Self and Identity, 2016
Vol. 15, No. 1, 1–18, http://dx.doi.org/10.1080/15298868.2015.1072580

Conceptual Organization of Self-representation: A Self-similarity Heuristic for Novel Person Representations

A. Nicole LeBarr, John G. Grundy, Mohsin Ali, and Judith M. Shedden

Department of Psychology, Neuroscience, & Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON, Canada L8S4K1

Implicit measures have revealed that cognitive representations of familiar individuals share associations with self-concept; however, this has yet to be established for novel individuals. We examined how self-similarity affects representation of information learned about new individuals. A novel version of the implicit association test (IAT), the self-similarity IAT, was developed to estimate the extent to which cognitive representations of new self-similar and self-dissimilar individuals are associated with self-representation. Categorization was faster when the self-similar individual was paired with self, not only for trait words related to the novel individuals, but also for unrelated demographic information pertaining only to self. This provides the first evidence using an implicit task that self-similarity may act as a heuristic for creating representations of new individuals.

Keywords: Self; Implicit association test; Heuristic; Association strength; Representation.

The self as a construct exerts top-down influence over perception (e.g., Gillihan & Farah, 2005; Markus, 1977; Markus, Smith, & Moreland, 1985). Self-conceptualization derives from one's idiosyncratic experiences with the world, leading to vastly different biases with which to organize incoming information. These biases apply to many types of input, including information acquired through social interactions. The present paper examines the processes through which novel individuals become incorporated into one's cognitive network, hypothesizing that similarity to self serves as a readily available heuristic by which this is done. Most studies that examine how self-concept affects representations of novel individuals have relied on explicit measures (e.g., judgments about the new person), whereas studies using implicit measures have focused on familiar rather than novel individuals. Therefore, the extant literature does not address whether we efficiently incorporate representations of new individuals according to self-similarity or whether the

Received 18 July 2014; accepted 10 July 2015; first published online 22 September 2015.

We thank Joey Legere for implementing the ReAL model analysis. We also thank Dr Christian Jordan and two anonymous reviewers for their helpful comments.

No potential conflict of interest was reported by the authors.

This research was funded by a Natural Sciences and Engineering Research Council of Canada (NSERC) to JMS [grant number 170353-2007].

Correspondence should be addressed to A. Nicole LeBarr, Department of Psychology, Neuroscience, & Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON, Canada L8S4K1. E-mail: lebarran@mcmaster.ca

effect of self-similarity extends only to deliberative processing of new individuals; the answer to this question may be revealed by the development and use of an implicit measure of self-representation. To address this idea, we designed a new version of the implicit association test (IAT), the self-similarity IAT, to examine the cognitive associations that develop when we encounter novel individuals that vary in their similarity to self.

Smith and Zarate's (1992) model of social perception suggests that we form unique exemplars to encode others and that these are incorporated into the cognitive network based on similarity to other previously encoded exemplars. Exemplars, in this context, refer to a cognitive representation of a person, including perceptual attributes that have been encoded for that person, which incorporate experiences with the person and biases of the perceiver (Smith & Zárate, 1992). In the present study, we show evidence for a self-heuristic consistent with this model, whereby we readily perceive self-similarity (i.e., having similar personality traits, interests, values, etc.) and automatically represent newly encountered individuals accordingly. We argue that representations of new individuals are organized according to degree of self-similarity; information about self-similar individuals is more strongly associated with one's self-concept than information about self-dissimilar individuals. Importantly, we would like to claim that the self-heuristic leads to a representation of the novel individual that not only affects the strength of association with concepts similar and dissimilar to self, but also affects processing of information that is distinct from that which determines the self-similarity. For example, suppose we know nothing about Kaleem, but we notice that he loves the same indie rock group that we do. This perception of similarity between ourselves and Kaleem may allow us to automatically form an association between representations of self and Kaleem, which could lead us to generalize some of our other attributes to him, even though they are completely unrelated to his music taste or our actual experience with him. This leads to the specific hypothesis tested in this paper, that this organization will bias categorization of information about self and others, including information that is both related and unrelated to the concepts that determine the initial self-similar or self-dissimilar associations.

Our concept of others is strongly related to our sense of self (see Decety & Sommerville, 2003; Markus & Wurf, 1987, for reviews), and therefore self-conceptualization should be viewed as an important mediator in social perceptions. For instance, research aimed at understanding social perceptions involving self and familiar individuals finds a positive correlation between participants' desire to be rated favorably on a personality dimension and the importance of this dimension when making judgments of significant others and acquaintances (Lewicki, 1983). Smith, Coats, and Walling (1999) have proposed a connectionist model of self–other overlap, hypothesizing that direct links exist between self-concept and representation of significant others. The model helps explain the finding that when participants make judgments about whether various traits are self-descriptive, performance is faster if those traits are also characteristic of a significant other (Aron, Aron, Tudor, & Nelson, 1991; Coats, Smith, Claypool, & Banner, 2000; Smith & Henry, 1996). This facilitation effect also occurs when the judgments are about the other instead of about the self; specifically, performance is facilitated on judgments of whether a relationship partner possesses a particular trait when there is a match with self (Smith et al., 1999). The observation of faster responses to these matches (i.e., both self and other have or do not have a trait) has been interpreted as evidence of an implicit process because traits possessed by the significant other affect reaction time for judgments that do not explicitly refer to that significant other. This line of evidence informs us that our cognitive representations of familiar others are structured by their similarity to self; however, it does not shed light on the processes that are involved in the formation of representations of novel individuals.

Studies examining the effect of self-relevant information on social perceptions of unfamiliar or newly encountered individuals have relied on more explicit self-report measures that imply the proposed organization but do not characterize it at a cognitive level. For example, when tasked to acquire information about a new individual, participants who identified strongly with a particular aspect of self-concept or self-schema (e.g., introversion, masculinity) chose to ask questions eliciting information related to that schema (Fong & Markus, 1982). Participants are also more likely to exaggerate differences and therefore make more extreme judgments of others for traits that are considered important to their sense of self (Tajfel & Wilkes, 1964). Moreover, participants who viewed bogus attitudinal questionnaires that were ostensibly completed by other individuals and varied in attitudinal similarity to themselves rated self-similar individuals more positively than self-dissimilar ones over several measures such as likeability, intelligence, morality, and adjustment (Byrne, 1961, 1969; Byrne & Griffitt, 1966; Byrne, Griffitt, & Stefaniak, 1967).

