THEORETICAL CONTRIBUTION

What Do Connectionism and Social Psychology Offer Each Other?

Eliot R. Smith
Purdue University

Social psychologists can benefit from exploring connectionist or parallel distributed processing models of mental representation and process and also can contribute much to connectionist theory in return. Connectionist models involve many simple processing units that send activation signals over connections. At an abstract level, the models can be described as representing concepts (as distributed patterns of activation), operating like schemas to fill in typical values for input information, reconstructing memories based on accessible knowledge rather than retrieving static representations, using flexible and context-sensitive concepts, and computing by satisfying numerous constraints in parallel. This article reviews open questions regarding connectionist models and concludes that social psychological contributions to such topics as cognition–motivation interactions may be important for the development of integrative connectionist models.

Probably every social psychologist, paging through the general journals of psychology (such as *Psychological Review* or *Psychological Science*) or the cognitive journals (such as the various sections of the *Journal of Experimental Psychology*) has noted the frequent appearance of articles applying connectionist or parallel distributed processing (PDP) models. Connectionism emerged as an intellectual movement in the early 1980s out of various earlier precursors (including Rosenblatt, 1962, and Grossberg, 1976), and its maturity was marked by the 1986 publication of the two-volume "PDP bible" (McClelland, Rumelhart, et al., 1986; Rumelhart, McClelland, et al., 1986). Today connectionism exerts a strong and growing influence in many areas of cognitive psychology, ranging from research on low-level visual perception to higher level processes such as language processing, categorization, and decision making.

Like an earlier major transition in psychology—that from behaviorism to information-processing cognitive models in the 1950s—the rise of the connectionist approach has been characterized as a scientific revolution or paradigm shift (Schneider, 1987). Enthusiastic proponents have written comments like the following:

> Connectionism . . . promises to be not just one new tool in the cognitive scientist's toolkit but, rather, the catalyst for a more fruitful conception of the whole project of cognitive science. (Clark, 1993, p. ix)

> Every once in a while, by some unknown means, people come up with ideas that change the way we think. I believe that connectionism embodies some genuinely original ideas. In particular, there is a novel way of representing knowledge—in terms of patterns of activation over units encoding distributed representations. These ideas have consequences that are just beginning to be explored. Imagine that it is 15 years ago and I propose to you that there is a type of knowledge representation that encodes both rule-governed cases and exceptions to the rules. Given the stock of theoretical ideas available at that time, my 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, 1993, p. 234)

> If a certain family of connectionist hypotheses turn out to be right, they will surely count as revolutionary. . . . There is no question that connectionism has already brought about major changes in the way many cognitive scientists conceive of cognition. . . . If we are right, the consequences of this kind of connectionism extend well beyond the confines of cognitive science, since these models, if successful, will require a major reorientation in the way we think about ourselves. (Ramsey, Stich, & Garon, 1991, pp. 199–200)

If social psychologists have dipped into the connectionist literature at all, they may be impressed by the sense of excitement

Preparation of this article was supported by Research Grant R01 MH46840 and Research Career Development Award K02 MH01178 from the National Institute of Mental Health. I am grateful to Jerry Busemeyer, Susan Coats, Greg Francis, Tory Higgins, Jim Sherman, Jaron Shook, and John Skowronski for comments on an earlier draft.

Correspondence concerning this article should be addressed to Eliot R. Smith, Department of Psychological Sciences, Purdue University, West Lafayette, Indiana 47907-1364. Electronic mail may be sent via the Internet to esmith@psych.purdue.edu.

Journal of Personality and Social Psychology, 1996, Vol. 70, No. 5, 893–912

evident in such statements as these, but they may wonder "what's in it for us?" For example, might these models help account for the way people flesh out observed information using accessible stored knowledge? Could they shed light on the relations between conscious and nonconscious processing, or on the mental representation of attitudes, person impressions, or stereotypes and the ways they are activated to affect judgment and behavior? In this article I propose answers to these questions. I believe not only that this intellectual movement can benefit our research and theory in social psychology but also that we have much of substance to offer in return.

This article is organized into several sections. After briefly outlining the fundamental assumptions of the symbolic models that are most commonly used in social psychology today, I present an overview of distributed connectionist models and their properties. I then describe reasons for seriously considering the implications of these models for significant issues in social psychology. Next I discuss several criticisms and open questions concerning connectionist models. Throughout these sections, I make a special effort to cite accessible or tutorially oriented discussions to which readers can refer for further detail. I conclude by reminding readers that the cognitive revolution that overthrew behaviorism was not so much of a revolution for social psychology, which had been "cognitive all along." Similarly, fundamental aspects of the connectionist revolution—particularly its focus on dynamic properties of cognition—may not be so revolutionary in social psychology either.

Traditional Symbolic Models and Connectionist Models

Symbolic Models

A brief review of the assumptions of traditional models can serve as background for the presentation of connectionist models. The theoretical assumptions of virtually all models in social psychology today (with only a few exceptions deriving from behaviorist or Gibsonian approaches) are representative of those that have prevailed throughout scientific psychology since the cognitive revolution of a generation ago. The essential ideas draw on language and logic (see Clark, 1993, Chapter 1; Fodor, 1987; Smolensky, 1989). Internal representations are constructed from languagelike symbols (concepts) that can be combined in structured ways to encode propositions. For example, a person could use the concepts of "Sam" and "honest" to construct a representation encoding the idea "Sam is honest." Thus, having a belief or a thought is very much like having a sentence in one's head (Churchland & Sejnowski, 1989). Such representations are manipulated by rules that perform logical inferences and the like, embodied in a computerlike symbol processor or "physical symbol system" (Newell, 1980).

