Social and Personality Psychology Compass 3/1 (2009): 64–76, 10.1111/j.1751-9004.2008.00160.x

Distributed Connectionist Models in Social Psychology

Eliot R. Smith*
Indiana University

Abstract

Distributed connectionist models of mental representation (also termed PDP or parallel distributed processing, or ANN or artificial neural networks) constitute a fundamental alternative to the associative or schematic models that have been much more prevalent in social psychology. A connectionist model is made up of a large number of very simple processing units, richly interconnected and able to send signals to each other depending on their momentary activation levels. No individual processing unit represents a meaningful concept; instead, overall patterns of activation hold representational meaning. This article emphasizes the novel properties of connectionist representation that might appeal to theorists and researchers in social psychology, including their context sensitivity and flexibility, ability to represent prototypes and exemplars within a single network, and ability to determine whether a stimulus is familiar even before the stimulus can be identified or categorized.

Virtually every theory in social psychology makes reference to mental representations, including beliefs, attitudes, stereotypes, person impressions, autobiographical memories, or the self-concept. Therefore, the assumed properties of mental representations fundamentally influence the predictions that will be made by our theories. Social psychologists have often unquestioningly adopted a standard set of assumptions about the nature of representations (such as associative network or schema models; Smith, 1998), assumptions that are pleasantly straightforward, familiar, and intuitive. This article describes an alternative, distributed connectionist models of representation, and discusses some of their less intuitive properties, with the goal of encouraging social psychologists to explore the implications of these models for their theories. As I will argue, connectionist models can explode misleading dichotomies, unify seemingly distinct types of processing, and naturally generate properties (such as the flexible, context-sensitive nature of representations) that are often found in social psychological research.

The properties of connectionist models (also termed PDP or parallel distributed processing, or ANN or artificial neural networks) are explored by running computer simulations (Van Overwalle, 2007). However, this article is not aimed at persuading social psychologists to conduct simulations or even to understand them in detail – partly because experience teaches that the vast majority of social psychologists are unresponsive to such persuasive attempts. Instead, the goal is to describe features of connectionist models in a non-technical way, at the conceptual level that will be most useful for typical social psychological theory building. Readers who wish to understand these properties in more depth are encouraged to follow the references to see how the properties arise and how they are revealed by computer simulations.

Fundamental Concepts

For detailed presentations on the basic properties of connectionist networks aimed at social psychologists, see Smith (1996, 1998) or van Overwalle (2007). Here is a quick overview.

Nodes and activation levels

A connectionist model is intended to be loosely analogous to a network of biological neurons. It consists of a large number of simple processing units or nodes (analogous to individual neurons), each of which has a level of activation that fluctuates over time. The nodes are richly interconnected and can send signals over these connections (analogous to synapses between neurons) to influence each other's activation levels. Each connection has a specific strength or weight, and the signal sent over the connection is multiplied by the weight. Each node integrates the input signals it receives over all of its incoming connections and uses that sum to determine its own activation level at the next moment. For example, assume nodes 1 and 2 have connections to node 3. If a_i is the activation level of node number i, and w_{ij} is the weight on the connection from node i to node j, then the total input received by node 3 is $a_1 * w_{13} + a_2 * w_{23}$. The activation level of node 3 at the next moment will depend on that summed input as well as the node's current activation level. In turn, node 3 will send an output signal that depends on its activation level to other nodes over its own outgoing connections.

What do nodes and their activation levels actually represent? In a 'localist' connectionist network, each node represents a meaningful concept or proposition (e.g., 'dog' or 'Jack loves Jill'), and the activation level might represent the network's degree of confidence that the concept is present or that the proposition is true. In contrast, in a 'distributed' network, only patterns of activation across many nodes, not individual nodes, have any conceptual meaning (Thorpe, 1995). This article focuses on the latter. To grasp the idea of representation in a distributed network, consider the pixels in a computer monitor. The brightness or color of any individual pixel has no meaning in itself (as it would in a localist interpretation). Yet, by taking on distinct patterns of brightness or color, the entire set of pixels can represent a very large number of distinct meaningful images (much larger than the number of individual pixels).