Research on both social projection (i.e., the false consensus effect) and self-anchoring offers additional examples of how self-perceptions may affect social perceptions of unfamiliar others. In both cases, participants hold the egocentric default belief that unfamiliar others are generally more self-similar than would be predicted by chance; favorable self-perceptions are automatically transferred to others in order to enhance self-esteem and feel a sense of belonging (Cadinu & Rothbart, 1996; Ross, Greene, & House, 1977). Interestingly, social projection and self-anchoring effects are most likely to occur for ingroups (which tend to be more associated with the self), implying that not only are we biased to think that others are similar to us, but that this bias is exaggerated when there is an established link between self and that group (Cadinu & Rothbart, 1996; Clement & Krueger, 2002; Krueger & Clement, 1996; Marks & Miller, 1987; Spears & Manstead, 1990; Ward, 1967).

This body of literature leads us to hypothesize that self-similarity acts as an automatic heuristic, so that new individuals are evaluated and categorized based on similarity to self-concept. The representation of the new individuals is biased by perceived similarity to self, leading to differences in strength of association between self and the other. We chose to measure this association implicitly because self-report and explicit judgments arise from conscious reflective processes that are subject to biases (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005). Further, to our knowledge, no research has used an implicit measure to look at the formation of representations of new individuals. We tested the self-heuristic hypothesis by developing a new version of the IAT, a test of cognitive association strength. The self-similarity IAT allows us to tap into the automatic cognitive associations of interest (Greenwald, McGhee, & Schwartz, 1998).

The IAT is a forced-choice task in which participants must categorize words as quickly and as accurately as possible into pre-specified categories (Greenwald et al., 1998). The categories are manipulated so that two different categories may require the same or different responses (e.g., left versus right key press); thus, any two categories can be paired together by requiring the same response. For example, in a simple categorization task, participants might sort words into categories representing magnitude of sound (e.g., loud/quiet) or size (e.g., large/small). The categories are paired by requiring a left response for words that belong to either the loud or large categories, and a right key response for words that belong to either the quiet or small categories. The relative strength of association between particular concepts is measured by comparing reaction times between paired conditions. For example, categorizing words into the loud category is faster for cognitively compatible category pairings (e.g., loud paired with large) than cognitively incompatible pairings (e.g., loud paired with small).

Our participants performed the self-similarity IAT, in which they categorized demographic words according to self (me/not me) and personality traits belonging to newly encountered individuals (self-similar/self-dissimilar). Self-similarity was manipulated in a previous learning phase, via sets of personality traits associated with each character which included a greater number of matches to self for the self-similar than the self-dissimilar character. The learning phase was meant to capture the implicit knowledge of self-similarity that we gain through brief real-life social encounters, in which traits are not formerly learned, but implicit knowledge of these traits is abstracted from experience. Although the self-similarity manipulation was not revealed to participants, implicit sensitivity to this similarity should affect representations of the new individuals relative to self-concept, consistent with Smith and Zárate (1992). Our first prediction is that the condition in which me and the self-similar character share a response key (and not-me and the self-dissimilar character share a response key) would be cognitively compatible with participants' internal representation, resulting in faster reaction times compared to the opposite (cognitively incompatible) pairing of categories.

Our second prediction is related to the demographic words specifically. It is possible that the self-similarity IAT effect is driven by associations among the trait words only, suggesting that the associations between self and the novel individual are specific to the trait information that was learned, consistent with Aron et al. (1991; Smith et al., 1999). The demographic words, however, are specific to the participant (e.g., first name, last name, date of birth, age, etc.) and are not relevant to the novel individuals. Since the demographic words are highly self-relevant and extremely well-practiced, they produce very fast responses, and one might expect that response times to demographic words would not be a significant contributor to the IAT effect. The model proposed by Smith et al. (1999) does not address this point. If demographic words do contribute to the IAT effect, this suggests that the associations formed due to similarity to a novel individual are not limited to overlapping traits.

Method

Participants were told that the main purpose of the experiment was to collect personality information to contribute to a database of student profiles. For each participant, personality traits, photographs, and demographic information were used to generate two profiles for the learning phase (a self-similar character and a self-dissimilar character) and to generate word lists of demographic and personality traits for the IAT. Participants learned about the two characters (learning phase) and were tested on this information (recognition phase). Participants then completed the IAT. The session ended with a questionnaire. Each of these phases is described in detail below.

Participants

A total of 33 undergraduate students (9 males, mean age = 19, SD = 1.32) were recruited from the introductory psychology research course at McMaster University and compensated with course credit. Participants had normal or corrected-to-normal eyesight.

Materials

Demographic Information

A demographic questionnaire was used to collect 16 words belonging to the me category for the IAT (e.g., name, age, phone number, academic program, career aspiration). A corresponding set of 16 words for the not-me category was generated from a list of sample answers that participants confirmed did not describe them.

Personality Traits

We drew our personality-trait words from Anderson's (1968) list of 555 trait words, for which he obtained likeableness ratings from 100 university students. This list is still widely used and his likeableness ratings have been replicated in recent years (e.g., Cartwright, 1997; Fisak, Tantleff-Dunn, & Peterson, 2007; Krienen, Tu, & Buckner, 2010; Mobbs et al., 2009). We reduced the original 555 words to 40 words that were relevant to university student self-concepts and were neutral enough to minimize responding according to social desirability. This was done by selecting the most neutral words (2.25–4.75 mean ratings for likeableness on a seven-point Likert scale), and eliminating additional words rated by at least two of five independent judges to be overly negative or positive, or obscure. Trait words for the three profiles (16 words each) were drawn from this set of 40 neutral trait words as described in the next section (see Figure 1 for list of traits).

Profiles

Participants completed a questionnaire, indicating on a seven-point Likert scale how well each of the 40 personality traits described themselves. The ratings ranged from 1 (this trait describes me very poorly) to 7 (this trait describes me very well). For each participant, we selected traits for the self-profile, self-similar profile, and self-dissimilar profile from this list of 40 traits, sorted by most to least self-descriptive (traits 1–40). To create the participant's self-profile, we used the 16 most self-descriptive traits (traits 1–16). To create the self-dissimilar character's profile, we used the 16 traits that had been rated as least self-descriptive by participants (traits 25–40). To create the self-similar character's

Participant 2

Which characteristics best describe you?			
	persuasive		perfectionistic
	obedient	x	forward
	sophisticated		excitable
x	thrifty		prideful
	sentimental		impulsive
x	objective	x	conservative
x	mathematical		unpredictable
	meditative	x	blunt
	fearless	x	emotional
	daring		bashful
x	sensitive		restless
	moralistic	x	choosy
	reserved		opportunist
	persistent		theatrical
	meticulous		impressionable
	bold	x	skeptical
	innocent	x	cunning
x	methodical	x	daydreamer
x	nonchalant	x	conventional
	self-contented	x	opinionated

FIGURE 1 Sample character profile. Each profile included a photo of the character (presented in colour in the experiment), their pseudonym ("Participant 1" or "Participant 2"), and a table listing all 40 personality traits, with X's indicating the 16 traits that best described that character.

profile, we avoided an exact match with the self-profile by selecting only the first eight traits that were chosen for the self-profile (traits 1–8); the remaining traits were the eight next most self-descriptive traits (traits 17–24), which were not included in the self-profile. There was no repetition in traits assigned between the self-similar and self-dissimilar profiles, and between the participant's own and self-dissimilar profiles. Prior to the learning phase, participants were shown the self-profile and confirmed that it was representative of the traits that they possessed.

