Overlapping Mental Representations of Self, In-Group, and Partner: Further Response Time Evidence and a Connectionist Model

Eliot R. Smith, Susan Coats and Dustin Walling

*Pers Soc Psychol Bull* 1999; 25; 873
DOI: 10.1177/0146167299025007009

The online version of this article can be found at:
http://psp.sagepub.com/cgi/content/abstract/25/7/873

Published by:
SAGE Publications
http://www.sagepublications.com

On behalf of:
Society for Personality and Social Psychology, Inc.

Additional services and information for *Personality and Social Psychology Bulletin* can be found at:

Email Alerts: http://psp.sagepub.com/cgi/alerts

Subscriptions: http://psp.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations (this article cites 5 articles hosted on the SAGE Journals Online and HighWire Press platforms):
http://psp.sagepub.com/cgi/content/abstract/25/7/873#BIBL
Overlapping Mental Representations of Self, In-Group, and Partner: Further Response Time Evidence and a Connectionist Model

Eliot R. Smith
Susan Coats
Dustin Walling
Purdue University

People's reports on their own characteristics are facilitated by matches with their relationship partner's or in-group's perceived characteristics and inhibited by mismatches. These results suggest that mental representations of the self, partner, and in-groups overlap. A previously untested implication is that judgmental facilitation or inhibition should operate in the reverse direction as well. The authors find that reports on the characteristics of a relationship partner or in-group are facilitated or inhibited by matches or mismatches with the participant's own characteristics. The size of the effect for the partner is linearly related to perceived relationship closeness. These findings, in combination with the previous ones, suggest that representations of self, other people, and social groups are constructed on-line rather than being stable, static entities, and the authors advance a connectionist model of this construction process.

People are not isolated and independent individuals but are pervasively influenced by their social surroundings that affect all types of thoughts, feelings, and behavior. To what extent do people's mental representations of the self and others reflect that fundamental fact? That is, do we maintain separate and noninteracting mental representations of the self, other people, groups, and other social objects? Or do these representations interact and influence each other, just as real people and social groups do?

In recent years, research has made major strides toward answering this question. Aron, Aron, Tudor, and Nelson (1991) hypothesized that a close relationship partner becomes part of the mental representation of the self. Using married students as research participants, they found evidence for overlapping representations. The participants completed pencil-and-paper questionnaires about their own and their partner's traits. Then, they made yes/no self-descriptiveness judgments on a computer on the same traits. Traits on which the participants perceived that they matched their partners were associated with faster response times (RTs) than traits on which participants differed from their partners. The logic of this method is similar to the logic of Stroop interference. As is well known, reporting the color of the ink in which a word is printed is speeded if the meaning of the word matches the color (e.g., the word green printed in green). On the other hand, the report of the color is slowed if the word mismatches the color, such as red printed in green. Different attributes of the same object (such as the meaning and the ink color of a word) can create facilitation if they match or interference if they mismatch, when someone tries to report on one of the attributes. In the same way, if mental representations of two persons or groups overlap so that they are effectively a single representation, reports on attributes of one will be facilitated or inhibited by matches and mismatches with the second.

Smith and Henry (1996) used the same method as Aron et al. (1991) and found that moderately salient in-groups (liberal arts or engineering majors, or fraternity/sorority members or nonmembers) also became part of the psychological self. As in the Aron et al. (1991) study, self-descriptiveness judgments were faster for traits on

Authors' Note: The research reported here was supported by the National Institutes of Mental Health under grants R01 MH46840 and K02 MH01178. Address correspondence to Eliot Smith, Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907-1364; e-mail: esmith@psych.purdue.edu.

© 1999 by the Society for Personality and Social Psychology, Inc.
which participants matched their in-groups than for traits on which they mismatched. A recent study by Smith, Coats, and Walling (1997) replicated this basic pattern of RT facilitation for other types of groups: university affiliation, day or night person, and first or later born person. Recent research by Brewer and Pickett (in press) also replicated the RT effect for students’ university affiliation in a control condition. Compared to this condition, a manipulation that increased students’ motivation to identify with their university in-group increased the size of the RT effect, whereas a different manipulation that motivated them to individuate themselves decreased the size of the effect. These findings, which are predicted by Brewer’s (1991) optimal distinctiveness theory, provide further evidence for the validity of the RT effect as a measure of people’s identification with a group.