Among the most important assumptions of these symbolic models are the following.

1. Representation construction. Cognitive representations (such as beliefs, schemas, attitudes, stereotypes, or person impressions) are dynamically constructed by the perceiver out of simpler, atomic representations (concepts).

2. Discrete representations. Representations are stored and maintained as discrete and separate units; one can be added, changed, or accessed without changing or accessing others.

3. Representation–process distinction. Representations of mental content are distinct from the processes that operate on them; unless altered by some process, representations are static and unchanging, like words on a page.

4. Process. Rule-governed processes operate on representations to transform them or to generate new representations encoding inferences, plans for behavior, and the like.

These traditional assumptions about representation and process, which prevail throughout most of cognitive as well as social psychology, rest on a metaphor of the mind as symbol processor. This theoretical viewpoint grew out of research on solving well-defined problems (Newell & Simon, 1972), in which a clear-cut solution is to be attained by searching a defined and limited body of knowledge. For example, applied to person perception, the assumption would be that there is a fixed, limited set of schemas, stereotypes, or traits that can be used to characterize a person, and the perceiver's job is simply to find the best-fitting one and to apply it. In Wyer and Srull's (1989) model of person perception, the perceiver is said to search through schemas or concepts, stored as discrete representations in a "storage bin", until one is found that adequately fits the available information about the target person.

Symbolic models are appealing for many reasons. They have obviously been fruitful in motivating research in social and cognitive psychology. The best-known theoretical landmarks of our field are of the symbolic sort (e.g., Hamilton, Katz, & Leirer, 1980; Higgins, 1989; Wyer & Srull, 1989). An additional underlying reason may be the good fit between symbolic models and naive psychology (which philosophers term *folk psychology;* see Clark, 1993, Chapter 10). Just as in everyday life we explain actions by attributing propositionally represented beliefs and goals to actors, psychologists do the same thing, although in more sophisticated ways.

Connectionist Models

Connectionist models rest on a very different set of fundamental assumptions (see Churchland & Sejnowski, 1992, Chapter 3, or the first several chapters of Rumelhart, McClelland, et al., 1986 for more extensive tutorial introductions). They can be discussed at two distinct levels.

Level 1: Units, links, and activation. At the most fundamental level, a connectionist network contains many simple processing units, interconnected by unidirectional links that transmit activation. Units are often assumed to perform a particularly simple computation: forming a weighted algebraic sum of all their inputs (which may be positive or negative in sign) and generating output that is a monotonic but nonlinear function of the summed input. All the complexity of a connectionist model resides in the overall "architecture" of the model and in the pattern of interconnections among units. Ordinarily the architecture is fixed; for example, a model may have two layers of units: an input layer that receives input from external sources (as well as possibly from other units) and an output layer that sends output to the outside world (and possibly to other units). If a network has more than two layers, there may be "hidden units" that have neither input nor output connections (see Figure 1). The pattern of interconnections among units may be assumed to be fixed, established a priori to permit

the network to perform some task (see, e.g., Rumelhart, Smolensky, et al., 1986). More often, however, the weights on the interconnections among units are assumed to be shaped by a learning process, as discussed below.

More specifically, each unit is characterized by a time-varying amount of activation (ranging from a minimum value, often 0, to a maximum value). A unit's current activation depends on the activation flowing to it from other units over incoming links as well as, possibly, its prior activation level. In turn, this unit sends activation over its outgoing links to other connected units. If the activation of Unit i is denoted by a_i and the strength or weight on a link from Unit i to Unit j is denoted by w_{ji}, which may be positive or negative, then the total input of Unit j is

$$\sum_i a_i w_{ji}$$

and in a simple model the unit's activation $a_j = f(\Sigma a_i w_{ji})$. Here f is a nonlinear function, often sigmoidal in shape.

Level 2: Patterns, representations, and computation. The low-level description in terms of units and activation leaves unanswered the question of what a unit represents semantically: What is the relationship between the unit activations at the lowest level and the meaning of what the network is processing? Most connectionist models use distributed representations: They identify a semantically meaningful processing state or mental state with a pattern of activation across many units (McClelland & Rumelhart, 1986). Activity of a single unit has no fixed meaning independent of the pattern of which it is a part.

Of course, meaningful activation patterns arise from activity at the lower level. A pattern is elicited in the network as activation flows across links, beginning with a particular pattern of activation received by a set of input units from the outside world. The entire set of connection strengths determines the way activation flows and therefore the activation pattern that will result from any given input. Recall that in symbolic models, symbols are entities that both carry semantic meaning and are the units on which the model's processes operate. A crucial contrast is that in distributed connectionist models, semantic interpretation is attached only to patterns that involve many units, whereas the rules that define the actual operation of the system (the equations that govern the computation and spread of activation levels) are at a lower level. Furthermore, these equations, dealing with continuous flows of activation, are fundamentally different in character than discrete symbol-manipulation rules. Connectionist models of this sort, termed *subsymbolic* by Smolensky (1988), are the primary focus of this article.