Each self-similar and self-dissimilar profile was associated with a front-facing color photo of the character's face. For this purpose, four faces (two male and two female) with a neutral expression were selected from our laboratory database so that we could present faces of the same sex as the participant. One face was associated with the self-similar character's profile and the other with the self-dissimilar character's profile (counterbalanced across participants). During learning, the face was identified by a pseudonym (i.e., Participant 1) and presented along with a list of all 40 personality traits; an "x" marked each of the 16 traits that best described that character (see Figure 1, for example).

Learning Phase

Participants were told they would be learning about real people who had completed the same experiment. In the learning phase, participants studied each character's face and personality-trait profile for three minutes (order of profile presentation was counter-balanced). After the three-minute study period, a memory test was administered. A photo array of six faces (including the two learned faces and four new faces) and a table listing all 40 personality traits was presented; the task was to (1) select the two learned faces and (2) complete the personality profile for each one by placing an "x" beside each applicable personality trait. Each time the participant failed to reach 100% accuracy, a shortened, one-minute learning phase was repeated. All participants reached 100% accuracy within two to six attempts ($M = 3.42$), demonstrating successful association between the faces and the profiles.

Implicit Association Test

Stimuli

Participants viewed stimuli in a dimly lit room. A chin rest maintained a 90 cm distance from the 19-inch color CRT display (resolution of 1600×1200, frame refresh rate = 75 Hz). Presentation experimental software (Version 15.0, www.neuro-bs.com) was used to control stimulus presentation and record responses on a Pentium 4 computer using the Windows XP operating system. Stimuli were presented on a black background; text stimuli were presented in 20-point Helvetica font, with a vertical visual angle of .45° (horizontal visual angle varied with the length of the word). The face photos associated with the self-similar and self-dissimilar characters were used as target faces in the IAT; photographs subtended a visual angle 2.99° wide and 3.75° high. Category labels appeared on the top left versus right sides of the screen (counterbalanced), 2.86° horizontally and 1.43° vertically from the center of the screen. Words to be categorized were presented at screen center. In all blocks, the trait words and the participant 1 (i.e., self-similar) and participant 2 (i.e., self-dissimilar) categories appeared in a different font color (white) than the demographic words and the me and not-me categories (green), to clarify which category pair to use for categorization on a particular trial.

Procedure

Within each block, the relevant category labels and/or photographs remained on the screen for the duration of that block. For example, the me label might be on the left side of the screen, in which case the not-me label would be on the right. Likewise, the participant 1 label with photo might be on the left side of the screen, in which case the participant 2 label with photo would be on the right. On each trial, one word was presented in the center of the screen until a response was made. Participants indicated with a button press, as quickly and accurately as possible, whether the word belonged to the category presented on the left (requiring a left response: "z" key) or right (requiring a right response: "/" key) side of the screen. Demographic words were to be categorized according to the me or not-me categories. Trait words were to be categorized according to the participant 1 (i.e., self-similar) or participant 2 (i.e., self-dissimilar) categories. Accuracy and reaction time for categorization were recorded. Following incorrect responses, a red X appeared above the word, both of which remained on the display until the correct response was selected. The center word was removed from the display upon a correct response; the inter-trial interval was 500 ms.

Practice blocks consisted of 64 trials (32 demographic words presented twice or 32 trait words presented twice, in random order). There were five practice blocks. The first three provided practice with categorizing self-similar and self-dissimilar characters' traits; the characters' photographs and pseudonyms (participant 1 and participant 2) were displayed on the left versus right side of the screen (counterbalanced). The fourth practice block provided practice categorizing the demographic words; the category names me and not-me were displayed on the left versus right sides of the screen (counterbalanced). Participants were given more practice blocks for the trait than demographic information because knowledge of the trait information was new, whereas the demographic information represents a well-practiced and existing categorization. After the fourth practice block, the first experimental block was presented, followed by a fifth practice block (described below), and finally the second experimental block.

The two experimental IAT blocks consisted of 128 trials (32 demographic words and 32 trait words, each presented twice, in random order). Participants sorted both trait and demographic words across four categories, two represented on the left and two on the right side of the screen. There were two possible pairings of categories. One consisted of the self-similar and me categories on one side and the self-dissimilar and not-me categories on the other side (hypothesized compatible pairing). The other consisted of the self-similar and not-me categories on one side and the self-dissimilar and me categories on the other side (hypothesized incompatible pairing). The two category pairings were presented in two separate blocks, with block order and side of the screen of each pairing counterbalanced across participants. Thus for any particular participant, me was paired with the self-similar character in one IAT block, and with the self-dissimilar character in the other IAT block. In between the two experimental IAT blocks, the final practice block was presented to provide practice with the new mapping between the left and right responses for the categories that changed sides in the second IAT block. Table 1 presents a schematic representation of the sequence of blocks, and Figure 2 shows an example of a trial sequence in an experimental IAT block.

Follow-up Questionnaire

In the follow-up questionnaire, participants were asked to indicate which character they liked best, to replicate Byrne's (1969) finding that participants prefer self-similar to self-dissimilar individuals. This ensured that we had created characters that were self-similar and self-dissimilar to the participant. In addition, participants were asked to indicate which

TABLE 1 Schematic Illustrating Progression of Blocks in the IAT Phase

Block	Task	Number of trials	Left response key category	Right response key category
1, 2, 3	Practice categorizing traits	64	"Participant 1" photo (i.e., self-similar)	"Participant 2" photo (i.e., self-dissimilar)
4	Practice categorizing demographic information	64	"Me"	"Not-me"
5	IAT block 1 categorizing traits and demographic information	128	"Participant 1" and "Me"	"Participant 2" and "Not-me"
6	Practice (response remapping) categorizing demographic information	64	"Not-me"	"Me"
7	IAT block 2 categorizing traits and demographic information	128	"Participant 1" and "Not-me"	"Participant 2" and "Me"

of the two characters they would rather (1) meet in person, (2) work with on a task in the lab, and (3) reward with $10. Along with the question asking which character they liked best, these questions were used to generate a composite score, indicating the degree to which participants preferred the self-similar character over the self-dissimilar character in a number of real-life situations. At the end of the questionnaire, participants were also asked to guess the experimental manipulation.