This evidence suggests that an in-group or a relationship partner becomes part of the self (in the sense that the group or partner’s characteristics influence self-reports). This conclusion implies another hypothesis that has not yet been empirically tested. Does the self also become part of the mental representation of the partner or in-group? That is, do self-characteristics affect people’s judgments about their partner or their in-group? By analogy with the term implicit memory, which refers to effects of past experiences on tasks that do not apparently require memory at all (Smith, 1998), these could be called implicit effects on judgment, as self-characteristics are predicted to affect judgments about others that do not on their face implicate the self at all. This study is designed to test these hypotheses, which flow directly from the model of overlapping self and other representations but have not yet been empirically tested.

Related research by Srull and Gaelick (1983) has investigated the impact of the self on judgments, arguing that the self constitutes a “habitual reference point.” The research showed that others are judged to be more similar to the self than the self is to others and, based on Tversky’s (1977) theory of feature similarity, concluded that the self is indeed used as a habitual reference point in making comparisons and similarity judgments. Although this finding is conceptually related to our current hypothesis, our study and this previous research actually use quite different approaches and have different goals. In particular, Srull and Gaelick asked participants to make explicit similarity judgments, and their theoretical claim is simply that when people explicitly compare the self to another person, the self (rather than the other) tends to be used as the reference point in the comparison. By contrast, the present work focuses on implicit effects of the self on judgments that do not overtly make reference to the self at all, such as judgments about whether another person or social group possesses various characteristics. If they are detected, the implications of such effects go considerably beyond explicit judgments about the similarity of the self to others.

Two additional questions can be answered in this study, along with our basic hypotheses that the self will implicitly affect judgments about an in-group or relationship partner. First, for the in-group, do perceptions of the relevant out-group matter? Smith and Henry (1996) showed that all self responses that matched the in-group were facilitated. Perceptions of the out-group seemed not to matter; that is, responses on traits for which the self mismatched the out-group, or matched the in-group on traits that are particularly distinctive (differ from the out-group), were not associated with any additional effects on RT. Perhaps, though, when people are judging the in-group directly (rather than the self as in Smith & Henry, 1996), distinctiveness from the out-group may matter (Brewer, 1991). So responses for the in-group on traits that are perceived to differ from the out-group, as well as traits on which the in-group matches the self, may be facilitated.

For the partner, does relationship closeness matter? A clear prediction is that closer relationships should involve more overlap of self and partner representations, a situation that is graphically illustrated in Aron, Aron, and Smollan’s (1992) Inclusion of Other in the Self (IOS) scale. The IOS scale is a single-item pictorial measure intended to tap perceived closeness in a relationship. Respondents are presented with seven Venn-like diagrams representing different degrees of overlap of two circles, and are asked to select the picture that best describes their relationship. If this measure truly taps the extent of overlap of mental representations, respondents whose relationships are closer (as measured by the IOS) should show greater RT effects of matching or mismatching their self traits when describing their partner’s traits in the computer task. An interesting question is whether the size of this effect smoothly and linearly increases with relationship closeness, or whether only those relationships over a specific threshold produce effects on self and other representations. In other words, does relationship closeness act as a quantitative or dichotomous variable?

In summary, this study was designed to answer four questions: (a) When people describe an in-group, will characteristics that match versus mismatch the self be associated (respectively) with RT facilitation and inhibition? (b) Will in-group characteristics that mismatch the relevant out-group also be associated with facilitation? (c) When people describe the characteristics of a close relationship partner, will characteristics that match versus mismatch the self be associated (respectively)
with RT facilitation and inhibition? (d) Will the degree of relationship closeness (measured by the IOS scale) moderate this effect?

METHOD

Participants and Materials

Participants were 87 Introductory Psychology students who received course credit for taking part in the study. Students currently involved in a romantic relationship of at least 3 months duration were eligible to participate. We used the list of 90 heterogeneous traits that was previously employed by Aron et al. (1991) and Smith and Henry (1996). Example traits are creative, sensible, shy, and ambitious. We also used the IOS scale developed by Aron et al. (1992).

Procedure

Participants reported to the laboratory in groups of 2 to 5. They were informed that they were going to make judgments on pencil-and-paper questionnaires and then make similar judgments using the computer. Participants were then each given a questionnaire packet. Five of the pages in the packet directed participants to describe themselves, Greeks (i.e., fraternity/sorority members), non-Greeks, their romantic partner, and a stranger (Bill Cosby, included as a filler to obscure the focus of the research). Each of these five pages contained the list of 90 traits, and participants rated the descriptiveness of each trait for the specified target using a 7-point scale ranging from not at all to extremely. Instructions on the pages for group descriptions stressed that whereas not all members of any group are exactly alike, group members tend to be similar on many traits and that participants should provide their personal opinions about the general characteristics of these groups. The sixth page in the questionnaire packet asked whether participants belonged to a Greek organization and presented participants with the IOS scale. Participants were instructed to choose the picture that best described their relationship with their romantic partner. The six questionnaire pages were randomly ordered for each participant.