Connection weights and learning

The connection weights in a network are not constant, but are gradually altered by a learning rule, which uses information that is locally available at a particular node (input signals and activation levels) to fine-tune the weights. Because the connection weights control the network's processing, learning produces changes in the network's processing dynamics as a result of experience. Stated differently, the weights constitute a connectionist network's stored knowledge, which controls the patterns of activation the network will generate based on specific inputs.

Connectionist networks can perform several useful functions. For example, a network can map input patterns into output patterns; for example, turning inputs representing the spelling of a word into outputs representing the word's pronunciation (Plaut, McClelland, Seidenberg, & Patterson, 1996), or turning inputs representing attributes of an object into an output representing the positive or negative evaluation of that object. Other connectionist networks generate a complete pattern of activation that has been experienced in the past, given inputs of a partial, incomplete version of that pattern. For example, a network might take inputs representing the sight of a furry tail in the distance and the sound of a bark and generate a pattern representing the concept of dog. This could be thought of as categorizing the input as a dog, or as retrieving the dog concept from memory in response to relevant cues, but it amounts to reconstructing a complete previously experienced pattern of activation from a partial version of it. The activation pattern the network tends to produce (given specific inputs) is termed an *attractor* (Hertz, 1995).

Perhaps the most powerful property of connectionist networks is that configurations of weights that perform useful functions like those just described can be produced by learning. No external process has to explicitly set the connection weights. Thus, for example, a network can learn to generate pronunciations from spelling patterns by being trained with a large number of (spelling, pronunciation) pairs, with a learning rule fine-tuning the connection weights after the presentation of each pair. Importantly, the network can eventually generate pronunciations not only for specific words that were presented during training, but also for other similar words: it has the ability to generalize. In the same way, a network can learn to reproduce patterns from partial cues by being repeatedly exposed to those patterns themselves. The weight changes during the learning process make the network better able to reconstruct the patterns of activation that it frequently entered in the past. In other words, learning creates and strengthens attractors corresponding to those patterns.

The main limitation of localist networks (compared to distributed networks) is that they cannot be constructed by learning. In a localist network, assuming a set of nodes representing distinct concepts (e.g., dog, fur, bark) are already present, a learning rule could record co-occurrences among those concepts by strengthening the links connecting them accordingly. Once the links are strengthened, activation of one node (e.g., bark) could cause frequently co-occurring nodes (dog, fur) to become active, a desirable property. The problem is that such a network cannot learn any new concepts. Because each concept is represented by its own node, a new concept would have to be represented by an entirely new node, and so some type of external process (outside the network itself) would have to create a new node and link it appropriately into the network. In contrast, a distributed connectionist network can learn a new pattern (representing a new concept) across the existing nodes through the operation of the standard learning rule. Thus, as distributed patterns representing different types of dogs are processed by the network, it would learn to reproduce the 'dog' pattern given a subset of its attributes. And a pattern representing a new concept (such as cat), which would include both similarities and differences from the dog pattern, could also be learned if it began to appear (McClelland & Rumelhart, 1986).

Properties of Connectionist Representations

Connectionist representations have unique properties that often challenge everyday intuitions about memory. To illustrate this point, this article describes several dichotomies – pairs of seemingly opposite properties – that are reconciled and integrated in connectionist models. As Seidenberg (1993, p. 234) wrote: 'Imagine that ... there is a type of knowledge representation that encodes both rule-governed cases and exceptions to the rules. Given the [prevailing] stock of theoretical ideas ... [this] proposal could only be taken as vacuous. Yet encoding both types of knowledge is what some kinds of connectionist networks do ... Here is something that is not a wave and not a particle, but acts like both.' Seidenberg's case is our first example.