Results

Of the 33 participants, three participants accurately guessed the manipulation and four participants recognized that one of the characters was similar to themselves but felt that it was unrelated to the experiment. These seven participants are included in the data presented here, as removing them from analyses did not alter the results. All reaction-time analyses were conducted on mean values for correct trials for each participant, with a 3 standard deviation outlier cut-off.

IAT Results

We examined the effect of self-similarity on IAT categorization accuracy and reaction time. The category pairing variable is defined as compatible pairing (me paired with the self-similar character, and not-me paired with the self-dissimilar character) versus incompatible pairing (me paired with the self-dissimilar character, and not-me paired with the self-similar character). A 2 x 4 repeated-measures ANOVA tested within-subject factors of IAT category pairing (compatible pairing, incompatible pairing) and word type (self-similar character traits, self-dissimilar character traits, me words, not-me words). Where appropriate, Greenhouse-Geisser correction was applied; we report original degrees of freedom, mean squared error (MSE), and corrected *p*-values. Where appropriate, the Bonferroni correction for multiple comparisons was applied.

The ANOVA for accuracy data revealed several significant effects. A main effect of IAT category pairing, F (1, 32)=7.72, MSE = .004, $p = .009$, showed greater accuracy for the compatible pairing compared to the incompatible pairing. Word type, F (3, 96) = 38.40, MSE = .010, $p < .001$, and the interaction of IAT category pairing and word type, F (3,96) = 3.64, MSE = .004, $p = .037$, were also significant. Paired sample *t*-tests

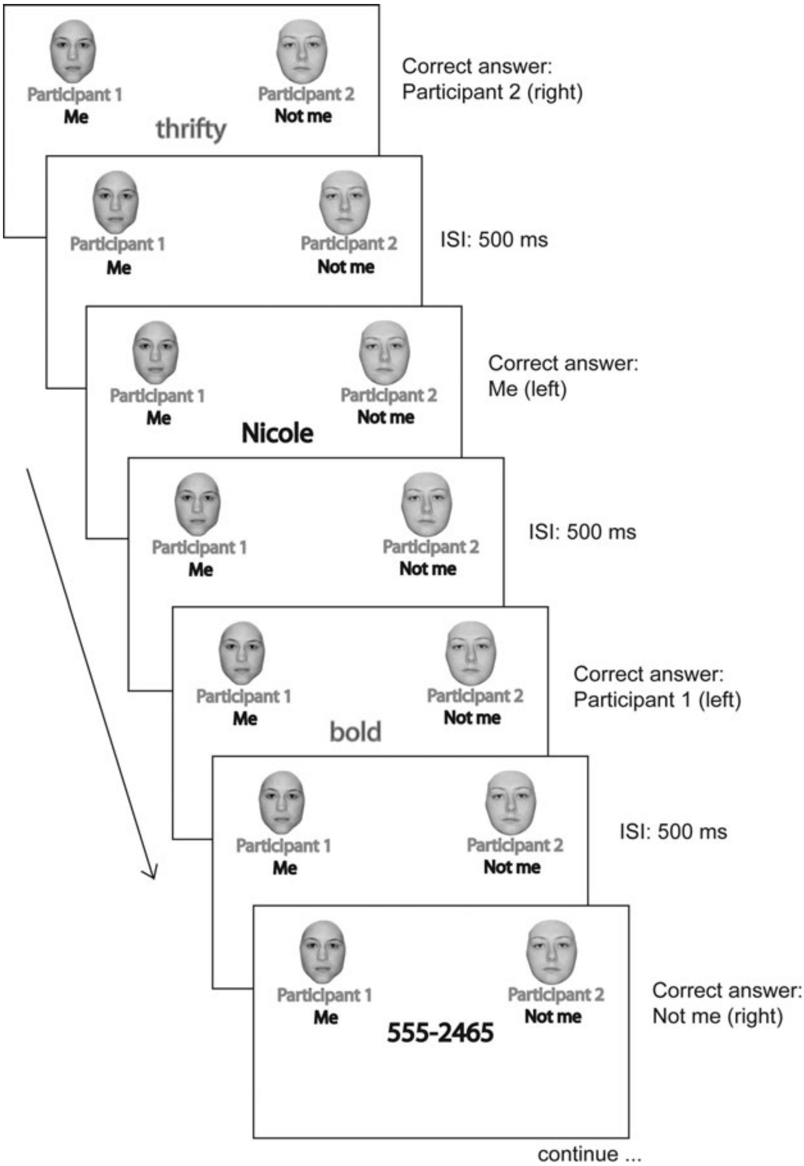

FIGURE 2 Example trial sequence for IAT block.

revealed that accuracy was significantly greater for self-similar character trait words in the compatible compared to the incompatible pairing, t (32) = 2.73, p = .01, while none of the other comparisons reached statistical significance. Table 2 presents a summary of the accuracy means and standard errors.

The ANOVA for reaction-time data revealed a significant main effect of IAT category pairing, F (1, 32) = 8.48, MSE = 14,931, p = .006, showing faster reaction time for the compatible pairing compared to the incompatible pairing. Word type, F (3, 96) = 102.05, MSE = 16,467, p < .001, was also significant, but there was no significant interaction of IAT category pairing and word type, F (3, 96) = .93, MSE = 4190, p = .411. To further examine word-type differences, paired sample t-tests showed that me words (M = 670.91

TABLE 2 Means for proportion accuracy data by IAT pairing and word type (numbers in brackets are within-subject standard errors)

Word type	Compatible pairing	Incompatible pairing	Mean difference
"Me"	.95 (.01)	.95 (.01)	0 (.01)
"Not-me"	.94 (.01)	.94 (.01)	0 (.01)
"Self-similar"	.86 (.02)	.80 (.02)	.06 (.02)
"Self-dissimilar"	.82 (.01)	.80 (.01)	.02 (.01)
Mean (all word types)	.89 (.004)	.87 (.004)	.02 (.01)

Note: Proportion Accuracy: Mean (SE).

ms) were categorized faster than not-me ($M = 733.92$ ms), $t\ (32) = -7.79$, $p < .001$, self-similar ($M = 991.19$ ms), $t\ (32) = -11.52$, $p < .001$, and self-dissimilar words ($M = 957$ ms), $t\ (32) = -12.46$, $p < .001$. Not-me words were also categorized faster than self-similar, $t\ (32) = -9.69$, $p < .001$, and self-dissimilar words, $t\ (32) = -9.49$, $p < .001$, which did not significantly differ from one another. This gradient in reaction time likely reflects differential familiarity with these categories. For example, participants are most familiar with the demographic me words (a robust and established category of self-knowledge) and least familiar with the traits possessed by the self-similar and self-dissimilar characters (newly encountered individuals). Table 3 presents a summary of the reaction-time means and standard errors.