After completing the questionnaire packet, participants were seated at a computer for a second task, which consisted of two blocks of trials. The order of the two blocks was random. In one block participants were instructed to describe their romantic partner and in the other they were instructed to describe their group. “Their group” referred to membership in the Greek system: Participants who were members of the Greek system were asked to describe Greeks, whereas nonmembers were asked to describe non-Greeks. In each block, the same 90 traits presented in the questionnaire were displayed on the computer screen, in a different random order for each participant. Participants were asked to respond to the traits by pressing a “yes” key for those traits they considered to be descriptive of the target (either their romantic partner or group) or a “no” key for those traits they did not consider to be descriptive of the target. Participants were asked to respond to the traits as quickly and accurately as they could. Following the computer task, participants were thanked, debriefed, and dismissed.

RESULTS

In-Group and Self

Analysis. The data were analyzed using multiple regression. This approach was used because (a) the predictors are a mix of categorical and continuous variables, (b) several of the independent variables are defined by participants’ own judgments and hence may be intercorrelated, and (c) regression allows for the testing and estimation of each effect controlling for other effects that might influence RTs. The RT for each individual trait judgment regarding the in-group served as the dependent variable. First, because of the within-subjects nature of the design, a vector was dummy coded for participants to control for overall mean differences in participants’ response times (see Cohen & Cohen, 1983). There were four other independent variables. Self is the dichotomized version of the self-descriptiveness rating for the specific trait from the questionnaire (as in Smith & Henry, 1996, responses ranging from 1 to 3 were scored as no and 5 to 7 as yes; midrange responses of 4 were considered as missing). Out-group is the questionnaire rating of whether the trait describes the out-group, dichotomized in the same way. In-group is the yes or no response recorded by the computer on the in-group descriptiveness task. Evaluation is a continuous score summarizing eight judges’ evaluation of each trait, ranging from –8 to +8.

In this analysis, our focal prediction is for an interaction of Self × In-Group, with faster RTs associated with in-group responses that match self responses. This effect should be observed even controlling for other effects, such as a main effect of In-Group (yes responses are typically faster than no) or In-Group × Evaluation (subjects may be faster to say yes to positive traits and no to negative traits regarding their in-group). Another possible effect is an interaction of In-Group × Out-Group, or faster responses for traits on which the in-group mismatches (is distinctive from) the out-group.

Because the predictors are intercorrelated, main effects and interactions are not all orthogonal (independent of one another, as in a balanced ANOVA design). Therefore, a hierarchical regression procedure
was used. First, the full model (with the vector defining individual participants and the full set of interactions of the other four factors) was estimated. If the four-way interaction was nonsignificant, it was dropped and the model reestimated. This procedure continued, dropping high-order interactions one at a time until only significant effects remained; this defines the final model.

We had hoped to obtain adequate numbers of both Greeks and non-Greeks as participants so that those two groups would each serve as in-group and out-group for different participants. The analysis would have included Greek/non-Greek as an additional factor interacting with the others. Unfortunately, only 8 Greek participants took part in the study. Therefore, our main analyses omit these individuals and focus on non-Greek participants only. Thus, the in-group is always non-Greeks, and the out-group is always Greeks.

Results. The final model included two significant two-way interactions as well as the four main effects; no higher order interactions reached significance ($p > .05$). The predicted Self $\times$ In-Group interaction, $F(1, 3492) = 7.13, p < .01$, is shown in Figure 1. Participants were faster to make in-group descriptiveness judgments that matched their self-perceptions. In addition, an interaction of Self $\times$ Evaluation emerged, $F(1, 3492) = 9.20, p < .01$, indicating that participants gave either yes or no responses more quickly about their group’s standing on positive traits that they saw themselves as possessing, and on negative traits that they did not possess, compared to other traits.