Rules and exceptions

This issue has been explored most extensively regarding word pronunciation. Standard theories proposed that general rules govern pronunciation of most words (e.g., SAVE, PAVE, WAVE), so specific exceptions (HAVE) whose pronunciation departs from the rule must be learned individually. A variety of evidence appears to support these ideas, such as (a) the fact that novel nonwords (MAVE) are usually pronounced according to the rules and (b) the fact that frequency in the language correlates with the speed of pronunciation of exception words but not rule-based words (consistent with the notion that the former but not the latter need to be memorized and represented individually). However, Plaut et al. (1996) developed a model using a single connectionist mechanism (a network trained with examples of words based on their real language frequencies) to account for all this evidence. In this model, every word that is learned forms its own attractor in the network. Groups of similarly spelled words with similar sounds produce some generalization, so that a similar but not previously encountered input may also be pronounced in the same way. After training, the network of Plaut et al. (1996) pronounces regular words (SAVE, GAVE) correctly and can also pronounce HAVE appropriately because its pattern forms its own specific attractor. Novel strings like MAVE will be pronounced like SAVE because more regular words than exception words are similar to the non-word. This means that the intuitively appealing dual-route idea is not necessary, and a single connectionist mechanism can adequately explain evidence that previously seemed to require separate processes of rule application and exception memorization.

Representation retrieval versus context-sensitive reconstruction

Traditional theories invoking associative networks or schemas (as well as our intuitive understanding based on the metaphor of memory 'storage' and 'retrieval') postulate that once a representation of an object or concept is constructed, it is stored in memory to be retrieved intact at a later time (or possibly forgotten). Each memory representation is separate and discrete. These traditional theories are consistent with our everyday intuitions, shaped by experiences of storing and retrieving papers in filing cabinets. In contrast, a connectionist memory does not hold a discrete representation for each object or concept; in fact, it maintains no discrete representations at all. Representations are not retrieved but reconstructed as states of activation in the network, generated by the current inputs and the configuration of weights. These alternative assumptions, fleshed out into formal theories, fit data from cognitive experiments on learning and memory (e.g., McClelland, McNaughton, & O'Reilly, 1995). Conceiving of representations as reconstructed (rather than 'searched for' or 'retrieved') allows us to understand that all aspects of the person's state (e.g., mood, goals, physical location) will influence the exact details of what is reconstructed; in other words, reconstructions will differ across time and contexts. This type of context sensitivity is characteristic of human memory function (Clark, 1993). For example, specific features of a dog such as 'barks' and 'fur' may lead to the reconstruction of a representation of a fierce guard dog (in the context of a security area at a military installation) or a tiny Pekingese (in the context of a grandmother's living room) (Conrey & Smith, 2007).

If an infinite variety of context-specific representations of any concept can be reconstructed, how can a perceiver recognize them all as instances of the same thing? That is, how can we account for stability as well as context sensitivity in our conceptual knowledge? One plausible answer is that experience with a domain allows the person to learn to activate roughly the same network pattern in different contexts – to rely on the invariant features that define the concept under surface variations. Thus, a young child may have very different concepts for 'fierce guard dog' and 'tiny lapdog', but an older child or adult may attain the understanding that they are both dogs (Keil, 1989; Conrey & Smith, 2007). The construction of a (more or less) context-free representation is an intellectual accomplishment resting on an ability to focus on the most essential features, rather than being the unproblematic initial or default state (Clark, 1993).

Exemplars versus prototypes

Social psychologists have often contrasted exemplars and prototypes as models of representation for certain types of concepts, such as social groups (Smith & Zarate, 1992). That is, one could maintain representations of a number of individual group members including their common properties (e.g., a number of individual politicians), or use a single abstracted representation of a typical or representative group member (a prototypical politician). The latter possibility is often intuitively favored based on the notion of 'economy of storage'. It is also consistent with evidence that exemplars that are highly typical of a category can be readily categorized even if they have never been encountered before, which suggests that the average or prototype is part of the category representation. However, other evidence suggests that people often maintain representations of specific exemplars (Park & Judd, 1990; Smith, 1998; Whittlesea, 1987). For example, old (previously encountered) exemplars can be categorized more readily than new exemplars that are equally typical of the category, suggesting that representations of specific exemplars are maintained and used.