We were also interested in how our effect related to explicit measures and found that the size of the IAT effect correlated with character preference, as measured on the follow-up questionnaire. As expected, participants were significantly more likely to prefer the self-similar character than the self-dissimilar character, $\chi^2\ (1) = 8.76$, $p = .003$, replicating Byrne's (1969) preference finding. The degree of this preference effect was positively correlated with the size of the IAT effect for reaction time, $r = .398$, $p = .022$.

Several researchers have suggested an alternative scoring method for the IAT (Greenwald, Nosek, & Banaji, 2003; Lane, Banaji, Nosek, & Greenwald, 2007). This method uses the difference score between category pairing conditions (me/dissimilar block minus me/similar block) divided by the inclusive standard deviation to calculate D. Using this method, we calculated $D = .34$, which was significantly greater than zero, ($M = .08$, SD $= .197$), $t\ (32) = 2.37$, $p = .024$, indicating a significant IAT effect consistent with the analyses presented above.

The IAT effect calculation collapses over categories to compare responses in compatible and incompatible blocks. To address our second hypothesis, we analyzed whether word categories contributed differentially to the IAT effect. There are two possibilities: (1) the associations could be directional, so that only the trait words, and not the demographic words, are affected by the compatibility pairing; or (2) the associations

TABLE 3 Means for Reaction Time Data by IAT Pairing and Word Type (Numbers in Brackets are Within-Subject Standard Errors)

Word type	Compatible pairing	Incompatible pairing	Mean difference
"Me"	658 (14)	684 (13)	− 26 (11)
"Not-me"	714 (15)	754 (11)	− 40 (12)
"Self-similar"	969 (20)	1014 (23)	− 45 (27)
"Self-dissimilar"	925 (18)	989 (19)	− 64 (25)
Mean (IAT effect)	816 (8)	860 (8)	− 44 (26)

Note: Reaction time (ms): Mean (SE).

could be bidirectional, so that both trait words and the more stable demographic words are affected by the compatibility pairing. The second possibility is most interesting, because it suggests that processing of the highly self-relevant demographic information does not dominate, and that the newly formed representations of the novel individuals can affect processing of self-relevant information. Thus, it is important to ask whether the IAT effect was driven solely by the self-similar/self-dissimilar trait words or also by the me/not me demographic words. We performed a repeated-measures ANOVA comparing the size of the IAT effect (incompatible RT minus compatible RT) across all four word types (me, not me, self-similar, self-dissimilar). We found no significant effect of word type on the size of the IAT effect, F (3,96) = .927, MSE = 8379, p = .431, supporting the hypothesis that the size of the IAT effect was similar across all four word types. We address this important observation further in the discussion.

Recoding in the IAT

We addressed the recoding interpretation of IAT results that claims that one of the response tasks may be dominant, and that response speed differences are due to response activation conflict, similar to a Stroop task (De Houwer, 2001, 2003; De Houwer, Geldof, & De Bruycker, 2005; De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009; Fazio & Olson, 2003; Meissner & Rothermund, 2013; Mierke & Klauer, 2001; Rothermund & Wentura, 2001). For example, if the demographic task is dominant, then the words belonging to the trait task might on some trials be categorized according to the me/not-me categories. This would produce faster responses for compatible pairings because trait words describing the self-similar character require the same response key as the demographic me words. In contrast, this strategy would activate an incorrect response for incompatible pairings, therefore producing greater response conflict, leading to slower categorization.

If participants were, in fact, recoding categories in the IAT, we would expect to find an effect of the degree of trait self-relevance on reaction time. Trait words rated as highly self-similar (or self-dissimilar) should be faster to categorize into the me (or not me) category than trait words rated as neutral with respect to self. To address this hypothesis, we examined whether categorization speed was affected by participants' original ratings of the trait words. In each case, we compared traits rated as strongly self-relevant (i.e., strongly self-similar or strongly self-dissimilar) to traits rated neutrally with respect to self. If trait words are categorized according to me/not-me, reaction time should be shorter for strongly self-similar (6–7 rating on scale) or self-dissimilar traits (1–2 rating on scale) compared to weakly self-similar (4–5 rating on scale) or self-dissimilar traits (3–4 rating on scale) for the compatible block. Paired sample t-tests revealed no significant differences in reaction time between strongly self-similar traits and weakly self-similar traits (mean difference = –10.747 ms), t (31) = − .251, p = .804, nor between strongly self-dissimilar traits and weakly self-dissimilar traits (mean difference = –15.147 ms), t (28) = − .396, p = .695.[1]

As a further test of the recoding hypothesis, we used the ReAL model to estimate variance explained by the recoding parameter (Re) of the model (Meissner & Rothermund, 2013). This is a multinomial processing model that uses correct and incorrect responses in the compatible and incompatible blocks, and takes into account task-repeat vs. task-switching trials, to estimate the contribution of three processes: recoding as described above (Re), evaluative associations based on categories, driven by the compatible and incompatible category pairings (A), and identification based on item labels, driven by the simple match between an item and its category label (L). The ReAL model is most appropriate for a modified IAT, which increases the number of errors; our design did not generate enough errors to test the model for individual participants, therefore we used the RT means. Another limitation to using the ReAL model, as described in Meissner and Rothermund (2013), is that it assumes that associations are directional so that no associations are activated from the trait words to the me/not-me categories. This assumption does not hold for our data; therefore, we modified the model to allow for bidirectional associations. The implementation of the ReAL model on our data provides a non-significant result for Re ($p = .98$), consistent with the other tests we performed above, suggesting that recoding is not a concern.[2]

Additional Analyses

Finally, we investigated two additional issues identified by reviewers. One was whether individual differences in ratings of the traits correlated with the IAT effect. Participants were slightly biased to rate traits as self-relevant ($M = 4.501$, SD = .069), $t\,(32) = 7.256$, $p < .001$, and some traits were rated more self-relevant than others, $F\,(39{,}1248) = 6.376$, MSE = 1.758, $p < .001$. After Bonferroni correction, 6 of the 40 traits significantly differed from the average trait rating; however, there were no significant correlations between the 40 individual trait self-relevance ratings and the size of the IAT. The final analysis ruled out the possibility that the observed IAT effects might be due to differences in memory for self-similar compared to self-dissimilar characters. We ran paired sample *t*-tests comparing performance on participants' first attempt at the memory task before they were given feedback to correct their errors. There was no significant effect for hits, $t\,(32) = .072$, $p = .943$, false alarms, $t\,(32) = -.937$, $p = .356$, and d' sensitivity, $t\,(29) = -.131$, $p = .897$, nor for the total number of attempts required to reach the 100% accuracy criterion, $t\,(32) = -.725$, $p = .474$.