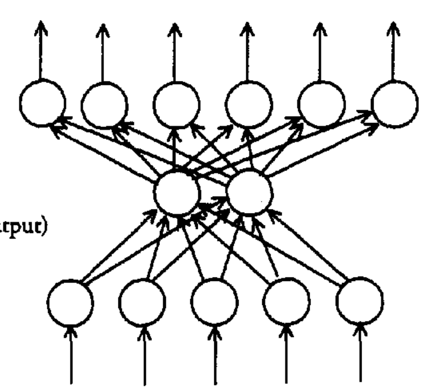

Figure 1. Example of a three-layer feed-forward network. Input units receive activation from sensory receptors or other networks, and output units send their activation onward. "Hidden units" are those that are neither inputs nor outputs. Overall, a network such as this can transform an input pattern (vector of activation values) to a distinct output pattern (also represented by a vector).

The *stored knowledge* of a connectionist network is encoded in the set of connection weights. In a sense, this set constitutes a single representation in which representations of all learned patterns are superposed or "mushed together." An individual *memory trace* is the change in the connection weights in the network produced by the learning algorithm when one input pattern is processed on one occasion. Each memory trace is embodied in changes in many weights, just as each weight is a part of many memory traces. *Retrieval* amounts to reinstatement of a previously experienced pattern of activation, which can be elicited by a particular set of cues presented to the network as inputs.

Distributed representations require new ways of thinking about the nature and function of memory. Traditional symbolic models conceptualize memory in terms of a static file cabinet or storage bin metaphor. Discrete representations are thought of as being inscribed on separate sheets of paper that are stored side by side but can be independently accessed. Even the terms *memory storage, search,* and *retrieval* invoke this familiar metaphor. Connectionist models use a very different type of representation, conventionally termed *distributed* but more correctly characterized as *superposed* (van Gelder, 1991).[1] There is no discrete location for each representation. Instead, the whole network of connection weights is a single representation that contains information derived from many past experiences. Accessing one representation necessarily accesses all, because all representations are encoded in the same set of connection weights. Similarly, adding a new experience changes many weights and therefore alters (perhaps minimally) the representation of all (van Gelder, 1991, p. 45). Instead of *search* and *retrieval,* access to memory might be better thought of with metaphors involving similarity and *resonance.* Thus, we might say that a new stimulus resonates with, and activates, representations in memory that resemble it (Estes, 1994, p. 14; see Ratcliff, 1978).

Though a wide variety of connectionist models with many detailed differences have been proposed for various tasks, some properties generally apply to models that use distributed representations (Churchland & Sejnowski, 1992, Chapter 4; Plate, 1994; Rumelhart & Todd, 1990; van Gelder, 1991).

1. Explicit similarity. With a symbolic representation, units are all distinct and unrelated; units representing related con-

[1] The reason is that a representation can be "distributed" over many units in a nonmeaningful way (Concept A represented by Units 1, 2, and 3; Concept B by Units 4, 5, and 6; etc.). However, the properties of such a representation do not differ substantially from those of a localist (one unit per concept) representation. The key property is not distribution but superposition: the idea that the same units participate in different representations by taking on different patterns of activation.

cepts are just as different as units representing completely unrelated ideas. In contrast, appropriate learning rules construct distributed representations so that similar patterns represent similar concepts. This means, among other things, that the similarity of two concepts can be easily computed from their representations. This property is important in many ways, for example, in matching new exemplars against category prototypes or in retrieving instances from memory that are similar to a newly encountered stimulus.

2. Prototype extraction. If a number of representations are all similar in important respects—say they are multiple instances of a given category, such as politicians—a distributed representation can ignore the details of their differences while preserving their common characteristics. Because of the explicit-similarity property, when the representations of all the instances are stored in a single set of connections, those respects in which they are similar will be reinforced, whereas their differences will tend to cancel out. The resulting representation will emphasize the central features shared by most of the instances; in effect, a prototype representation has been created.

3. Generalization. The other side of the prototype-extraction coin is the ability to generalize. Once a representation of the general characteristics of a category has been formed, new category instances can be treated appropriately on the basis of their similarity to the prototype. Similar inputs (whether new or previously encountered) will tend to be treated similarly, and this is usually desirable.

4. Redundancy. Distributed representations are usually redundant, so that given a portion of a pattern as input the network can reconstruct the whole pattern (pattern completion property). Also, damage to a small number of units or links may somewhat degrade the network's performance but should not completely destroy it (graceful degradation property).

5. Parallel constraint satisfaction. Connectionist networks have the ability to settle into the overall pattern that best fits the current input in light of stored representations of past experiences (Smolensky, 1989). This is just another way of looking at the pattern-completion property: The input can be viewed as a partial pattern, and the network "decides" what complete pattern is most consistent with the input.

In summary, the properties of connectionist systems that use distributed representations are quite different from those of symbolic systems. In the latter, the representation of each separate item uses distinct resources (e.g., units). In contrast, in a distributed system, representations are superposed. In general, this means that each item cannot be retrieved in precisely the same form as it was initially stored; it is better to think of a representation as being *re-created* or *evoked* than as being *searched for*. McClelland, Rumelhart, and Hinton (1986) emphasized this point:

> In most models, knowledge is stored as a static copy of a pattern. Retrieval amounts to finding the pattern in long-term memory and copying it into a buffer or working memory. There is no real difference between the stored representation in long-term memory and the active representation in working memory. In [distributed] models, though, this is not the case. In these models, the patterns themselves are not stored. Rather, what are stored are the *connection strengths* between units that allow these patterns to be re-created.(p. 31)

The re-creation will often be imperfect and subject to influence from the person's other knowledge (such as schemas and scripts)—but this characteristic is typical of actual human memory performance (Carlston & Smith, in press; van Gelder, 1991).