Connectionist models offer an integrated representation that has properties of both exemplars and prototypes. They automatically extract communalities when learning about a series of exemplars, yet also store the specific details of any particular exemplar that is frequently or recently encountered (McClelland et al., 1995). After experience with a variety of dogs, a perceiver will be able to readily recognize a typical dog (even one that has never been seen before). But the neighbor's dog Sassy, frequently seen on the street, will also be readily recognized even if Sassy is quite different from the typical dog (e.g., because she is very small or unusually colored). No explicit process of summarizing or averaging ever needs to be invoked, so this type of model avoids difficult questions like 'when should I switch from storing information about individual exemplars to storing an abstracted summary instead, to save storage space?'.

Causal ordering among cognitions

Do beliefs cause attitudes or vice versa? Certainly, learning something negative about a person can make you like the person less, but it is also true that dislike arising for irrelevant reasons can be rationalized by adopting beliefs that the person has negative characteristics. Social psychologists are often uncomfortable with observations like this partly because useful data-analytic methods like path analysis require clear and unambiguous assumptions about causal precedence. Lengthy debates in the literature, for example, have dealt with the relative causal priority of 'cognition' versus 'affect'. Connectionist models strongly suggest that such questions will often be unanswerable. In a connectionist network, multiple cognitions can adjust to each other to achieve the overall best fit (maximum coherence or harmony) in the system (Thagard, 1989). A positive attitude will fit well with positive beliefs, a negative attitude with negative beliefs. Once the system has reached the state in which all elements fit best, a change in any of the elements (e.g., a belief or an attitude) will cause all the cognitive elements to adjust accordingly to find a new optimum. Therefore, it is impossible to impose any unidirectional causal ordering, for a change in any cognitive element can produce changes in the others.

Top-down versus bottom-up processing

Often, top-down processing (driven by the perceiver's prior knowledge or expectations, such as social stereotypes) is contrasted with bottom-up processing (driven by the currently available perceptual information, such as a target person's behaviors). Various factors are said to encourage perceivers to rely more on one versus the other type of processing. Again, connectionist networks eliminate this seeming dichotomy by making it clear that all processing is necessarily both top-down and bottom-up. The generation of any activation pattern or representation in a network depends on both the current inputs (bottom-up) and the connection strengths in the network (which constitute the network's stored knowledge, top-down). Neither type of processing can occur without the other also being involved.

Knowing that something is familiar before knowing what it is

Suppose you are looking through a file cabinet to find a particular document. If it is there, you won't know until you find it, and if it is not there, you will not know until you have unsuccessfully searched through all the records. This argument suggests that it is impossible for a memory system to indicate whether a representation is old or new (familiar or unfamiliar) before being able to identify what it is. Yet, people can have a subjective sense that an object is familiar before they can identify precisely what it is or when they have seen it before (i.e., before retrieving a memory representation of the object that was formed on a previous occasion). For example, subliminal exposures (not consciously perceptible) of a stimulus can produce feelings of familiarity that result in increased liking for the stimulus (Bornstein & D'Agostino, 1994), without the perceiver's even being able to report the fact of the previous encounter. As Smith (2000) described, several different types of connectionist network have this apparently logically impossible property: they can generate a signal indicating the familiarity versus novelty of a stimulus pattern, before they settle into a state corresponding to identifying or categorizing the stimulus. The details of how they do this are technical, but aspects of the flows of activation in the network early in the process can indicate whether the current inputs represent a new or familiar pattern, before the network settles into its final activation pattern (corresponding to identifying or categorizing the pattern). Familiarity information is available before identity information.