To test whether the limitation to non-Greek participants might have made a notable difference in the results, the remaining participants were added to the analysis and effects of the Greek/non-Greek variable’s interactions with all the terms in the equation were tested. None of these terms was individually significant, nor was the entire group, $F(6, 3829) < 1$. Though the power of this test is low because of the limited number of Greek participants, these results suggest that it is the properties of in-groups that drive the results rather than the properties of a specific group. Similarly, Smith and Henry (1996) reported that their results were statistically consistent across four participant groups (Greeks, non-Greeks, liberal arts majors, and engineering majors).

Relationship Partner and Self

Analysis. As before, the RT for each individual trait judgment regarding the partner served as the dependent variable, and a predictor vector was dummy coded for participants to account for overall mean differences in participants’ response times. There were three other independent variables. Self is the dichotomized version of the self-descriptiveness judgment for the specific trait.

![Figure 1: Response time (RT) for judgments about in-group by in-group responses (abscissa) and self-responses (legend).](image-url)

Partner is the yes or no response recorded by the computer on the partner descriptiveness task. Evaluation is a continuous score summarizing eight judges’ evaluation of each trait, ranging from −8 to +8. As before, the prediction is for an interaction of Self $\times$ Partner, with faster RTs associated with partner responses that match the self responses.³

Results. The final model arrived at by the hierarchical regression approach included two significant two-way interactions as well as the three main effects. The predicted Self $\times$ Partner interaction, $F(1, 5439) = 12.25, p < .001$, is shown in Figure 2. Participants were faster to make partner descriptiveness judgments that matched their self-perceptions. In addition, an interaction of Partner $\times$ Evaluation emerged, $F(1, 5439) = 26.10, p < .0001$, indicating—not surprisingly—that responses attributing positive traits to the partner and denying negative ones were relatively fast. (Smith & Henry, 1996, found a comparable effect of Self $\times$ Evaluation on the speed of participants’ reports of their own characteristics.)

We also wished to test whether participants who viewed their relationship with their partner as closer would show the basic effect more strongly. We therefore conducted an analysis with the IOS scale, Aron et al.’s (1992) single-item self-rating of relationship closeness, as an additional predictor. We predicted an interaction of IOS by Self-Rating $\times$ Partner, with the basic Self $\times$ Partner interaction being stronger for participants who view their relationships as closer.

When the three-way interaction of Self-Rating $\times$ Partner $\times$ IOS was tested (in an analysis that also included the appropriate hierarchically lower order interactions and
main effects), it proved significant, \( F(1, 5378) = 6.47, p < .02 \). In this analysis, IOS was considered as a continuous variable, so this test shows that its effect has a significant linear component. To assess whether the effect was actually linear or involved some sort of threshold, the RTs for Self × Partner matches and mismatches were calculated separately for subjects who used each of the seven possible responses on the IOS scale. The result is shown in Figure 3. The size of the interaction effect (difference between matches and mismatches) increases smoothly as the IOS increases. Deviations from the linear trend were nonsignificant, \( F(5, 5378) < 1 \). Thus, it seems that the IOS scale taps quantitative degrees of relationship closeness that have graded effects on the overlap of self and partner mental representations. Especially in view of the nonsignificance of departures from a linear effect, there is little evidence in these data for the idea that closeness has effects only above some threshold.

**DISCUSSION**

**Summary of Results**

The results of this study support three of our four predictions. Based on the idea that people’s mental representations of the self overlap with representations of the partner and the in-group, we predicted that reports about both the partner’s and in-group’s characteristics would be facilitated or inhibited by matches or mismatches with the self. These hypotheses were strongly confirmed. This study thus complements Aron et al. (1991) and Smith and Henry (1996), who also found evidence for overlap of self with other representations in the form of effects in the reverse direction (effects of partner or group on self-reports).

Results showed that the self/in-group match had the primary effect. In-group/out-group match—that is, the distinctiveness of the in-group on a particular trait—had no detectable effect. This result is similar to the findings of Smith and Henry (1996), who also found that it was simply the match between self and in-group, regardless of the out-group’s standing on the trait, that affected self-report RTs.

Finally, the effect of the self on reports about the partner was moderated by relationship closeness (measured by the IOS) as predicted. The IOS measure of closeness thus has implications not only for various relationship processes and outcomes (see Aron et al., 1992) but also for mental representations and judgment processes, which affect the strength of the RT facilitation effect. The effect seems to be linear in nature rather than showing a sharp discontinuity between close and nonclose relationships. However, recall that all participants had been in a romantic relationship of at least 3 months duration, so it is possible that all of their relationships—even those that were described with the lowest values on the IOS scale—are effectively close and differ sharply from any nonclose or acquaintance relationship. What we can say is that within these relationships, the degree of closeness has a relatively linear effect on the magnitude of the RT facilitation.