Learning in Connectionist Models

A network may learn a set of connection weights that permit it to perform some task such as mapping a given set of input patterns into desired outputs. As an example, consider a categorization task in which a pattern of stimulus attributes (encoded as varying degrees of activation) is applied to the input units of a network, and one of several possible output patterns becomes active to represent the network's decision as to the category membership of the stimulus. In constructing a network to perform such a categorization task, a learning algorithm is generally used. Initially all connection weights are given values of zero or random values. A training pattern is presented to the input, and the network's output is observed. Using one of several specific procedures (e.g., backpropagation), the weights are then adjusted incrementally to reduce the discrepancy between the network's output and the correct output (reflecting the known category membership of the training stimulus). This process is repeated many times with a given set of training stimuli. After enough training, the weights usually stabilize at values that give adequate performance at categorizing the training stimuli. The network can then be tested by presenting it with new stimuli (not part of the training set) and observing how it categorizes them. This is a description of *supervised learning*, in which the correct outputs for the training patterns are known and used in the training process. The process is somewhat analogous to the statistical technique of regression analysis, for the network learns which input features to use in predicting the output category membership.

Other types of learning are "unsupervised" in the sense that target output values are not required during training. Unsupervised nets can, for example, detect sets of features that covary across a number of input patterns. The process is analogous to the statistical technique of factor analysis, which uncovers patterns of covariation within a single set of variables (not divided into independent and dependent variables). In statistical analysis, such patterns (i.e., factor scores) in turn may serve as inputs for further analysis; in a connectionist model, patterns detected by unsupervised learning can serve as higher level input features to be further processed by other networks.

The most significant points about connectionist learning procedures are that learning is incremental, taking place after the presentation of each training stimulus; learning modifies the connection weights in the network; and (for supervised learning) the modifications are in the direction of reducing the discrepancy between the network's actual response to the current stimulus and the known correct response.

Types of Distributed Connectionist Models

Feed-forward (pattern transforming) networks. Most connectionist models fall into two broad categories. A feed-forward network like that shown in Figure 1 has links from input units, perhaps by way of intervening layers of hidden units, to output

units. When an input stimulus is presented as a pattern of activation levels to the input units, activation feeds forward through the network, ultimately producing a distinct pattern on the output units. Among the applications are categorization (where the input patterns represent exemplars and the output patterns represent category labels) and transformation of information from one representational system into another (e.g., mapping visual appearance of letters into semantic representations of words or phonological representations of pronunciation). A well-known example is NETtalk (Sejnowski & Rosenberg, 1987), which learned, with supervised training, the mapping from English spelling to pronunciation. Feed-forward networks come in many varieties, for example, with units having continuous-valued or binary activation levels, and with and without hidden units. In whatever guise, the function of a feed-forward network is to compute a mapping or transformation from one domain into another.

Recurrent (memory) networks. Some networks are not strictly feed-forward in structure. Networks that involve feedback of activation, whether with interconnections among units in a single layer or with connections back from a later layer to an earlier one, are called *recurrent networks* (see Figure 2). Recurrent connections allow units to influence and constrain each other in finding the best overall pattern that fits the input. For example, two units may have reciprocal excitatory connections and tend to turn each other on even if an input pattern directly activates only one of the two. In a recurrent network, over time the flows of activation settle on a state that (locally) optimizes the fit of the activation pattern to the various constraints represented by the between-unit connections as well as the current inputs. Such a state is called an *attractor* of the network (see Churchland & Sejnowski, Chapter 3). A single network may have many attractor states that it will reach given different starting points (patterns of activation).

If the attractor states represent learned patterns, a recurrent network can function as a content-addressable memory. The attractor dynamics are another way to view the pattern completion property described earlier. Presentation of a new pattern that is similar to a learned one—or a learned pattern in incomplete form or with random error added—will start the network out in the neighborhood of the attractor corresponding to that pattern, so over time the network will settle into the proper state: a representation of the learned pattern, with the noise or error "cleaned up." An example of a recurrent network was presented and fit to psychological data by McClelland and Rumelhart (1986). Thus, recurrent networks can function as content-addressable memories or as pattern cleanup devices that remove error and restore the canonical form of a pattern. Other types of recurrent networks can perform other tasks, such as the recognition or generation of sequences of patterns over time (a property that has applications in language processing; Jordan, 1989).

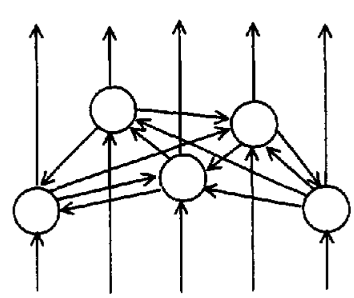

Figure 2. Example of a recurrent network. The units are densely interconnected so that they influence each other (as well as being affected by the input pattern). Such a network can operate over time to converge to an "attractor" state that depends on the weights on the within-network connection (and hence on past learning) as well as on the input.

