed between participants (Wald $Z = 12.53$), and between adjectives (Wald $Z = 4.83$). The trait match index slope did vary between participants (Wald $Z = 4.35$).

Convergent validity of IOS. In a separate model, we tested how the me/not-me procedure and thereby Aron and colleagues' concept of inclusion of the other in the self is related to other measures of interpersonal closeness. The (log-transformed) reaction time score was again used as dependent variable with the trait match index as continuous predictor. The different scales including MORQ CS, Communal Strength, and PAM were included as main effects and moderators of the effect of trait match, each in a separate model.[3] We predicted that the interaction of each closeness score with the trait match index would be significant, analogous to the interaction effect of IOS ratings and the trait match index.

The MORQ CS indicated a small interaction effect with the trait match index, $B = -34$ [-64, -4], as well as the PAM, $B = -35$ [-66, -4]. The Communal Strength score also indicated an interaction effect with the trait match index, $B = -59$ [-89, -30] (Table 3). For all models the trait match index indicated a main effect.

The original main model was rerun including the item asking for similarity of the target as covariate and its interaction with trait match. We did not expect this interaction to be significant. The interaction between the trait match index and similarity was significant, $B = -59$ [-92, -.25] (Figure 4), as well as the descriptiveness, $B = -31$ [-.44, -.19], and trait match main effect, $B = -155$ [-.196, -.114]. The interaction of descriptiveness with similarity was

[3] Although not explicitly mentioned in the pre-registered report the MORQ CS, Communal Strength, PAM, Similarity, and Perspective Taking items were also standardized in order to resemble the original model with the IOS rating and to combat multicollinearity problems.

also significant, *B* = 11 [.5, .21], although not when using the log-transformed reaction time score as DV (Table 4).

Finally, in a different model, perspective taking was added to the original main model as a covariate, and entering two-way interactions and a three-way interaction with the other predictors. We predicted that general trait perspective taking would amplify the interaction effect of IOS and the trait match index. We observed only a main descriptiveness effect as in the original model, *B* = -35 [-58, -12], but not the trait match main effect and none of the two-way interactions were significant except for the interaction between the trait match index and the IOS rating, *B* = -30 [-47, -13]. The three-way interaction with perspective taking, IOS rating and the trait match index was not significant either, *B* = -7 [-22, 8]. The IOS rating indicated a small main effect, *B* = 15 [.3, 31] (Table 4).

Additional analyses. As pre-registered, the same additional analyses were repeated as in the pilot study. First, the original ratings of self and the other target were utilized with the trait match index as continuous predictors in a model with the reaction time as dependent variable. We expected the trait match index to still be negatively associated with reaction times. Results indicated that the interaction between self and other ratings (trait match index) was still negatively associated with reaction times, *B* = -65 [-77, -53]. In addition, self ratings indicated a main effect, *B* = -30 [-39, -21] (Appendix Table 1).

Second, in order to ascertain that the effect of the trait match index was not caused by midpoint values (0), these were excluded for the trait match index in our main model. We expect the same effects as before. We observed a similar main effect of the trait match index after excluding all midpoints (Appendix Table 1).

Finally, both models with the manipulation and the IOS scale were repeated with the original Aron score replacing our trait match index. We expected similar direction though

considerably weaker effects. In general, similar effects were obtained using the original Aron score (see Appendix Table 2).

We also repeated the models using only the first set of 60 responses to judge whether future applications of the paradigm can be shortened. The general effects were similar, though the interaction between trait match index and the type of condition was not significant in the first model (see Appendix Table 3).

Finally, we calculated measures of association[4] for each participant for the set of adjectives by getting the association for each pair of adjectives. This procedure calculated test-retest reliabilities for the descriptiveness judgments. The overall set of adjectives across all participants indicated a Phi coefficient of $\phi = .78$.

General Discussion

The present study was designed in order to replicate and extend the *me/not-me* response time task targeted at implicitly measuring the concept of inclusion of the other in the self (IOS). We adapted the original paradigm with regard to testing environment (moving it online) and introduced changes regarding procedure and assessment that were assumed to enhance its predictive power and reliability. Nevertheless, the basic idea of Aron et al.'s (1991) procedure was retained, and we consider the current study an extended replication attempt of their classic finding. What we found were three replicated effects:

First, we successfully replicated the general *descriptiveness effect*, the tendency to judge traits faster as descriptive than as non-descriptive (H1). Similarly, we observed that traits exhibiting higher match with the target were rated faster in contrast to traits showing decreased match with the target (H2). Judging traits faster as descriptive than non-descriptive for the self has been observed on a number of occasions (Aron et al., 1991; Aron & Fraley,

[4] In the pre-registered report these were mistakenly denoted as *profile correlations*.

1999; Kuiper, 1981). In the presented studies (including our pilot studies) we have provided additional evidence for this effect. Replicating the descriptiveness pattern lends also further support to the possibility of running response latency tasks online. However, the focus of the present investigation is not on providing an in depth and detailed discussion of this effect.