Could these effects be explained by uncontrolled attributes of the individual traits rather than by self-other match versus mismatch, as we propose? Suppose we had found that traits that a given participant saw as describing the self were associated with quicker yes answers to questions about whether the trait described the in-group...
or the partner. Such a pattern of data could be explained by the assumption that these particular traits (which would be different for each participant) were more accessible, more self-relevant, or had other special characteristics that caused the relatively fast responses. No linkage between self- and other-representations would be required to account for such a finding. However, this is not what we found. Instead of the main effect of self-descriptiveness on response times (with traits describing the self being faster than other traits) that is implied by this explanation, we found that traits on which the response for the self matched the partner or in-group led to faster RTs. Traits on which the responses for self and partner (or in-group) were both negative were speeded as much as traits for which both responses were positive (see Figures 1 and 2). This pattern of data cannot be explained by the assumption that traits that describe the self are particularly high in accessibility, frequency of use, and so forth, because traits that do not describe the self also produce facilitation (in the no/no cell).

Given the data pattern we found, to speculatively account for our results without invoking the effect of a self-partner or self-group match on RT, one would have to assume that there is something special (for each participant) about those traits that describe both self and other and also those traits that describe neither self nor other. Yet, this special characteristic would have to be uncorrelated with (a) self-descriptiveness (which had no main effects) and (b) trait evaluation (which was controlled in the analyses). It is difficult to imagine what such a characteristic might be. Moreover, the clear dependence of the match-mismatch RT difference on relationship closeness (see Figure 3) also helps rule out such a potential confound. To account for these results, the hypothetical special characteristic of the traits would also have to (c) be present to a systematically greater extent in participants who viewed their relationships as closer on the IOS. Still, despite the implausibility of such a potential factor, in the absence of any practical or ethical way to manipulate people’s self-views, perceptions of their partners, or relationship closeness, the findings we report here—like those in much research on the self and relationships—remain correlational rather than experimental in nature.

A Connectionist Model of Self-Partner Overlap

Aron et al. (1992) did not offer a detailed model of mental representations and processes that might account for their initial findings in this area, contenting themselves with metaphorical terms like the idea of the partner becoming part of the self. How can these results be explained at the level of underlying mechanisms? First, consider how this might be done in a traditional social cognition theory such as Wyer and Srull (1989). To answer a trait question about the partner, presumably the perceiver would retrieve a mental representation of the partner from its storage bin in memory. That representation, consisting of a central node representing the person linked to nodes representing traits or other characteristics, would then be used to answer questions. In such an account, it is difficult to explain why or how the characteristics of another person (i.e., the self) should affect the process. Traditional theories portray representations of distinct persons as discrete and independent, and if the partner representation can be retrieved and its structure (the links between person and trait nodes) can be read off directly to answer questions, the patterns of RT interference and facilitation should not occur.

However, this traditional picture of mental representations and judgment processes is not the only possible one, despite its familiarity and seeming naturalness. And this picture is deeply inconsistent with recent research that suggests mental representations of the self and other people—in fact, all social and nonsocial objects—are flexible and context-sensitive constructions rather than fixed, stable things that are stored away and retrieved like crates in a warehouse (Barsalou, 1987; Smith, 1998). For example, Markus and Wurf (1987) argued that the self is flexible and dynamic, changing with social situations and the person’s current goals. Similarly, Turner, Oakes, Haslam, and McGarty (1994) wrote that “the concept of the self as a separate mental structure does not seem necessary, because we can assume that any and all cognitive resources—long-term knowledge, implicit theories, . . . and so forth—are recruited, used, and deployed when necessary” (p. 459) to construct a self-representation appropriate for the current situation. Of course, these points apply to representations of other people and social groups as well as the self.

New models of the fundamental nature of mental representation, now being developed within cognitive science and social cognition, allow us to understand how representations can be flexibly reconstructed rather than stored and retrieved as static entities. In particular, connectionist representations have this property (see Smith, 1996). In such models, semantically meaningful representations (e.g., those used to answer questions) are transient patterns of activation rather than stable, fixed structural connections between nodes. Connectionist networks do consist of nodes or units connected by links over which activation can be sent, but individual nodes and links cannot be directly accessed and used to answer questions or retrieve specific information. The
underlying network structure can affect judgments and overt responses only insofar as it affects the spread of activation and therefore the nature of the representation (pattern of activation) that is constructed.