Summary

As these examples have shown, connectionist networks can often unify the poles of conceptual dichotomies or show the way out of seeming dilemmas. Our intuitions about mental representation and memory are based on familiar metaphors of storage and retrieval as in a file cabinet. Because connectionist networks operate very differently, it is hard for us to grasp intuitively that processing is top-down and bottom-up at the same time, that memory always generates a context-sensitive reconstruction, that exemplars and prototypes can be integrated into a single representational framework, or that a memory system can indicate whether a stimulus is familiar before being able to tell what it is. But these counterintuitive properties are a key to understanding the unique capabilities of connectionist networks – and to incorporating their strengths into theoretical models in social psychology.

Current Models in Social Psychology and Future Directions

A handful of connectionist models have been developed within social psychology, including Smith and DeCoster (1998) on stereotype learning and application, Queller and Smith (2002) on stereotype change, Kashima and Kerekes (1994) and van Overwalle and Labiouse (2004) on impression formation, van Overwalle and Siebler (2005) on persuasion, Smith, Coats, and Walling (1999) on self-perception and group identification, and Monroe and Read (2008) on attitude structure and function. Several models are collected in edited books by Read and Miller (1998) and van Overwalle (2007). The latter book usefully includes software allowing the reader to duplicate and modify the published simulation runs. Other discussions of the implications of connectionist thinking include Bassili and Brown

(2005) on attitude change and Mischel and Shoda's (1995) CAPS model of personality and social behavior. However, the great majority of connectionist models in the literature deal with issues in cognitive or developmental rather than social psychology; Elman, Grossberg, McClelland, O'Reilly, Plaut, and Seidenberg are a few of the major contributors in these areas and the interested reader is urged to seek out their work.

As these examples suggest, virtually all connectionist models within social psychology have dealt with specific cognitive structures and processes, such as stereotyping or attitude structure. Two important future directions for the cognitive sciences as a whole include modeling of *complete agents* by incorporating perception and action rather than modeling specific, isolated tasks; and modeling *interactions of multiple agents* over time. Connectionist models are well suited for exploring these two areas of embodied and distributed cognition (see Smith & Semin, 2004).

Complete-agent models (whether implemented as embodied robots or as simulated agents within a simulated environment) are not limited to a specific task but seek to capture the entire set of processes by which an agent senses its environment and behaves adaptively, for example by finding food or avoiding predators. An important tool is reinforcement learning (Sutton & Barto, 1998), a procedure that allows an agent to learn about the positive and negative consequences of its actions so that it can refine its behaviors in the future. Although reinforcement learning has not been applied to date in social psychology, it appears to offer a natural framework for modeling attitude formation. That is, the reinforcement signal reflecting the good or bad consequences of the agent's behavior could be used to attach information about positive or negative valence to representations of specific objects or environmental locations that have been encountered. In social psychology, Fazio, Eiser, and Shook (2004) modeled an agent that encounters many food items called beans in sequence, and can choose whether or not to consume each one. Good beans (if consumed) increase the agent's energy, but some beans are bad and consuming them causes illness and reduces energy. The catch is that if a bean is not consumed, the agent does not learn whether it is good or bad. Although this model does not use reinforcement learning, it illustrates the fundamental idea of the whole-agent approach: only through modeling perception and behavior as well as the details of knowledge representation can phenomena like attitudes be embedded in an overall framework displaying their functionality and adaptive value.

Another area for future advances is to conceptualize distributed cognition in interacting groups of agents in terms similar to connectionist networks. Each individual agent is regarded as a node, and connections allow information to flow from one agent to others (Mason, Conrey, & Smith, 2007). Connection weights represent how much each agent pays attention to or uses information provided by other agents. Hutchins (1991) pioneered this approach with a (localist) connectionist model of each individual supplemented by connections between individuals. The connections implemented the idea that communication makes the activation value of a specific node in one person's mind (representing the strength of the person's belief in that idea) similar to its level of activation in the other person's mind. Related work, combining connectionist models of individual agents with consideration of information flows between agents, has been conducted by van Overwalle and Heylighen (2006).

Conclusions

The limited impact to date of connectionist models in social psychology (compared to other areas of psychology) suggests addressing the question: why should social psychologists pay any attention to connectionist models? Briefly, here are two answers.