Second, we replicated earlier findings (Aron et al., 1991; Aron & Fraley, 1999; Smith et al., 1999) by observing that traits are judged more quickly as descriptive or non-descriptive for the self if they show increased trait match with the other target (*distinctiveness effect*). By comparing our operationalization of the trait match between the self and the target (*trait match*) with past measurement we provided a more nuanced picture of this relationship. We did not exclude cases in which ratings do not overlap, but merely treated the trait match index as the interaction between ratings for the self and the target. Supplementary analyses confirmed that the negative effect for the trait match index on response times was not influenced by ratings of zero and thus a match index of zero, where one might argue that the lack certainty leads to a confusion and slower reaction times.

Third, confirming the central hypothesis, that effect of trait match was moderated. It is important to realize that the general facilitating effect of trait match on response time is not indicative of the IOS concept yet. Instead, because IOS is a conceptualization of *closeness*, the influence of trait match should be enhanced for close targets. In line with our predictions and Aron's model, the influence of match was moderated by the experimental manipulation (close vs. distant, H3). The negative effect of the trait match index on response times was stronger for participants choosing a close target in contrast to participants completing the study with regard to a distant target. The effect was statistically significant with a p value of .013, which was within the limits of our preregistration, but should probably still be interpreted with caution (G. M. Sullivan & Feinn, 2012). In addition, we also observed a conceptually similar, and stronger interaction of the trait match index with the IOS self-report

rating (H4). Increased ratings on the IOS rating resulted in a stronger facilitating effect of the trait match index on response times, similar to past findings (e.g., Smith et al., 1999).

A closer look at the interaction effects reveals a more detailed picture of the mechanisms at play. Considering the interaction of trait match with the experimental manipulation, response times for high trait match traits did not differ between participants in the close or distant condition. On the other hand, low trait match resulted in increased (i.e., slower) response times in the close condition in comparison to the distant condition. This effect is in line with Aron and colleagues' (1991) original interpretation that low overlap between two close individuals results in confusion and inhibited responses and thereby longer response times. This argument is supported by the observation that low overlap in the distant condition resulted in significantly faster response times. There existed no confusion because the distant relationship is thought to have low overlap in the first place.

The same direction was also observed for the interaction between the trait match index and the IOS rating. However, in addition we also observed that high overlap resulted in faster response times in the distant in comparison to the close condition. This finding seems counterintuitive to the general prediction of the IOS concept. The two experimental conditions successfully induced different levels of IOS ratings, still it might be that targets in the distant condition were not *distant* enough. The original studies by Aron et al. (1991) and Smith et al. (1999) used a controlled neutral target by employing a media personality (i.e., Bill Cosby). In fact, in our study the overall trait match index was on average still positive in the distant condition.[5] Finally, the difference between high overlap responses times was considerably smaller and should be interpreted with caution.

---

[5] The trait match mean in the close condition was, $M = .11$ ($SD = .23$), while it was smaller in the distant condition, $M = .07$ ($SD = .22$).

We observed a slightly different picture in the main study in contrast to the pilot study. While we observed no difference between response times for high and low trait match for low IOS ratings, high trait match was rated faster for low IOS ratings in the main study. Considering the low power and the exploratory nature of our pilot study this difference seems negligible.

Similarity and IOS

In a series of studies, Mashek et al. (2003) provided evidence that the general mechanism of IOS is distinct from a general tendency of increased similarity. Therefore, we included a one-item measure of similarity with the target in order to assess its effect on the trait match index. Contrary to our predictions, we observed a significant two-way interaction between the two variables. As with IOS ratings, the negative effect of the trait match index on response times was stronger for high ratings of similarity. High overlap resulted in faster response times for individuals high on similarity in contrast to scores low on similarity. Considering this evidence, we cannot conclude that similarity is completely distinct from the IOS process. Further research should explore this link in more depth.

Convergent validity of IOS

We included a number of scales designed to assess the concept of *interpersonal closeness*. In general, all three measures, the MORQ CS, PAM and the communal strength scale, significantly interacted with the trait match index and could thus be considered equivalent to the IOS ratings. The negative effect of the trait match index was stronger for higher ratings of these measures. However, the magnitude of effects differed across these measures. While the effect was comparably smaller for the MORQ CS and the PAM, the communal strength measure exhibited a stronger effect. A possible post hoc explanation might be that both the PAM and the MORQ CS assess a more varied construct of *interpersonal closeness*, including cognitive, motivational and affective components. On the

other hand, the communal strength scale focuses on motivational tendencies. Obviously, this unexpected observation should be replicated and backed up by additional measures before taking it for granted.

Although the concept of perspective taking has been associated with IOS on earlier occasions (Galinsky & Moskowitz, 2000), we did not observe any evidence that perspective taking moderated the link between the trait match index and the IOS item.