Multiple modules. A single network module, whether a pattern transformer or a content-addressable memory, can be only a component of a complete cognitive system. Proposed connectionist models for significant, realistic tasks such as sentence comprehension or question answering therefore generally involve multiple interconnected modules. An example is Miikkulainen's (1993) DISCERN model, which learns to process simple stories based on scripts such as a restaurant visit or an airplane flight. DISCERN uses two separate memory modules for storing its lexicon of word meanings and the representations of stories that it has processed, and several feed-forward processing modules for parsing input sentences, answering questions, and the like.

As another example, Rueckl (1990) presented a distributed connectionist model of *repetition priming*, the effect of prior presentations of words and nonword strings on people's ability to later recognize them in very brief visual presentations (a form of implicit memory). His model includes a module for visual features connected to a recurrent net that represents orthographic patterns. Two additional modules, also linked to the orthographic module, represent semantic and phonological information respectively. The presentation of visual features at the input elicits a pattern of activity in the visual module that represents the visual appearance of the stimulus. The flow of activation then follows the mappings from visual to orthographic patterns and thence to patterns of semantic features and phonological features in the latter two modules. Feedback paths between modules and within the recurrent orthographic module influence the way the system converges or "relaxes" into a final state reflecting not only the nature of the input but also learned visual, orthographic, and phonological constraints. Rueckl found that this model accounted qualitatively for his data patterns. For example, the effects of repeated priming presentations on later identification performance differs for words versus nonwords; the reason is the lack of any encoding for nonwords (which, of course, have no meanings) in the semantic module. In a recent study, Rueckl and Olds (1993) found that attaching arbitrary "meanings" to nonwords changed the way they were affected by priming, making them act identically to words!

As a final example, McClelland, McNaughton, and O'Reilly (1995) presented a model aimed at explaining how humans can both quickly form memory traces of unique events and also integrate many experiences over time so that expectancies can be based on the general long-term statistical structure of the world rather than on a highly variable sample of recent events. In a connectionist framework, these two types of memory pose competing demands: for rapid changes in connection weights so that the details of a new experience can be preserved, and for

slow, incremental changes so that a network's weights summarize a large amount of experience rather than being haphazardly pushed one way or the other by each new event. These two forms of memory correspond roughly to episodic and semantic memory, and McClelland and his colleagues assumed (on the basis of neuropsychological evidence) that they are anatomically mediated by the hippocampus and the neocortex. Their model has two modules. The one analogous to the hippocampus uses a learning algorithm that rapidly acquires new memories. In a process akin to *consolidation,* a newly formed memory is later transferred by repeated presentations to the other module, whose weights change only gradually and incrementally with experience. Consolidation is known, on independent grounds, to take considerable time in humans—up to years—and the authors suggested that it is necessarily slow so that new knowledge can be integrated nondisruptively into the stably structured representations maintained in the neocortical system.

In summary, single networks can be developed for specific isolated tasks (e.g., NETtalk; Sejnowski & Rosenberg, 1987). However, models of higher level tasks such as story comprehension or repetition priming, or models that are intended to capture the relationships among different types of performance (such as rapid vs. slow memory functions for McClelland et al., 1995), generally use several interconnected modules.

Distinctions Between Associative Networks and Connectionist Models

In the 1970s, theorists developed the simple stimulus–response (S–R) links postulated by behaviorist models into more sophisticated associative theories of memory (e.g., J. R. Anderson & Bower, 1973), which in turn were the direct forerunners of the earliest associative models in social cognition (e.g., Hamilton et al., 1980; Hastie & Kumar, 1979). There is some potential for confusion between these familiar types of associative networks with spreading activation and the distributed connectionist models that are the focus of this article. After all, both involve nodes connected by links. However, the differences are fundamental (see J. A. Anderson, 1995; Barnden, 1995a; McClelland et al., 1995, pp. 428–429; Thorpe, 1995; Touretzky, 1995).

1. Associative networks are models of representational structure only, not process; additional processes must be postulated to construct and retrieve information (e.g., the productions of J. R. Anderson, 1983). In contrast, connectionist networks constitute the processor as well as the knowledge representation; flows of activation are the only processing mechanism.

2. Associative models use localist representations: A node represents a concept or proposition. In contrast, most connectionist models use distributed representations in which a single node has no meaningful semantic interpretation; only a pattern of activation has any meaning.

3. Associative networks are assumed to be rapidly constructed and dynamically modified by interpretive processes, for example, during the second or two it takes to comprehend a sentence. In contrast, connectionist networks are generally assumed to have a fixed topology, with the connection weights changing only slowly as learning occurs.

4. In associative networks, activation is usually assumed to spread both ways over links (Node A can spread activation to a connected Node B, or B can activate A). In contrast, in a connectionist network excitatory or inhibitory activation is ordinarily assumed to spread only one way (links are directional).

A summary of all of these points is that modern associative models have been developed by theorists concerned with understanding the representation and processing of linguistically encoded information, and they serve those purposes well (Barnden, 1995b). In contrast, connectionist networks have been constructed for a much greater variety of functions, ranging from early perceptual processing on the input side to motor control on the output. They are also generally developed with a greater concern for neural plausibility, though they still involve simplifications and idealizations rather than detailed matches to the properties of actual biological neurons. Perhaps most important, they use distributed representations and therefore acquire the properties (such as explicit similarity and pattern completion) outlined earlier.