First, the properties of connectionist models sketched above can liberate our theoretical thinking. No longer need a theorist agonize over (for example) whether to assume exemplar or prototype representations, or whether to assume that attitudes are retrieved or actively constructed, for a connectionist theory can integrate both of these apparently polar opposites. As long ago as the 1930s, Gestalt theorists tried to capture ideas like these, for example in emphasizing that wholes could have 'emergent' properties that differed from the sum of their parts, and of course Gestalt ideas had great influence within social psychology. Now, we can go beyond fuzzy words like 'holism' and 'Gestalt' to specify actual mechanisms that display these properties, grounded in concrete results of computer simulations.

Second, the properties discussed in this article falsify the common belief that the distinction between connectionist and other modes of representation (e.g., schemas or associative networks) is too low-level an issue to have any implications for social psychological theory. Representational structures constrain cognition and behavior, so different assumptions about representation afford different conclusions about social behavior (Smith, 1998). I hope that this article demonstrates some of the attractive properties of connectionist models, which invite attention and increased adoption by social psychological theorists. At the very least, I encourage social psychologists to reject the standard, intuitive file-cabinet metaphor for memory representation, as well as the language supporting that metaphor (e.g., terms such as 'storage' and 'retrieval'). Human memory does not have much in common with a filing cabinet, for example in its ability to perform flexible, context-sensitive reconstruction, and displacing our comfortable intuitions is the first step freeing our theoretical imagination.

Acknowledgements

Preparation of this article was facilitated by NSF Grant BCS-0527249. The author is grateful to Austin Chapman, Elizabeth Collins, Elise

Percy, Charles Seger, and Ashley Waggoner for helpful comments and suggestions.

Short Biography

Eliot R. Smith is Chancellor's Professor of Psychological and Brain Sciences at Indiana University, Bloomington, having moved from Purdue University in 2003 to join the IU faculty. He received his PhD from Harvard University. Much of Dr. Smith's research focuses on prejudice and intergroup behavior, especially the role of the emotions that people experience when they think of themselves as members of a group (such as a political party or an ethnic, religious, or national group). He also pioneered the study of connectionist models of mental representation within the field of social psychology. Other major research interests include person perception and stereotyping, as well as embodied cognition, the investigation of the role of the body in shaping people's thoughts, feelings, and behaviors – especially toward other people. His research has been funded extensively by the National Science Foundation and the National Institute of Mental Health. Dr. Smith has been honored with the Gordon Allport Prize for contributions to the study of intergroup relations, the Society for Personality and Social Psychology's Theoretical Innovation Prize, and the Thomas M. Ostrom Award for contributions to social cognition. He has served as Editor of *Personality and Social Psychology Review*, and Associate Editor of *Journal of Personality and Social Psychology: Attitudes and Social Cognition*.

Endnote

* Correspondence address: Department of Psychology, Indiana University, Bloomington, 1101 E. Tenth Street, Bloomington, IN 47405, USA. Email: esmith4@indiana.edu

References

Bassili, J. N., & Brown, R. D. (2005). Implicit and explicit attitudes: Research, challenges, and theory. In D. Albarracin, B. T. Johnson & M. P. Zanna (Eds.), *The Handbook of Attitudes* (pp. 543–567). Mahwah, NJ: Lawrence Erlbaum Associates.

Bornstein, R. F., & D'Agostino, P. R. (1994). The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual fluency/attributional model of the mere exposure effect. *Social Cognition*, **12** (2), 103–128.

Clark, A. (1993). *Associative Engines: Connectionism, Concepts, and Representational Change*. Cambridge, MA: The MIT Press,

Conrey, F. R., & Smith, E. R. (2007). Attitude representation: Attitudes as patterns in a distributed, connectionist representational system. *Social Cognition*, **25**, 718–735.

Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology*, **87** (3), 293–311.

Hertz, J. (1995). Computing with attractors. In M. A. Arbib (Ed.), *Handbook of Brain Theory and Neural Networks* (pp. 230–234). Cambridge, MA: MIT Press.