The Me/Not-Me Procedure

Considering the evidence of the present study, the me/not-me procedure seems to represent a valid method of measuring inclusion of the other in the self implicitly. We did not only observe an effect of our experimental manipulation on the trait match index, but also by the one item IOS rating. Low overlap between individuals resulted in slower response times in the close condition in contrast to the distant target condition. Convergent validity was provided with three constructs measuring different aspects of *interpersonal closeness*. Contrary to our predictions and the literature on IOS increased similarity also resulted in facilitation and faster response times of ratings with high overlap. Higher inclusion of the other in the self fosters an inhibition or confusion effect of traits that do not overlap. The more the other target is part of the self, the harder it gets to distinguish between the other and the self (Aron & Fraley, 1999; Aron et al., 2013). In contrast, it seems that similarity facilitates the recognition of traits that do overlap in close relationships.

The present investigation suggested that our adaptions and extensions of the original me/not-me task were improving methodological and statistical power, as well as ease of application. All files including the jsPsych script and our syntax preparation file are uploaded to our project page (https://osf.io/pazfj/). The present results suggested that the paradigm might be applicable with reducing the response latency task to 60 items, thereby facilitating online implementation, but 120 items would assure higher statistical power. We also hope

that the task might be used as a dependent variable in experimental research that tries to assess IOS implicitly. To date this powerful approach has been neglected in the literature on interpersonal closeness.

The me/not me procedure has a double purpose: On the one hand, it tests a hypothesis derived from the notion of including others in the self and the self-expansion model (Aron & Aron, 1986). On the other hand, it is seen as a measurement of that construct. Our results support both, but it is clear that we did not provide any evidence that applying this complicated measure has any advantage over applying the one-item graphical scale (which was in fact used to validate the result of the procedure). It remains a task for future work to identify when that procedure can tell us more about social-cognitive processes than the graphical measure itself; the rich work on implicit measures of attitudes and stereotypes can be a guide to that.

Limitations

The present approach also includes a number of limitations. Although the general procedure was successfully implemented in an online environment as done in previous attempts (de Leeuw & Motz, 2015; Schubert et al., 2013), the exclusion rate with about 20% was considerably high. These exclusions were mostly based on attention checks pointing at the idea that such checks should be commonly used when recruiting participants on MTurk (Peer, Vosgerau, & Acquisti, 2014). In addition, the diversity of possible settings differs substantially from controlled laboratory settings.

The online environment presented further complications such as the variability in browser compatibility or plug-in configurations. Our script programmed in jsPsych (de Leeuw, 2015) did not run properly using the Internet Explorer. Limiting participants to a specific browser type does undermine generalizability of our research. The browser variability also poses additional constraints on testing and designing the research protocol.

Finally, the present study did only focus on a sample of US Americans recruited on MTurk. In order to derive broader conclusions about the applicability and functionality of the me/not-me task it is important to test diverse populations. Considering that the task relies on few cultural diverse components this might not pose strong invariance concerns.

The present study presented a replication and extension of the neglected implicit *me/not-me* response latency measure of IOS. We have employed a number of adaptions that increased the methodological and statistical power, as well as the simplification of the paradigm. Our experiences suggest that implementation of response latency measures in an online environment presents a strong, promising, but not yet straightforward path. Our general evidence indicates that the me/not-me procedure is a valid implicit measure of inclusion of the other in the self. Nevertheless, its mechanisms, as well as the IOS concept itself, still have to be explored in more depth.

References

Agnew, C. R., van Lange, A., Rusbult, C. E., & Langston, C. A. (1998). Cognitive interdependence: Commitment and the mental representation of close relationships. *Journal of Personality and Social Psychology*, *74*(4), 939–954. https://doi.org/10.1037/0022-3514.74.4.939

Anderson, N. H. (1968). Likeableness Ratings of 555 Personality-Trait Words. *Journal of Personality and Social Psychology*, *9*(3), 272–279. https://doi.org/10.1037/h0025907

Aron, A., & Aron, E. N. (1986). *Love and the expansion of self: Understanding attraction and satisfaction* (Vol. x). New York, NY, US: Hemisphere Publishing Corp/Harper & Row Publishers.

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*(4), 596–612. https://doi.org/10.1037/0022-3514.63.4.596

Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationships as including other in the self. *Journal of Personality and Social Psychology*, *60*(2), 241.

Aron, A., & Fraley, B. (1999). Relationship closeness as including other in the self: Cognitive underpinnings and measures. *Social Cognition*, *17*(2), 140.

Aron, A., Lewandowski Jr, G. W., Mashek, D., & Aron, E. N. (2013). The self-expansion model of motivation and cognition in close relationships. In J. A. Simpson & L. Campbell (Eds.), *The Oxford handbook of close relationships* (pp. 90–115). Oxford, UK: Oxford University Press.

Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and Social Psychology Bulletin*, *23*(4), 363–377.

Barnhoorn, J. S., Haasnoot, E., Bocanegra, B. R., & van Steenbergen, H. (2014). QRTEngine: An easy solution for running online reaction time experiments using Qualtrics. *Behavior Research Methods*, 1–12.

Clark, M. S., & Lemay, E. P. (2010). Close relationships. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (Vol. 2, pp. 898–940). Hoboken, NJ: Wiley.