Discussion

Previous research suggests that unique self-conceptualizations can bias our perceptions of others, indicating that self may be an important heuristic used in structuring our cognitive representations of the outside world (e.g., Byrne, 1961; Hoyle, 1993; Lewicki, 1983; Markus et al., 1985; Ross et al., 1977). We used a self-similarity IAT to determine whether new representations of others are structured relative to self-concept. Our findings suggest that representations of others relative to self-concept are formed soon after the first encounter and that this process may be considered a default heuristic used in acquiring information about other people. Moreover, the results show that the newly formed associations between self and the self-similar character affect categorization of self-relevant knowledge that shares no overlap with the new individual. Collectively, the results provide support for the idea that self-similarity is an important mediator in social perceptions.

Participants were introduced to two new individuals who were presented as other students participating in the same study but were actually characters constructed to be self-similar or self-dissimilar to the participant, based on personality traits varying with respect to self-relevance. Participants learned to recognize the newly encountered individuals by studying face photos and personality traits. Following the learning phase, the self-similarity IAT was used to estimate associations of self with the two characters. More specifically, under different category pairing conditions, the self-similarity IAT required categorization of demographic words into me and not-me categories, and categorization of personality-trait words into the newly learned self-similar and self-dissimilar character categories.

Response pairing with the categories me versus not me influenced categorization of the self-similar and self-dissimilar trait words describing the new individuals. Items were categorized faster when self was paired with the self-similar character compared to the self-dissimilar character. Faster reaction time for a category pairing in the IAT represents greater associative strength between those categories (Greenwald et al., 1998). In this case, the IAT effect was sensitive to the cognitive associations established during the learning phase (via trait overlap) between me and the self-similar character versus the self-dissimilar character. This finding is consistent with previous research using other implicit measures, specifically, the facilitation of judgments for trait matches between self and a significant other (or ingroup) (Aron et al., 1991; Coats et al., 2000; Smith & Henry, 1996; Smith et al., 1999). Smith et al. (1999) attributed this effect to the idea that shared traits can proliferate network activations to self-representation in two ways: directly from the trait to self-representation, and indirectly by way of the link between the representations of the significant other (or ingroup) and self.

Importantly, the influence of the category pairing was bidirectional. In addition to the effect on shared traits, we found that categorization of demographic me and not me words was also influenced by the category pairing with self-similar versus self-dissimilar characters. Indeed, we found that me, not-me, self-similar, and self-dissimilar words showed similar-sized reaction time differences between the compatible (me/self-similar character and not-me/self-dissimilar character) and incompatible (me/self-dissimilar character and not-me/self-similar character) pairings, indicating that all four word types contributed to the IAT effect. Note that the demographic words were specific to the participant (e.g., first name, last name, date of birth, hometown, age, etc.) and were unrelated to the newly learned individuals.

It is important to understand why the self-similarity IAT effect was evident not only for traits that are shared between self and the self-similar character, but also for demographic information that was solely self-relevant. We would like to argue that the IAT compatibility effect occurs due to the enhanced association strength between self-concept and the self-similar character representation, and that this representation goes beyond possessing particular traits. Although our findings align with the trait self-relevance judgment tasks, which show facilitation when self and other either both possess or both do not possess a trait (Aron et al., 1991; Coats et al., 2000; Smith & Henry, 1996; Smith et al., 1999), our results also suggest that association between self and newly encountered others can affect processing of other information in the network. Therefore, our implicit measure offers a novel contribution to the literature, indicating that cognitive representation of novel individuals relative to self-concept can lead us to go beyond the information given.

This idea is consistent with literature on self-referencing, whereby simply leading participants to associate self with a particular stimulus results in enhanced positivity toward related stimuli (Gawronski, Bodenhausen, & Becker, 2007; Perkins & Forehand, 2012; Prestwich, Perugini, Hurling, & Richetin, 2010). For instance, when participants completed an IAT block in which they categorized self-relevant stimuli and brand A using one response key and non-self-relevant information and brand B using a second response key, subsequent implicit and explicit measures revealed that they viewed brand A more favorably than brand B (Perkins & Forehand, 2012); the authors theorize that this effect is due to the mere cognitive association between self and the previously unrelated target object. This self-referencing effect is comparable to our observed effect, which is that the similarity between self and other on a few traits affects cognitive associations with other concepts in the network.

Our results suggest that person representations are incorporated into our cognitive networks according to self-similarity, consistent with Smith and Zarate's (1992) model of social perception. Smith and Zárate (1992) hold that representation of an individual as an exemplar in one's cognitive network is based on the perceived similarity of that individual to existing exemplars. Perceptions of similarity are based on the specific dimension that one is attending to at the time of encoding. Most of the participants in our experiment were unaware that the characters varied with respect to self-similarity and yet they were highly sensitive to this manipulation, representing self-similar and self-dissimilar individuals with varying association strength to a self-exemplar. This observation implicates self-relevance as a dimension upon which exemplars are encoded, providing support for Smith and Zárate (1992) model, as well as the finding that social perceptions are generally based on dimensions and traits that participants possess, find desirable, or for which they have expertise (Fong & Markus, 1982; Lewicki, 1983; Markus et al., 1985).

We have shown that implicit associations between self and other influence retrieval of information that is related to the other person, but also information about self that is unrelated to the other person. This finding exemplifies our tendency to use self-similarity as a heuristic to "go beyond the information given" (Andersen, Reznik, & Chen, 1997). Research on the false consensus effect and social projection suggests that participants possess the egocentric default belief that others are more self-similar than they are in reality (Hoyle, 1993; Ross et al., 1977). Interestingly, more recent research has demonstrated that the false consensus effect is more likely to be engaged for judgments of the ingroup than the outgroup (Cadinu & Rothbart, 1996; Clement & Krueger, 2002; Krueger & Clement, 1996; Marks & Miller, 1987; Spears & Manstead, 1990; Ward, 1967). The exaggeration of this bias may occur due to the strong prior link between self and the ingroup, which could drive the projection of unrelated self-relevant traits upon the ingroup. This process resembles the result presented here: that the link between self and the self-similar character facilitates IAT performance even for demographic words that are unrelated to the newly encountered individual. In this way, the false consensus effect may exemplify a downstream consequence of the implicit processes observed in our experiment.

Another link with existing literature relates to an explicit measure that correlated with the self-similarity IAT effect. Specifically, the IAT effect was positively correlated with preference for the self-similar character in a variety of real-life scenarios. When others are already familiar, faster categorization for shared traits in a self-relevance judgment task is mediated by an explicit measure of relationship closeness with the other, which represents the participant's own estimate of how much they overlap with the other (Smith et al., 1999). Even though the participants in our study do not have a relationship with the characters, it may be that character preference offers a similar index of the cognitive association strength between self and newly encountered individuals.