Hutchins, E. (1991). The social organization of distributed cognition. In L. Resnick, J. Levine

& S. Teasley (Eds.), *Perspectives on Socially Shared Cognition* (pp. 283–307). Washington, DC: American Psychological Association.

Kashima, Y., & Kerekes, A. R. Z. (1994). A distributed memory model of averaging phenomena in person impression formation. *Journal of Experimental Social Psychology*, **30** (5), 407–455.

Keil, F. (1989). *Concepts, Kinds, and Cognitive Development*. Cambridge, MA: MIT Press.

Mason, W. A., Conrey, F. R., & Smith, E. R. (2007). Situating social influence processes: Dynamic, multidirectional flows of influence within social networks. *Personality and Social Psychology Review*, **11**, 279–300.

McClelland, J. L., & Rumelhart, D. E. (1986). A distributed model of human learning and memory. In J. L. McClelland, D. E. Rumelhart & The PDP Research Group (Eds.), *Parallel Distributed Processing* (Vol. 2, pp. 170–215). Cambridge, MA: MIT Press.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, **102** (3), 419–457.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, **102** (2), 246–268.

Monroe, B. M., & Read, S. J. (2008). A general connectionist model of attitude structure and change: The ACS (Attitudes as Constraint Satisfaction) Model. *Psychological Review*, **115** (3), 733–759.

Park, B., & Judd, C. M. (1990). Measures and models of perceived group variability. *Journal of Personality and Social Psychology*, **59** (2), 173–191.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S. & Patterson, K. E. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domain. *Psychological Review*, **103**, 56–115.

Queller, S., & Smith, E. R. (2002). Subtyping versus bookkeeping in stereotype learning and change: Connectionist simulations and empirical findings. *Journal of Personality and Social Psychology*, **82**, 300–313.

Read, S. J., & Miller, L. C. (1998). *Connectionist Models of Social Reasoning and Social Behavior*. Mahwah, NJ: Lawrence Erlbaum Assoc.

Seidenberg, M. S. (1993). Connectionist models and cognitive theory. *Psychological Science*, **4**, 228–235.

Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology*, **70** (5), 893–912.

Smith, E. R. (1998). Mental representation and memory. In D. Gilbert, S. Fiske & G. Lindzey (Eds.), *Handbook of Social Psychology* (4th edn, Vol. 1, pp. 391–445). New York, NY: McGraw-Hill.

Smith, E. R. (2000). Subjective experience of familiarity: Functional basis in connectionist memory. In H. Bless & J. Forgas (Eds.), *The Role of Subjective States in Social Cognition and Behavior* (pp. 109–124). Philadelphia, PA: Psychology Press.

Smith, E. R., & DeCoster, J. (1998). Knowledge acquisition, accessibility, and use in person perception and stereotyping: Simulation with a recurrent connectionist network. *Journal of Personality and Social Psychology*, **74**, 21–35.

Smith, E. R., & Semin, G. R. (2004). Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology*, **36**, 53–117.

Smith, E. R., & Zárate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, **99**, 3–21.

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.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, **12**, 435–467.

Thorpe, S. (1995). Localized and distributed representations. In M. A. Arbib (Ed.), *Handbook of Brain Theory and Neural Networks* (pp. 549–552). Cambridge, MA: MIT Press.

van Overwalle, F. (2007). *Social Connectionism*. New York, NY: Psychology Press.

Van Overwalle, F., & Heylighen, F. (2006). Talking nets: A multiagent connectionist approach to communication and trust between individuals. *Psychological Review*, **113**, 606–627.
van Overwalle, F., & Labiouse, C. (2004). A recurrent connectionist model of person impression formation. *Personality and Social Psychology Review*, **8**, 28–61.
van Overwalle, F., & Siebler, F. (2005). A connectionist model of attitude formation and change. *Personality and Social Psychology Review*, **9**, 231–274.
Whittlesea, B. W. A. (1987). Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 3–17.