Clark, M. S., & Mills, J. (1979). Interpersonal attraction in exchange and communal relationships. *Journal of Personality and Social Psychology*, *37*(1), 12–24. https://doi.org/10.1037/0022-3514.37.1.12

Craemer, T. (2014). Implicit Closeness to Blacks, Support for Affirmative Action, Slavery Reparations, and Vote Intentions for Barack Obama in the 2008 Elections. *Basic and Applied Social Psychology*, *36*(5), 413–424. https://doi.org/10.1080/01973533.2014.935380

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research. *PLoS ONE*, *8*(3), e57410. https://doi.org/10.1371/journal.pone.0057410

Dalsky, D., Gohm, C. L., Noguchi, K., & Shiomura, K. (2008). Mutual Self-Enhancement in Japan and the United States. *Journal of Cross-Cultural Psychology*, *39*(2), 215–223. https://doi.org/10.1177/0022022107313863

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, *10*, 85.

de Leeuw, J. R. de, & Motz, B. A. (2015). Psychophysics in a Web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods*, *48*(1), 1–12. https://doi.org/10.3758/s13428-015-0567-2

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12.

Fazio, R. H., & Olson, M. A. (2003). Implicit Measures in Social Cognition Research: Their Meaning and Use. *Annual Review of Psychology*, *54*(1), 297–327. https://doi.org/10.1146/annurev.psych.54.101601.145225

Fiske, A. P. (2004). Relational models theory 2.0. In N. Haslam (Ed.), *Relational Models Theory: A Contemporary Overview* (pp. 3–25). Mahwah, NJ: Erlbaum.

Gächter, S., Starmer, C., & Tufano, F. (2015). Measuring the Closeness of Relationships: A Comprehensive Evaluation of the "Inclusion of the Other in the Self" Scale. *PLOS ONE*, *10*(6), e0129478. https://doi.org/10.1371/journal.pone.0129478

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*(4), 708–724. https://doi.org/10.1037/0022-3514.78.4.708

Haslam, N., & Fiske, A. P. (1999). Relational models theory: A confirmatory factor analysis. *Personal Relationships*, *6*(2), 241-250.

Hodges, S. D., Sharp, C. A., Gibson, N. J. S., & Tipsord, J. M. (2013). Nearer My God to Thee: Self–God Overlap and Believers' Relationships with God. *Self and Identity*, *12*(3), 337–356. https://doi.org/10.1080/15298868.2012.674212

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods*, *39*(1), 101–117. https://doi.org/10.3758/BF03192848

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Kuiper, N. A. (1981). Convergent Evidence for the Self as a Prototype The "Inverted-U RT Effect" for Self and Other Judgments. *Personality and Social Psychology Bulletin*, *7*(3), 438–443. https://doi.org/10.1177/014616728173012

Le, B., Moss, W. B., & Mashek, D. (2007). Assessing Relationship Closeness Online Moving From an Interval-Scaled to Continuous Measure of Including Others in the Self. *Social Science Computer Review*, *25*(3), 405–409. https://doi.org/10.1177/0894439307297693

Mashek, D. J., Aron, A., & Boncimino, M. (2003). Confusions of Self With Close Others. *Personality and Social Psychology Bulletin*, *29*(3), 382–392. https://doi.org/10.1177/0146167202250220

Mills, J., Clark, M. S., Ford, T. E., & Johnson, M. (2004). Measurement of communal strength. *Personal Relationships*, *11*(2), 213–230.

Otten, S., & Epstude, K. (2006). Overlapping Mental Representations of Self, Ingroup, and Outgroup: Unraveling Self-Stereotyping and Self-Anchoring. *Personality and Social Psychology Bulletin*, *32*(7), 957–969. https://doi.org/10.1177/0146167206287254

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods*, *46*(4), 1023–1031. https://doi.org/10.3758/s13428-013-0434-y

Reimers, S., & Stewart, N. (2014). Presentation and response timing accuracy in Adobe Flash and HTML5/JavaScript Web experiments. *Behavior Research Methods*, *47*(2), 309–327. https://doi.org/10.3758/s13428-014-0471-1

Ross, M. J., Mueller, J. H., & De La Torre, M. (1986). Depression and trait distinctiveness in the self-schema. *Journal of Social and Clinical Psychology*, *4*(1), 46.

Sande, G. N., Goethals, G. R., & Radloff, C. E. (1988). Perceiving one's own traits and others': The multifaceted self. *Journal of Personality and Social Psychology*, *54*(1), 13–20. https://doi.org/10.1037/0022-3514.54.1.13

Schubert, T. W., Murteira, C., Collins, E. C., & Lopes, D. (2013). ScriptingRT: A Software Library for Collecting Response Latencies in Online Studies of Cognition. *PLoS ONE*, *8*(6), e67769. https://doi.org/10.1371/journal.pone.0067769

Schubert, T. W., & Otten, S. (2002). Overlap of Self, Ingroup, and Outgroup: Pictorial Measures of Self-Categorization. *Self and Identity*, *1*(4), 353–376. https://doi.org/10.1080/152988602760328012

Schubert, T. W., Zickfeld, J. H., Seibt, B., & Fiske, A. P. (2016). Moment-to-Moment Changes in Being Moved Match Changes in Perceived Closeness, Weeping, Goosebumps, and Warmth: Time Series Analyses. *Cognition and Emotion*. https://doi.org/10.1080/02699931.2016.1268998

Schultz, P. W. (2002). Inclusion with nature: The psychology of human-nature relations. In P. Schmuck & P. W. Schultz (Eds.), *Psychology of sustainable development* (pp. 61–78). Dordrecht: Kluwer Academic Publishers.