To demonstrate that a simple connectionist model can account for the results reported here, we conducted a small-scale simulation. This simulation will be described as if the other person is a relationship partner, but all assumptions hold without change if the "other" is assumed to be an in-group instead.

Connectionist models can be classified as localist or distributed, based on the nature of the representations they use. In localist models, a node is interpreted as representing a meaningful object or concept, whereas in distributed models meaning is attached to a pattern of activation across a number of nodes and no individual unit has a particular, fixed meaning. Distributed models have important advantages, including their greater plausibility as descriptions of biological neural networks, the ability to learn, and to automatically generalize from one concept to other similar concepts (Smith, 1998). However, in this article we describe a localist representation because of the one great advantage, simplicity and ease of interpretability. The assumptions of this model are the following.

1. Basic network assumptions. The basic architecture of the model is the Interactive Activation and Competition (IAC) model proposed by McClelland and Rumelhart (1981) as a model of memory structure and retrieval. The IAC model uses a localist representation in which concepts (such as a person or a trait) are represented by distinct units. The activation of each unit varies over time according to the equations of the model, but is constrained to be within the range of −0.2 (minimum activation level) to +1.0 (maximum activation). Units are interconnected by bidirectional links whose weights are fixed at either +1 or −1. We used the IAC implementation provided by the Brainwave simulator (version 1.1; http://www.psy.uq.edu.au/~brainwav/). We left all model parameters at their default values rather than conducting any search for parameter values that would make our simulations work better.

2. Specific pattern of connections. We set up units and links as shown in Figure 4. The units representing self and partner are connected with a link having a positive weight, representing the linkage between mental representations that we postulate. Each of these units is also connected to several units representing traits, with +1 links to traits that are believed to characterize the self or other and −1 links to traits that are thought not to describe the person. See Kunda and Thagard (1996) for similar examples of translating psychological situations into network connections.

3. Learning. In this article, we simulate only the processing that occurs in this connectionist network after it is set up, not the process by which the network is constructed. However, we assume that a network with these functional connections would result from an associative learning process as the individual repeatedly considers or encounters information suggesting that the self or the other does (or does not) possess specific traits and information about the self-other relationship. For example, we suggest that repeatedly encountering information suggesting that "I am not polite," "my partner is polite," and "I love my partner" would result in the construction of positive self-partner and partner-polite links and a negative self-polite link.

4. Question answering. Given the connectionist network depicted in Figure 4, we model the process of answering a question about whether a trait describes the partner as follows. The relevant person node (in this case, partner) is fixed at its maximal activation level of +1.0; then, activation is allowed to flow through the network for a number of cycles until activation values have reached asymptote and are no longer changing. A yes or no response is then given, based on the positive or negative activation level attained by the trait. To generate predictions about response times, in common with virtually every psychological model of response generation ever proposed, we assume that the closer the activation level is to the high and low ends of the scale (i.e., the +1.0 and

---

**Figure 4** Interactive activation and competition (IAC) model of self, other, and traits: solid lines represent positive links between nodes, dotted lines represent negative links.
Basic results. The simulation results show that when the partner unit is turned on as just described, final activation levels for traits 1 and 2 (both connected by positive links to the partner unit) are .90 and .46, respectively, corresponding to yes answers. Trait 1, on which the link to the self is also positive (matching the link to the partner), has an activation level much closer to the +1.0 maximum level and is therefore expected to give rise to a shorter RT than trait 2. Similarly, activation levels for traits 3 and 4 (both connected by negative links to the partner unit) are −.15 and −.18, respectively, corresponding to the no answers. Trait 4, on which the link to the self is also negative (matching the link to the partner), has an activation level much closer to the −0.2 minimum and is therefore expected to give rise to a shorter RT than trait 3. Thus, the model predicts the same pattern of results as we obtained in this study. The Figure 4 network is fully symmetrical; therefore, if the labels of the self and partner units are swapped, exactly the same simulation shows that the model also predicts the results of Smith and Henry (1996) and Aron et al. (1991).