Localist constraint-satisfaction networks, which have recently been applied in social psychology (e.g., Miller & Read, 1991) and are sometimes labeled *connectionist,* actually share more important properties with associative models than with distributed connectionist models (see Barnden, 1995b). These networks involve discrete nodes that represent semantically meaningful features or propositions, connected by positive or negative links that encode the covariational or inferential links among nodes. For example, a node representing "John loves Mary" and one meaning "Mary loves John" would presumably be connected with a positive link, for it is likely that both of these propositions are true if either one is. Such networks are well suited to representing certain types of information, and spreading-activation processes can be used to model the simultaneous satisfaction of a number of constraints represented by the links. However, they should not be confused with distributed connectionist models. Like other types of associative networks, localist constraint-satisfaction networks model structure only (not process). They use localist representations that are assumed to be rapidly constructed by interpretive processes. They spread activation both ways over links constructed to reflect positive or negative implicational relations among concepts; the links do not arise from a learning process. Therefore, their properties make them much more akin to associative networks than to connectionist models that use distributed representations (Thorpe, 1995; Touretzky, 1995).

Implications for Social Psychology

The reader may be persuaded by the above material that connectionist models are interesting and deserve the attention of psychologists who study perception or memory, yet still wonder whether these models have any direct implications for social psychological theory and research. In this section I present four types of argument for the idea that there are such implications.

General Intuitive Arguments

First, there are some general reasons to favor connectionist models over symbolic ones. These points were elaborated by Churchland and Sejnowski (1989) and in several chapters in Rumelhart, McClelland, et al. (1986).

1. Inspection of the brain reveals a massive number of richly

interconnected, very simple processors (neurons). The brain appears a much more promising candidate as hardware for connectionist networks than as hardware for the familiar sort of symbolic computer.

2. Artificial intelligence researchers, despite much focused effort, have generally failed in efforts to get symbolic models to do things that are trivially easy for humans to do, such as recognizing our friends' faces or walking around without bumping into things. On the other hand, operations that are difficult for humans (such as complex mathematical calculations) are trivial for current symbolic systems. These points suggest that the human mind–brain and symbol-processing systems may have distinct architectures that are well suited for different types of problems.

3. Even if it were plausible that adult humans are fundamentally processors of languagelike symbols, this is not a plausible description of, say, dogs or human infants. Nonverbal or preverbal creatures presumably lack the syntactic and semantic (conceptual) powers to reason symbolically, yet they generally behave quite effectively. The idea that their cognition is of a fundamentally different sort from that of adult humans makes no sense in evolutionary or developmental terms.

Of course, none of these simple arguments can be fully convincing in itself, yet they exemplify the sorts of considerations that have made connectionist models attractive to many researchers.

Integration With Cognitive Psychology

A second point is that many cognitive psychologists have found connectionist models to be useful—not only researchers studying basic processes of vision or memory but also those interested in such higher level phenomena as categorization, decision making, and relations between judgment and memory (e.g., Kruschke, 1992; Weber, Goldstein, & Busemeyer, 1991). These areas of research have close connections with important social psychological processes such as stereotyping and social judgment, so it seems reasonable to predict that social psychologists interested in similar issues should also find these models useful.

The present time may parallel the late 1970s, when some social psychologists came to realize that the newly flourishing field of cognitive psychology had developed powerful models of mental representation and process (such as associative network/spreading-activation and schema theories) that appeared to have potential implications for social psychology. Borrowing from such models, initially in such pioneering studies as those of Hamilton et al. (1980) and Hastie and Kumar (1979), sparked the development of the entire field of social cognition. Today the time may be ripe for similar borrowing of connectionist models, which are now being widely applied within cognitive psychology. Such borrowing should be valuable because the adoption of a common theoretical language will facilitate productive theoretical interchange and integration between social and cognitive psychology. The history of social cognition since the 1970s shows the benefits of such interchange (Devine, Ostrom, & Hamilton, 1994).

Connectionist Accounts for Known Phenomena

Beyond the general virtues of conceptual integration, there are a number of specific ways in which connectionist models can at least give us new metaphors and change the way we think about social psychological phenomena—and perhaps also set us looking in new theoretical and empirical directions. First I describe ways in which connectionist models can account for phenomena that are also predicted by existing social psychological theories.

Explicit memory: Recall and recognition. Humphreys and his associates (Chappell & Humphreys, 1994; Humphreys, Bain, & Pike, 1989; Wiles & Humphreys, 1993) have extensively investigated the ability of a class of multiple-module connectionist models, schematically portrayed in Figure 3, to fit psychological findings regarding explicit memory retrieval. Recurrent network modules learn to store semantic representations of knowledge by means of the pattern-completion property. Other networks perform input and output mappings (e.g., translating from visual features of letters into the central representation of a word's meaning). One version of this model (Chappell & Humphreys, 1994) fits many detailed data patterns from studies of recognition and cued recall. In this model, explicit memory depends on re-activation of representations in the central recurrent network memory, whereas some types of implicit memory (such as repetition priming) are due to weight changes in the input–output pattern associator networks, as I discuss shortly.