Slatcher, R. B. (2010). When Harry and Sally met Dick and Jane: Creating closeness between couples. *Personal Relationships*, *17*(2), 279–297. https://doi.org/10.1111/j.1475-6811.2010.01276.x

Smith, E. R., Coats, S., & Walling, D. (1999). Overlapping mental representations of self, in-group, and partner: Further response time evidence and a connectionist model. *Personality and Social Psychology Bulletin*, *25*(7), 873–882.

Smith, E. R., & Henry, S. (1996). An in-group becomes part of the self: Response time evidence. *Personality and Social Psychology Bulletin*, *22*(6), 635–642.

Starzyk, K. B., Holden, R. R., Fabrigar, L. R., & Macdonald, T. K. (2006). The Personal Acquaintance Measure: a tool for appraising one's acquaintance with any person. *Journal of Personality and Social Psychology*, *90*(5), 833–847. https://doi.org/10.1037/0022-3514.90.5.833

Sullivan, G. M., & Feinn, R. (2012). Using Effect Size—or Why the P Value Is Not Enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. https://doi.org/10.4300/JGME-D-12-00156.1

Sullivan, M. P., & Venter, A. (2005). The Hero Within: Inclusion of Heroes into the Self. *Self and Identity*, *4*(2), 101–111. https://doi.org/10.1080/13576500444000191

Tropp, L. R., & Wright, S. C. (2001). Ingroup identification as the inclusion of ingroup in the self. *Personality and Social Psychology Bulletin*, *27*(5), 585–600.

van Veelen, R., Otten, S., Cadinu, M., & Hansen, N. (2015). An Integrative Model of Social Identification Self-Stereotyping and Self-Anchoring as Two Cognitive Pathways. *Personality and Social Psychology Review*, 1088868315576642.

Waugh, C. E., & Fredrickson, B. L. (2006). Nice to know you: Positive emotions, self–other overlap, and complex understanding in the formation of a new relationship. *The Journal of Positive Psychology*, *1*(2), 93–106. https://doi.org/10.1080/17439760500510569

Tables

Table 1

Unstandardized coefficients and inferential statistics for the Pilot Study 3. The log-transformed reaction time score is utilized as dependent variable.

	B	95% CI	df	*t*	*p*
Model 1					
Me/Not Me	-.028	[-.040, -.015]	3045	-4.40	<.001
Trait Match	-.083	[-.129, -.036]	84	-3.54	.001
Condition	-.021	[-.082, .040]	40	-.68	.498
Me/Not Me X Trait Match	.022	[-.021, .064]	2179	1.01	.314
Me/Not Me X Condition	-.026	[-.043, -.009]	3409	-3.00	.003
Trait Match X Condition	-.031	[-.087, .026]	37	-1.09	.282
Model 2					
Me/Not Me	-.018	[-.028, -.009]	1669	-3.73	<.001
Trait Match	-.069	[-.109, -.028]	126	-3.37	.001
IOS	-.004	[-.035, .027]	37	-.25	.807
Me/Not Me X Trait Match	.019	[-.024, .062]	1718	.87	.386
Me/Not Me X IOS	-.006	[-.015, .003]	2164	-1.31	.192
Trait Match X IOS	-.031	[-.060, -.003]	34	-2.23	.032

Note. Me/Not Me was coded as -.5 (non-descriptive), .5 (descriptive); condition was coded as 0 (distant), 1 (close).

Table 2

Unstandardized coefficients and inferential statistics for the Main Study. The log-transformed reaction time score is utilized as dependent variable.

	B	95% CI	df	*t*	*p*
Model 1					
Me/Not Me	-.017	[-.022, -.012]	32335	-6.98	<.001
Trait Match	-.046	[-.064, -.029]	314	-5.23	<.001
Condition	.008	[-.013, .030]	354	.78	.436
Me/Not Me X Trait Match	.011	[-.003, .025]	35679	1.59	.111
Me/Not Me X Condition	.006	[.0002, .012]	36995	2.03	.042
Trait Match X Condition	-.029	[-.051, -.006]	244	-2.51	.013
Model 2					
Me/Not Me	-.014	[-.017, -.010]	26748	-7.40	<.001
Trait Match	-.063	[-.075, -.050]	394	-9.76	<.001
IOS	.010	[-.0001, .02]	356	1.96	.051
Me/Not Me X Trait Match	.014	[-.0003, .028]	35193	1.91	.056
Me/Not Me X IOS	-.0001	[-.003, .003]	36709	-.06	.951
Trait Match X IOS	-.020	[-.030, -.009]	220	-3.59	<.001

Note. Me/Not Me was coded as -.5 (non-descriptive), .5 (descriptive); condition was coded as .5 (distant), -.5 (close).