It was important to ensure that the IAT effect was not due to recoding, whereby in the compatible pairing block, the participants could use the me/not-me categories while ignoring the self-similar/self-dissimilar character categories (De Houwer, 2001, 2003; De Houwer et al., 2005, 2009; Fazio & Olson, 2003; Mierke & Klauer, 2001; Rothermund & Wentura, 2001), since traits were by definition related to self. If this were the case, we would expect reaction time to vary based on ratings of trait self-descriptiveness (Markus, 1977). We found no such effect on reaction time when comparing strongly self-similar and -dissimilar traits to weakly self-similar and -dissimilar traits. Support against the recoding hypothesis for the present data is also provided using a multinomial processing tree model (ReAL model; Meissner & Rothermund, 2013), which separates influences from evaluative associations and recoding processes. Therefore, the assumption remains that participants were attentive to all four categories in completing the IAT and that we have not obtained a result confounded by participants' use of the wrong category set for classifying the self-similar and self-dissimilar trait words.

Our study offers a novel use of the IAT to explore associations regarding self-similarity. In Greenwald and colleagues (2002) social knowledge structure, self is central and is linked to various person, group, and attribute concepts (which themselves are inter-linked) with varying associative strengths. The IAT has typically been used to measure associative strengths between self and a valenced attribute concept (i.e., positive versus negative words) to determine implicit self-esteem (Greenwald & Farnham, 2000; Karpinski, 2004), between self and a non-valenced attribute concept (e.g., masculinity versus femininity) to uncover implicit self-concept (Greenwald et al., 2002; Greenwald & Farnham, 2000; Nosek, Banaji, & Greenwald, 2002), and between self and a group concept (e.g., White versus non-White) to reveal implicit group identity (Devos & Banaji, 2005; Knowles & Peng, 2005). The present study is the first to use the self-similarity IAT, which served as a tool in measuring the associative strength between self and person concepts (i.e., self-similar versus self-dissimilar character) to infer the relationship between self-concept and the incorporation of new person representations into the social network.

Conclusions

In summary, we provide evidence to suggest that new individuals are quickly incorporated into cognitive representational networks according to self-similarity. Specifically, participants easily formed associations between their self-concept and a self-similar character despite learning relatively little information about the character a short time before performing the IAT, thereby substantiating the importance of self-conceptualization for social perceptions of newly encountered individuals. Importantly, this cognitive association affected the categorization of knowledge that shared no overlap with the new individual, indicating that the effect of self-similarity goes beyond the available information to alter other associations in the network. In addition, these findings suggest that even an established categorization of self-knowledge is sensitive to representations of newly learned individuals. In this way, we support the idea that when making judgments about self, we are likely to implicitly access representations of self-similar others. This study is one of the first to use an *implicit* technique to directly access the cognitive relationship (i.e., association strength) between self and representations of newly encountered individuals to demonstrate this type of bias, thus avoiding methodologies inherently biased by conscious awareness. Future research could use the self-similarity IAT to test hypotheses about relations between self and familiar others, or to explore sensitivity to very small and subtle manipulations of self-similarity (i.e., same birthday).

Notes

1. This conclusion is also supported by comparing two types of self-similar traits. Of the 16 traits that described the self-similar character, 8 traits were selected as self-similar stimuli on the basis that they were the highest rated traits by participants (more similar; traits 1–8 of the sorted 40 traits) and 8 traits were selected as self-similar stimuli that were not among participants' 16 highest rated traits (less similar; traits 17–24). Recoding would produce faster categorization of the 8 more similar vs. less similar traits. A paired sample *t*-test performed for the compatible block revealed no significant difference (mean difference = −31.128 ms), $t(32) = -1.149$, $p = .259$, lending support to the idea that the IAT effect was not driven by recoding.
2. The four L parameters (me, not-me, self-similar, self-dissimilar) all differed from zero (all $p < .0001$). The A parameters representing me and self-similar associations differed from .5

(all $p < .01$). The A parameters representing the not-me ($p = .16$) and self-dissimilar ($p = .93$) associations did not differ from .5. This suggests that the me and self-similar associations contributed more to the IAT effect than the not-me and self-dissimilar associations. Converging evidence from an analysis of switch costs in our experiment lends confidence to our interpretation of the ReAL model results. A paired sample *t*-test revealed no significant difference in the size of switch costs across the compatible ($M = 90.36$ ms) and incompatible ($M = 104.99$ ms) blocks, $t\ (32) = -.925$, $p = .362$.

References

Andersen, S. A., Reznik, I., & Chen, S. (1997). The self in relation to others: Cognitive and motivational underpinnings. *Annals of the New York Academy of Sciences*, *818*, 233–275. doi:10.1111/j.1749-6632.1997.tb48258.x

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, *9*, 272–279. doi:10.1037/h0025907

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*, 241–253. doi:10.1037/0022-3514.60.2.241

Byrne, D. (1961). Interpersonal attraction and attitude similarity. *The Journal of Abnormal and Social Psychology*, *62*, 713–715. doi:10.1037/h0044721

Byrne, D. (1969). Attitudes and attraction. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. New York, NY: Academic Press.

Byrne, D., & Griffitt, W. (1966). A developmental investigation of the law of attraction. *Journal of Personality and Social Psychology*, *4*, 699–702. doi:10.1037/h0023993

Byrne, D., Griffitt, W., & Stefaniak, D. (1967). Attraction and similarity of personality characteristics. *Journal of Personality and Social Psychology*, *5*, 82–90. doi:10.1037/h0021198

Cadinu, M. R., & Rothbart, M. (1996). Self-anchoring and differentiation processes in the minimal group setting. *Journal of Personality and Social Psychology*, *70*, 661–677. doi:10.1037/0022-3514.70.4.661

Cartwright, D. (1997). Lacunae and likeableness in the big five model of personality. *Personality and Individual Differences*, *22*, 855–865. doi:10.1016/S0191-8869(96)00276-0

Clement, R. W., & Krueger, J. (2002). Social categorization moderates social projection. *Journal of Experimental Social Psychology*, *38*, 219–231. doi:10.1006/jesp.2001.1503

Coats, S., Smith, E. R., Claypool, H. M., & Banner, M. J. (2000). Overlapping mental representations of self and ingroup: Reaction time evidence and its relationship with explicit measures of group identification. *Journal of Experimental Social Psychology*, *36*, 304–315. doi:10.1006/jesp.1999.1416

Cunningham, S. J., Turk, D. J., Macdonald, L. M., & Macrae, C. N. (2008). Yours or mine? Ownership and memory. *Consciousness and Cognition*, *17*, 312–318. doi:10.1016/j.concog.2007.04.003

De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology*, *37*, 443–451. doi:10.1006/jesp.2000.1464

De Houwer, J. (2003). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219–244). Mahwah, NJ: Erlbaum.