These distributed models assume that a stimulus item is represented in memory as a pattern or vector of features. In fact, McClelland and Chappell (1995) observed that this approach is becoming increasingly common in nonsocial memory models. The older models that have best survived detailed tests against psychological data (e.g., J. R. Anderson, 1983, or the SAM [search of associative memory] model of Gillund & Shiffrin, 1984, and its relatives) assumed an associative framework: An item in memory was conceptualized as a node with associative links (created by study) to nodes representing other items and the context. In contrast, the newer models (including those of Chappell & Humphreys, 1994; McClelland & Chappell, 1995; and McClelland & Rumelhart, 1986) consider a memory item as a pattern of features, a conceptualization that naturally maps onto a distributed representation as a pattern of activation

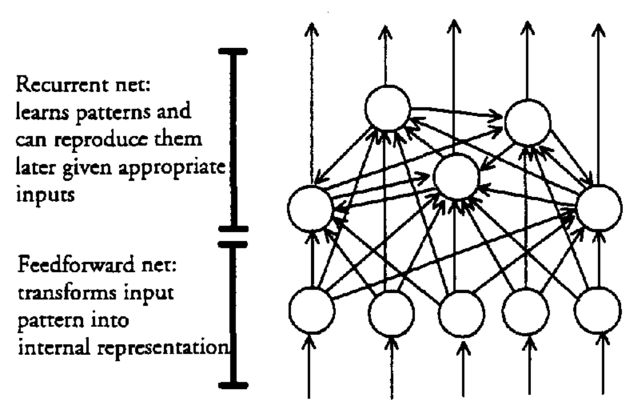

Figure 3. Example of a multiple-module system, with an input feedforward network transforming patterns of visual features to a distributed semantic representation, which is then stored in a central recurrent memory network.

across a set of units. The rise of this new approach has been driven by the recognition that the older associative models have great difficulty in accommodating certain newer findings regarding memory, such as the null list-strength effect (see McClelland & Chappell, 1995).

Associative retrieval. Despite representing stimuli as patterns of features rather than as discrete nodes connected by associative links, connectionist models of memory can account for the observation that when items of information are encountered or considered together, one item can later facilitate recall of the others (J. A. Anderson, 1995; Chappell & Humphreys, 1994; McClelland et al., 1995; Wiles & Humphreys, 1993). In general, the models combine the distributed patterns representing the several items into a single, larger pattern. The learning rule then changes the weights on connections among units in a way that subserves the pattern-completion property. At a later time, re-presentation of one or more of the "associated" stimulus items (a partial pattern) can lead to the completion of the pattern, as the remaining items (other parts of the pattern) are "retrieved." One version of this general idea, accounting for the associative binding together of different aspects of everyday experiences (such as verbal names and visual, auditory, and tactile attributes of an object; see Damasio, 1989, or Carlston, 1994) was described by Moll, Miikkulainen, and Abbey (1994). Once a pattern that binds together the different subpatterns is created and learned through connection weight changes, presentation of one aspect will lead to the reactivation of all. Thus the sight of a friend may reactivate her name, memories of her characteristic speech patterns, feelings about her, and other associated representations (Carlston, 1994). Moll et al. demonstrated that the network architecture they proposed has sufficient capacity (given the number of neurons estimated to exist in the appropriate brain regions) to store approximately a hundred million distinct associative memories, enough for several every minute over a human lifetime.

Semantic priming. Responding to a given stimulus facilitates a subsequent response to a semantically related stimulus; thus, people are faster to reply that *nurse* is a word after reading *doctor* than after reading an unrelated word (Meyer & Schvaneveldt, 1971). In a distributed memory model, this type of priming would not be explained as the result of activation spreading over links between nodes representing concepts, for concepts are not represented by single nodes. Instead, it could be explained in terms of pattern overlap (Masson, 1991). Distributed representations created by typical learning rules have the useful property that related items have similar representations (Churchland & Sejnowski, 1992; van Gelder, 1991). The pattern of activation representing *nurse* is more similar to that for *doctor* than it is to *tree* (similarity, loosely, is the correlation of the vectors of activation values). Thus, if a set of units is already in the *doctor* pattern, it takes less time and less definitive input information to change to the *nurse* pattern than it would from an unrelated starting pattern. This account of semantic priming, in contrast to the traditional spreading-activation account, predicts that priming should be abolished by a single unrelated intervening item. This prediction has been tested and confirmed (Masson, 1991).

Schemas. As noted earlier, a recurrent network can function as a content-addressable memory. If it learns many stimuli (patterns of activation), then at a later time input that is similar to one of the known patterns (e.g., a subset of a pattern, or a pattern with random noise added) can elicit the entire pattern. If a number of the learned patterns are related—say, they derive from a category prototype with random variations—their prototype will be learned. Then presentation of a new category member will elicit the prototype as a response (McClelland & Rumelhart, 1986; Rumelhart, 1992; Smolensky, 1989). For example, if the network has learned about a number of politicians who vary in many ways but are usually "verbal" and "ingratiating," its output will indicate that a newly encountered politician probably has those traits. As another example, if numerous presentations have led to the storage of a pattern representing a particular individual, then a new person with some physical resemblance to the known person may be inferred to share other characteristics as well. All of these performances are different expressions of the prototype extraction and pattern completion properties discussed above.

These performances correspond to schema-based or exemplar-based processing (e.g., Smith & Zárate, 1992; Wyer & Srull, 1989). In traditional social psychological theories, the perceiver is said to search memory (conceptualized as a file drawer or storage bin) to locate the discrete schema or pattern that best fits the input information. This idea requires various assumptions, such as: How is the search performed (top down in a storage bin)? What are the criteria for stopping the search? When will a schema versus a well-known exemplar be used as the basis for inference? How are schemas formed in the first place?