Table 3

Unstandardized coefficients and inferential statistics for the other interpersonal *closeness* measures in the Main Study. The log-transformed reaction time score is utilized as dependent variable.

	B	95% CI	df	*t*	*p*
Model 1					
PAM	.008	[-.002, .018]	339	1.50	.134
Trait Match	-.053	[-.065, -.042]	244	-9.35	<.001
PAM X Trait Match	-.012	[-.023, -.001]	214	-2.18	.030
Model 2					
CS	.006	[-.004, .017]	339	1.17	.243
Trait Match	-.053	[-.064, -.042]	245	-9.52	<.001
CS X Trait Match	-.023	[-.034, -.013]	210	-4.41	<.001
Model 3					
MORQ CS	.006	[-.004, .017]	338	1.15	.251
Trait Match	-.053	[-.064, -.042]	243	-9.29	<.001
MORQ CS X Trait Match	-.013	[-.024, -.002]	208	-2.39	.018

Table 4

Unstandardized coefficients and inferential statistics for the models including perspective taking and similarity in the Main Study. The log-transformed reaction time score is utilized as dependent variable. All two-way interactions except Trait Match x IOS were non-significant in Model 2 and are not represented.

	B	95% CI	df	*t*	*p*
Model 1 (Similarity)					
Me/Not Me	-.014	[-.018, -.010]	21839	-7.08	<.001
Trait Match	-.064	[-.078, -.050]	330	-9.19	<.001
Similarity	.002	[-.010, .013]	299	.27	.789
Me/Not Me X Trait Match	.01	[-.004, .026]	30505	1.44	.150
Me/Not Me X Similarity	.003	[-.0004, .006]	31267	1.74	.082
Similarity X Trait Match	-.023	[-.034, -.011]	176	-3.85	<.001
Model 2 (Perspective Taking)					
Me/Not Me	-.012	[-.019, -.005]	35100	-3.35	.001
Trait Match	-.023	[-.049, .002]	235	-1.81	.072
IOS	.006	[.0002, .012]	349	2.02	.044
Perspective Taking (PT)	-.016	[-.038, .006]	335	-1.36	.175
Trait Match X IOS	-.011	[-.017, -.005]	216	-3.50	.001
Trait Match X IOS X PT	-.003	[-.008, .003]	197	-1.03	.302

Note. Me/Not Me was coded as -.5 (non-descriptive), .5 (descriptive); condition was coded as .5 (distant), -.5 (close).

Figures

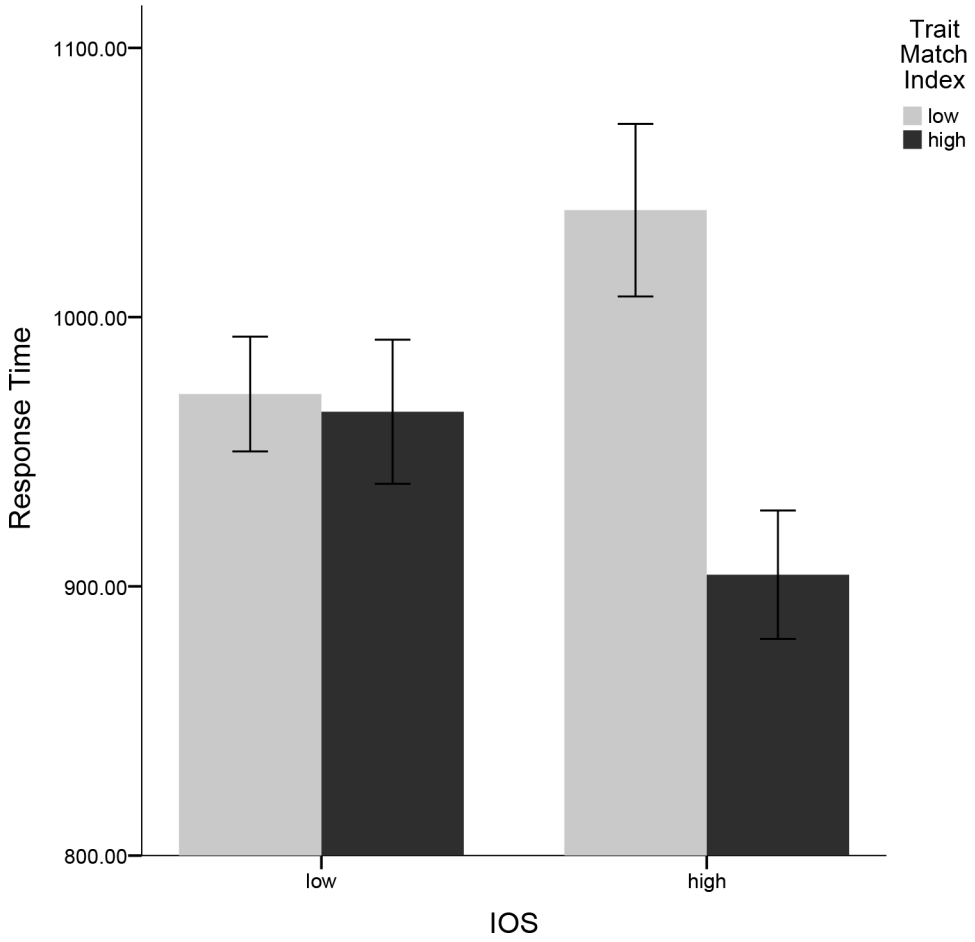

Figure 1. Bar graph of interaction between the trait match index and the IOS rating in the Pilot Study 3. Response time is given in ms. Error bars represent 95% confidence intervals and are only suitable for between subject comparisons. IOS is dichotomized by median split.