De Houwer, J., Geldof, T., & De Bruycker, E. (2005). The Implicit Association Test as a general measure of similarity. *Canadian Journal of Experimental Psychology*, *59*, 228–239. doi:10.1037/h0087478

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368. doi:10.1037/a0014211

Decety, J., & Sommerville, J. A. (2003). Shared representations between self and other: A social cognitive neuroscience view. *Trends in Cognitive Sciences*, *7*, 527–533. doi:10.1016/j.tics.2003.10.004

Devos, T., & Banaji, M. R. (2005). American = White? *Journal of Personality and Social Psychology*, *88*, 447–466. doi:10.1037/0022-3514.88.3.447

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327. doi:10.1146/annurev.psych.54.101601.145225

Fisak, Jr, B., Tantleff-Dunn, S., & Peterson, R. D. (2007). Personality information: Does it influence attractiveness ratings of various body sizes? *Body Image*, *4*, 213–217. doi:10.1016/j.bodyim.2007.03.004

Fong, G. T., & Markus, H. (1982). Self-schemas and judgments about others. *Social Cognition*, *1*, 191–204. doi:10.1521/soco.1982.1.3.191

Gawronski, B., Bodenhausen, G. V., & Becker, A. P. (2007). I like it, because I like myself: Associative self-anchoring and postdecisional change of implicit evaluations. *Journal of Experimental Social Psychology*, *43*, 221–232. doi:10.1016/j.jesp.2006.04.001

Gillihan, S. J., & Farah, M. J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychological Bulletin*, *131*, 76–97. doi:10.1037/0033-2909.131.1.76

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*, 3–25. doi:10.1037/0033-295X.109.1.3

Greenwald, A. G., & Farnham, S. D. (2000). Using the implicit association test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, *79*, 1022–1038. doi:10.1037/0022-3514.79.6.1022

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480. doi:10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216. doi:10.1037/0022-3514.85.2.197

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*, 1369–1385. doi:10.1177/0146167205275613

Hoyle, R. H. (1993). Interpersonal attraction in the absence of explicit attitudinal information. *Social Cognition*, *11*, 309–320. doi:10.1521/soco.1993.11.3.309

Johnson, Jr, R. (1988). The amplitude of the P300 component of the event-related potential: Review and synthesis. In P. K. Ackles, J. R. Jennings, & M. G. H. Coles (Eds.), Advances in Psychophysiology (3rd ed., pp. 69–137). Greenwich, CT: JAI Press.

Karpinski, A. (2004). Measuring self-esteem using the implicit association test: The role of the other. *Personality and Social Psychology Bulletin*, *30*, 22–34. doi:10.1177/0146167203258835

Knowles, E. D., & Peng, K. (2005). White selves: Conceptualizing and measuring a dominant-group identity. *Journal of Personality and Social Psychology*, *89*, 223–241. doi:10.1037/0022-3514.89.2.223

Krienen, F. M., Tu, P. C., & Buckner, R. (2010). Clan mentality: Evidence that the medial prefrontal cortex responds to close others. *The Journal of Neuroscience*, *30*, 13906–13915. doi:10.1523/JNEUROSCI.2180-10.2010

Krueger, J., & Clement, R. W. (1996). Inferring category characteristics from sample characteristics: Inductive reasoning and social projection. *Journal of Experimental Psychology: General*, *125*, 52–68. doi:10.1037/0096-3445.125.1.52

Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: What we know (so far) about the method. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 59–102). New York, NY: Guilford Press.

Lewicki, P. (1983). Self-image bias in person perception. *Journal of Personality and Social Psychology*, *45*, 384–393. doi:10.1037/0022-3514.45.2.384

Marks, G., & Miller, N. (1987). Ten years of research on the false-consensus effect: An empirical and theoretical review. *Psychological Bulletin*, *102*, 72–90. doi:10.1037/0033-2909.102.1.72

Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*, *35*, 63–78. doi:10.1037/0022-3514.35.2.63

Markus, H., Smith, J., & Moreland, R. L. (1985). Role of the self-concept in the perception of others. *Journal of Personality and Social Psychology*, *49*, 1494–1512. doi:10.1037/0022-3514.49.6.1494

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, *38*, 299–337. doi:10.1146/annurev.ps.38.020187.001503

Meissner, F., & Rothermund, K. (2013). Estimating the contributions of associations and recoding in the implicit association test: The ReAL model for the IAT. *Journal of Personality and Social Psychology*, *104*, 45–69. doi:10.1037/a0030734

Mierke, J., & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. *Experimental Psychology*, *48*, 107–122. doi:10.1026//0949-3946.48.2.107

Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., & …Dalgleish, T. (2009). A key role for similarity in vicarious reward. *Science*, *324*, 900–900. doi:10.1126/science.1170539

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*, 44–59. doi:10.1037/0022-3514.83.1.44

Otten, S. (2005). The in-group as part of the self: Reconsidering the link between social categorization, in-group favoritism and the self-concept. In M. D. Alicke, D. Dunning, & J. Krueger (Eds.), *The self in social perception* (pp. 273–292). Hove, UK: Psychology Press.

Perkins, A. W., & Forehand, M. R. (2012). Implicit self-referencing: The effect of nonvolitional self-association on brand and product attitude. *Journal of Consumer Research*, *39*, 142–156. doi:10.1086/662069

Prestwich, A., Perugini, M., Hurling, R., & Richetin, J. (2010). Using the self to change implicit attitudes. *European Journal of Social Psychology*, *40*, 61–71. doi:10.1002/ejsp.610

Ross, L., Greene, D., & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*, 279–301. doi:10.1016/0022-1031(77)90049-X

Rothermund, K., & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test (IAT). *Experimental Psychology*, *48*, 94–106. doi:10.1026//0949-3946.48.2.94

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*, 873–882. doi:10.1177/0146167299025007009

Smith, E. R., & Henry, S. (1996). An in-group becomes part of the self: Response time evidence. *Personality and Social Psychology Bulletin*, *22*, 635–642. doi:10.1177/0146167296226008

Smith, E. R., & Zárate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, *99*, 3–21. doi:10.1037/0033-295X.99.1.3

Spears, R., & Manstead, S. R. (1990). Consensus estimation in social context. *European Review of Social Psychology*, *1*, 81–109. doi:10.1080/14792779108401858

Tajfel, H., & Wilkes, A. L. (1964). Salience of attributes and commitment to extreme judgments in the perception of people. *British Journal of Social and Clinical Psychology*, *3*, 40–49. doi:10.1111/j.2044-8260.1964.tb00403.x

Ward, C. D. (1967). Own height, sex, and liking in the judgment of the heights of others. *Journal of Personality*, *35*, 381–401. doi:10.1111/j.1467-6494.1967.tb01435.x