Effect of degree of closeness. In additional simulations, the strength of the self-partner link, which was fixed at +1 to obtain the results just described, was varied to determine whether the model could reproduce the empirically obtained graded effects of relationship closeness. We assume that this link, representing the strength of the connection between self and other, differs between close and less close relationships but that none of the person-trait links differ. The results are shown in Table 1. As the magnitude of this link is decreased, the differences between activation levels reached by traits 1 and 2, and also traits 3 and 4, decline and eventually vanish when the self-other link reaches zero. Thus, varying a single parameter in the connectionist model, a parameter related in a meaningful and principled way to the concept of relationship closeness, reproduces the results found in research participants with varying degrees of perceived closeness.

There is one apparent difference between the model’s predictions shown in Table 1 and the actual empirical data from our study, which deserves mention. Table 1 shows that in the model, the effect of increasing relationship closeness is mainly a slowing down of responses on mismatching traits rather than a speedup on matching traits. In the data, the means for the matching and mismatching conditions that are shown in Figure 3 show a different pattern: Increasing IOS is associated with faster matching responses, whereas RTs for mismatching responses stay relatively constant across levels of IOS. This apparent discrepancy is easy to reconcile, however, if one realizes that the connectionist simulation is intended to model only the way the linkage of self and other nodes affects the judgment process and not other aspects of the situation that contribute to overall RT means. Another relevant aspect is that closer relationships are bound to be associated with higher levels of knowledge about the relationship partner, due to increasing levels of shared activities and self-disclosure in those close relationships. Greater knowledge about the partner, in turn, will create greater confidence and faster RTs when the research participants answer questions about their partners’ traits. This process will therefore generate a main effect of IOS on RTs: Members of closer relationships will respond faster overall. When this effect (which our simulation does not attempt to capture) is added to the model predictions (increasing RT for mismatches, flat RT slope for matches as IOS increases), the sum of the two corresponds exactly to the empirical results (flat RT for mismatches, decreasing RTs for matches as IOS increases). Thus, the difference between the model results and the empirical findings reflect not an error in the model but simply its omission of an unrelated but fairly obvious process, the fact that relationship closeness leads to deeper knowledge of the partner and hence to more confident and faster responses about the partner’s characteristics.

Implications

In this connectionist model, representations of the self and others are not isolated and independent structures that are stored separately in a metaphorical file cabinet or storage bin and retrieved as discrete entities. This is the way that representations of persons or other social objects are portrayed in most traditional theories of mental representation (Smith, 1998). Instead, self and other nodes are linked, both by the direct connection signifying the relationship between self and other and by indirect connections through commonly shared traits (or goals, activities, emotions, or other characteristics).

This model has a second, equally important distinction from existing theories of mental representation of the self and others. The answer to a question (such as “Are you polite!”) is assumed not to be read off from the structure of the connectionist network. That is, the perceiver cannot selectively access the specific connection between self and a trait node and give an answer based on that connection weight. Such structural details of the network are simply not retrievable, and in this model they do not function as representations in themselves. Instead, the network’s structure constitutes a substrate in which flows of activation can generate representations—patterns of activation—that in turn can be accessed and used in judgments. Thus, questions are
answered by allowing activation to flow through the network; in our example, by activating the relevant person node and waiting for activation to accumulate on the trait. The flow of activation is certainly influenced by the direct person-trait connection, but also by all other connections in the network. As Smith (1996) and Kunda and Thagard (1996) outlined, connectionist models generally operate not by retrieving a single, discrete memory representation, but by using multiple representations simultaneously in a process of multiple constraint satisfaction. In this case, the constraint satisfaction process allows representations of close others (a partner or in-group) to influence responses to questions about the self, or vice versa.

Thus, we do not believe that the results obtained by us, by Smith and Henry (1996), and by Aron et al. (1991) can be explained by the retrieval and use of discrete, context-invariant representations of the self, other persons, and social groups. Equally, we doubt that any kind of explicit comparison procedure that judges the match or mismatch between self and others is responsible for the systematic patterns of RT facilitation and inhibition. Instead, we attribute these results to the online construction and reconstruction of representations (i.e., patterns of activation) as the judgments are made, in a way that is affected by the entire structure of the network and therefore by related representations. The quotation from Turner and his colleagues (1994) cited above captures the gist of this process: Knowledge about the self is among the resources from which judgments about a partner or in-group can be constructed, as well as the converse. Thus, judgments about the self are affected by knowledge about the partner and the in-group (Aron et al., 1991; Smith & Henry, 1996), and judgments about the partner and in-group by knowledge about the self.