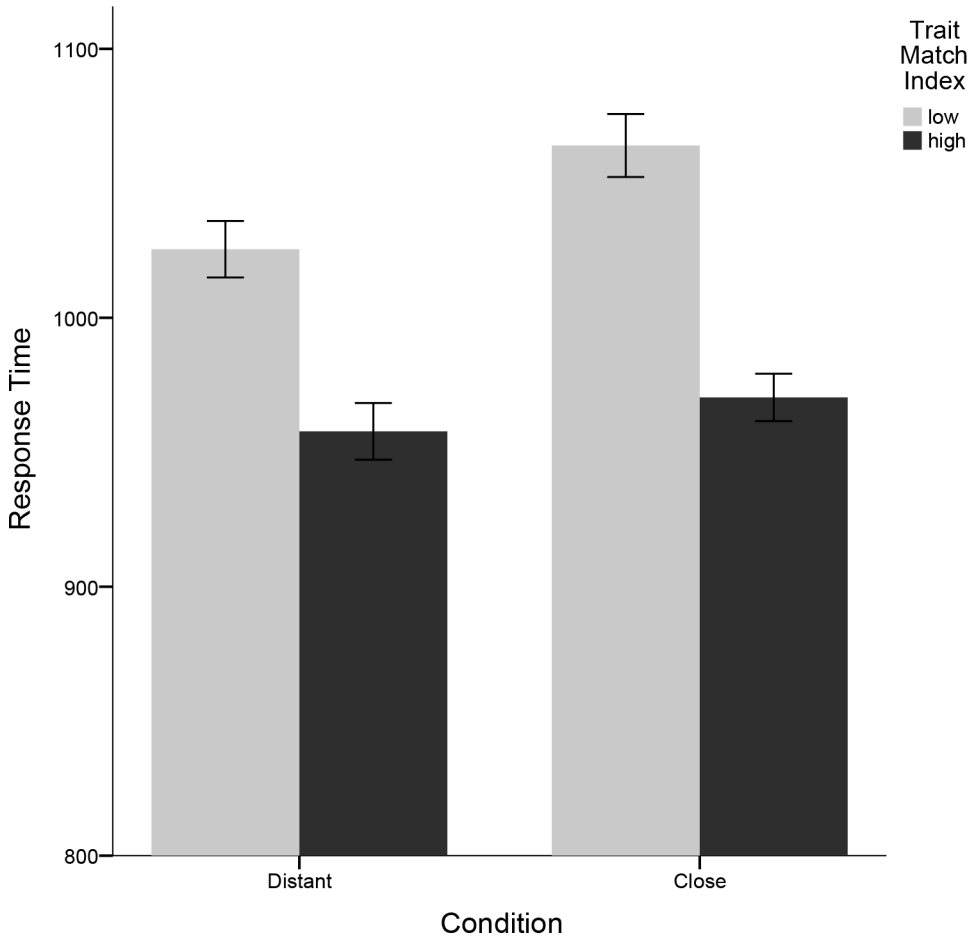

Figure 2. Bar graph of interaction between the trait match index and the experimental condition in the Main Study. Response time is given in ms. Error bars represent 95% confidence intervals and are only suitable for between subject comparisons. Trait match is dichotomized by median split.