We are arguing that the traditional view of mental representations as discrete, independent constructs makes it difficult to account for these results that suggest functional overlap between self- and other-representations (i.e., suggest that both representations are accessed in the process of answering questions about one). However, it might be thought that the view of self- and other-representations as linked (as in our simulation) is incompatible with the observation that personal and in-group identities can be functionally antagonistic. People sometimes view themselves in terms of distinct personal identities, and at other times as relatively interchangeable group members (see Turner et al., 1987). But this functional antagonism is simply another way of looking at the context-dependent nature of the online process of constructing and reconstructing mental representations. In one situation, individuating characteristics may be most prominent in the activation pattern that is the situation-specific self-representation, whereas in another, group or partner characteristics may be most important, due to motivational factors or situationally driven accessibility. As we have suggested before, connectionist models of representation offer persuasive accounts of context sensitivity (Smith, 1996, 1998) and therefore are entirely consistent with this type of individual-to-social identity continuum postulated by social identity theory.

Our theoretical model of this online construction process as implemented in the IAC network is only a preliminary effort, and many issues about what types of network will best fit empirical data, together with questions about what types of knowledge will have effects under what conditions, demand further research attention. Still, this connectionist model represents a fundamentally novel type of theory within social psychology, which may have broad applicability to understanding other types of social judgments and even social behavior.

More broadly, these results further support the fundamental idea that close relationships and group membership both involve some sort of merging of self and other (Aron et al., 1991; Turner et al., 1987). This process may deeply influence cognition, affect, and behavior in relationship and group contexts. For example, according to self-categorization theory, such outcomes as cooperation, altruism, and social influence are due to group membership becoming part of the self (Turner et al., 1987). Our study shows that the self is implicitly accessed and has effects when people construct descriptions of their partner or in-group to answer questions about their traits. We propose that the self is also implicitly accessed when people construct more complex judgments about a relationship partner or group, such as the following: “Should I cooperate with him [or them]?” “Should I

---

**TABLE 1: Simulation Results: Activation of Four Traits for Varying Values of Connection Between Self and Partner**

<table>
<thead>
<tr>
<th>Trait Number</th>
<th>Self-Partner Link = +1 (main simulation)</th>
<th>Self-Partner Link = .50</th>
<th>Self-Partner Link = .25</th>
<th>Self-Partner Link = .10</th>
<th>Self-Partner Link = 0.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trait 1 (yes/yes)</td>
<td>0.9</td>
<td>0.94</td>
<td>0.94</td>
<td>0.93</td>
<td>0.89</td>
</tr>
<tr>
<td>Trait 2 (yes/no)</td>
<td>0.46</td>
<td>0.61</td>
<td>0.73</td>
<td>0.83</td>
<td>0.91</td>
</tr>
<tr>
<td>Trait 3 (no/yes)</td>
<td>−0.15</td>
<td>−0.16</td>
<td>−0.18</td>
<td>−0.18</td>
<td>−0.19</td>
</tr>
<tr>
<td>Trait 4 (no/no)</td>
<td>−0.18</td>
<td>−0.19</td>
<td>−0.19</td>
<td>−0.19</td>
<td>−0.19</td>
</tr>
</tbody>
</table>
help her [or my group] at some cost to myself?” and “Should I do what he [or the group] wants me to?” The result should be an increased likelihood of positive answers to such questions. If so, the process investigated in this study could produce cooperation, altruism, and social influence. Understanding of such phenomena (as well as parallel outcomes in close relationships) may be advanced if we see them as stemming from implicit effects of the self on the construction of judgments about relationship partners or groups.

NOTES

1. In this design, there can be no neutral baseline condition against which to compare the match and mismatch conditions, so we can use the terms facilitation and inhibition only in a relative (not absolute) sense. Note, however, that even in priming paradigms in which a neutral priming condition is sometimes included, many issues cloud the interpretation of facilitation or inhibition in absolute terms (see Jonides & Mack, 1984).

2. The non-Greek in-group (also used in Smith & Henry, 1996) may seem to be a meaningless negation (something like the category “all objects that are not doughnuts”). But at our university, like many schools with large and influential fraternity/sorority systems, non-Greek students do tend to think of themselves as independent students and in other ways to consider Greek organization members and non-members as distinct and meaningful groups.

3. An in-group is defined by its differentiation or contrast with the out-group (Turner et al., 1987), so analyses involving the in-group also included perceptions of the out-group as a factor. However, the partner’s personal characteristics are in no way defined by contrast to Bill Cosby’s, so it makes little conceptual sense to include perceptions of the stranger in the analysis.

REFERENCES


Received September 19, 1997
Revision accepted June 5, 1998