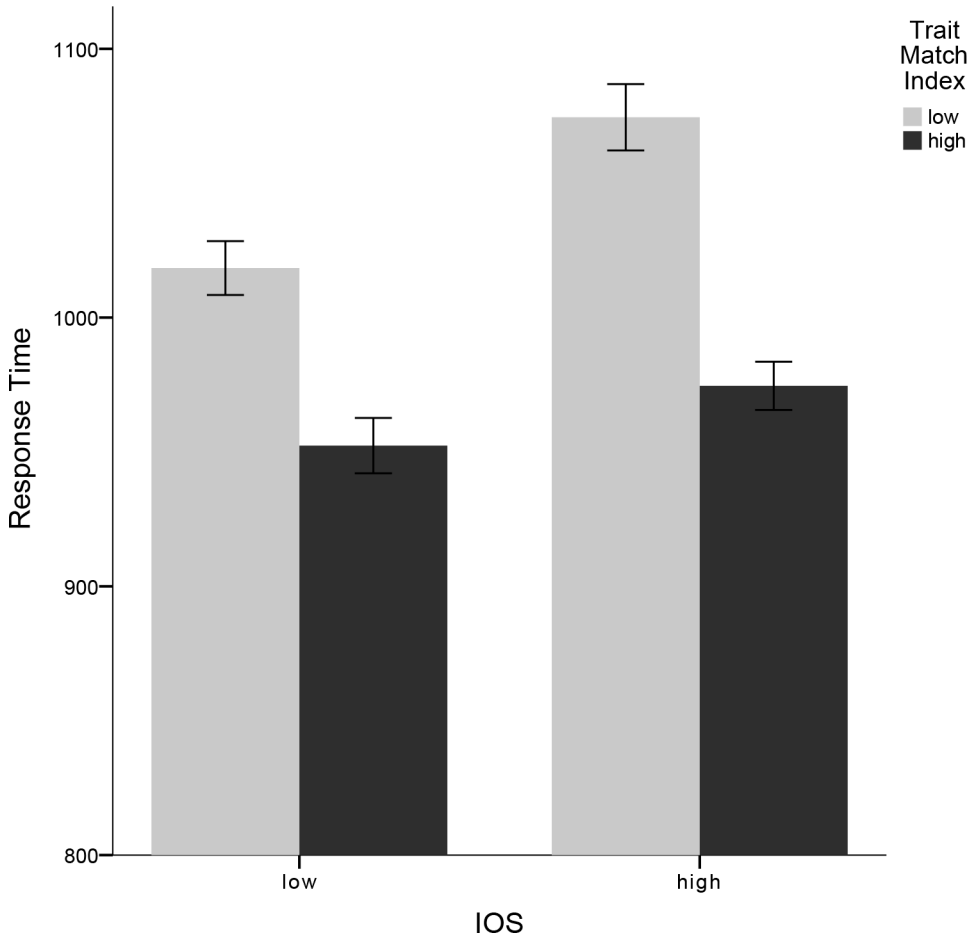

Figure 3. Bar graph of interaction between the trait match index and the IOS rating in the Main Study. Response time is given in ms. Error bars represent 95% confidence intervals and are only suitable for between subject comparisons. IOS and trait match are dichotomized by median split.

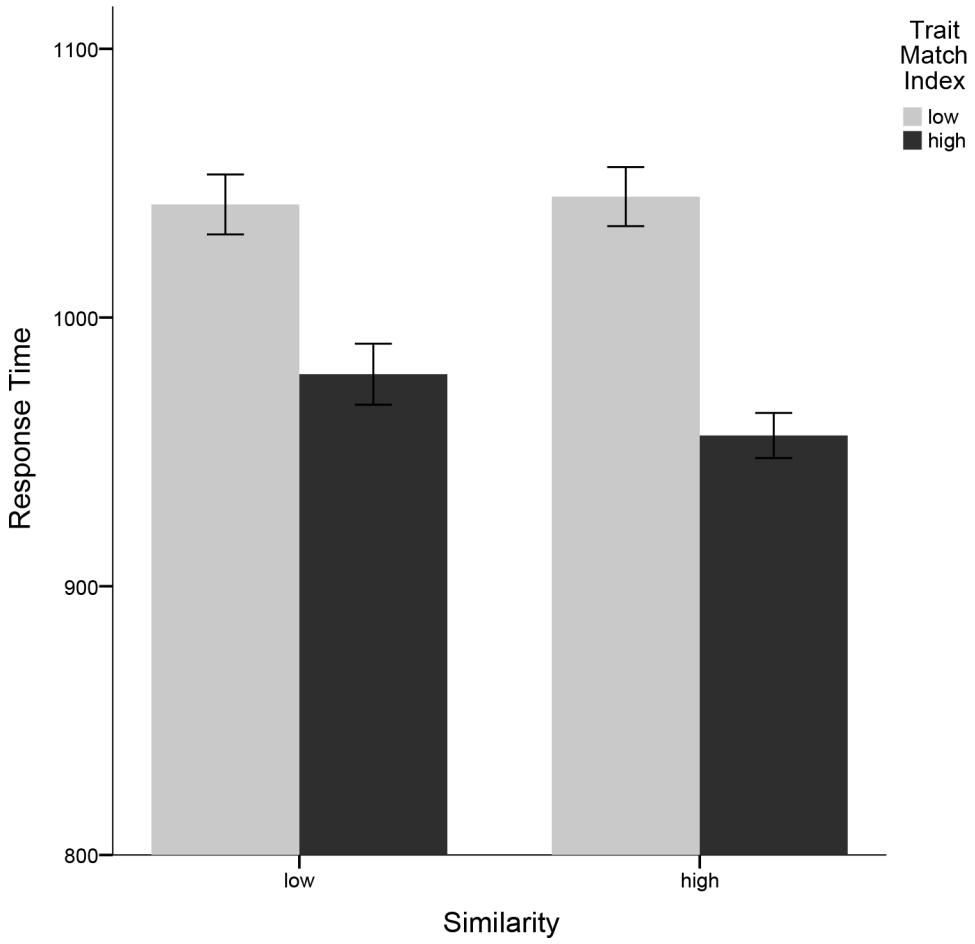

Figure 4. Bar graph of interaction between the trait match index and the similarity rating in the Main Study. Response time is given in ms. Error bars represent 95% confidence intervals and are only suitable for between subject comparisons. Similarity and trait match are dichotomized by median split.