In the connectionist model none of these questions arise. There is no search, and the learning process is explicitly modeled. The most important difference is that in a superposed connectionist model a schema is not a "thing" written on a sheet of paper in a file cabinet but rather is one among many potential patterns that is encoded in a set of connection weights and can be elicited as a pattern of activation—given the appropriate input cues. As Rumelhart, Smolensky, et al. (1986) explained, in this model

> there is no representational object which is a schema. Rather, schemata emerge at the moment they are needed from the interaction of large numbers of much simpler elements all working in concert with one another. Schemata are not explicit entities, but rather are implicit in our knowledge and are created by the very environment that they are trying to interpret. (p. 20)

Another way of saying this is that because patterns representing knowledge are not stored as discrete entities, there are not separate steps of "searching" for relevant prior knowledge that is then "used." Rather, all knowledge is implicit in the connection weights that gate the flow of activation between units, so all knowledge necessarily influences the course of processing.

Elicitation of norms and effects on social judgments. Kahneman and Miller (1986) outlined *norm theory* as an account of the way perceivers interpret and evaluate objects and events against a background of relevant alternatives. Their key claim is that an experience elicits the retrieval of representations of similar past experiences or of counterfactual alternatives (constructed on the basis of general knowledge). The elicited

norm then serves as a context that influences judgments of the experience's typicality or unusualness, comparative judgments of various kinds, and affective reactions to the experience. Kahneman and Miller (p. 136) emphasized that "each event brings its own frame of reference into being" by eliciting relevant events from memory, in contrast to the typical view that events are judged with reference to static, precomputed expectancies. As just outlined in the discussion of schemas as on-the-spot constructions, connectionist models offer a straightforward implementation of these ideas. A stimulus pattern that is input to a network can yield outputs that reflect the aggregation of similar past experiences as well as relevant generic knowledge. In this way, as Kahneman and Miller posited, the details of the event itself (serving as cues) activate the subset of stored knowledge that in turn serves as a context and background for judgments and responses to the event.

Repetition priming. *Repetition priming* is the facilitation of processing of a stimulus when the same stimulus has been processed in the same way on a previous occasion. In sharp contrast to semantic priming, this is a long-lasting phenomenon (up to months; Sloman, Hayman, Ohta, Law, & Tulving, 1988). It generally does not require explicit memory for the initial experience (Schacter, 1987; Smith, Stewart, & Buttram, 1992). Wiles and Humphreys (1993, pp. 157–163) investigated the possible mediation of such effects in distributed memory models and concluded that weight changes in networks that translate information from one representation to another (e.g., from letters to word meanings) are responsible. Learning changes weights incrementally after each individual pattern is processed. Because the changes are in the direction of more accurately and efficiently processing the given pattern, the pattern will have an advantage over a novel pattern for a period of time after a single presentation. Similar suggestions have been made by Humphreys et al. (1989), Rueckl (1990), Schacter (1994), and Moscovitch (1994). This theoretical idea locates repetition priming in input–output pattern transformation networks, separate from those that subserve explicit memory (central recurrent or pattern-completion networks; see Figure 3). It thereby explains why these two forms of memory are often found to be independent of one another, both in normal humans and in those with various forms of amnesia (see Schacter & Tulving, 1994). Repetition priming is not part of many mainstream symbolic theories in social psychology (e.g., Wyer & Srull, 1989) but has been empirically observed with social judgments (Smith et al., 1992) and can be explained by the exemplar model of Smith and Zárate (1992).

Flexibility and context sensitivity. Two important dynamic aspects of distributed representations are their flexibility and context sensitivity. In connectionist models, representations that are not currently active are not stored away inertly until accessed by a retrieval process. Instead, flows of activation through connection weights that are shaped by learning reconstruct a representation as a distributed pattern of activation. In this process, any other current sources of activation (e.g., patterns representing the person's mood, perceptually present objects, current concerns, or goals) will also influence the resulting representation. For instance, thinking of an "extravert" in the context of a noisy party might activate a representation that includes "telling jokes" and "being the center of attention," whereas in the context of a used-car lot the resulting representation might include features such as "pushy" and "impossible to discourage." Such thoroughgoing context sensitivity is an inherent property of distributed representations. Clark (1993, especially Chapters 2 and 5) summarized the implications of this fact:

> The upshot is that there need be no context-independent, core representation for [a concept]. Instead, there could be a variety of states linked merely by a relation of family resemblance. . . . A single . . . [concept] will have a panoply of so-called subconceptual realizations [i.e., activation patterns], and which realization is actually present will make a difference to future processing. This feature (multiple, context-sensitive subconceptual realizations) makes for the vaunted fluidity of connectionist systems and introduces one sense in which such systems merely approximate their more classical cousins. (pp. 24–25)

Current thinking in many areas of social psychology (including the self, attitudes, and stereotypes) emphasizes the flexibility and context sensitivity of mental representations. For example, Markus and Wurf (1987) advanced the notion of a "working self-concept," the contextually relevant set of self-attributes that are currently active. According to Wilson and Hodges (1992), attitudes too are constructed on the spot in a flexible and context-dependent manner rather than being retrieved from memory in invariant form every time they are accessed. It seems likely that all types of cognitive representations will be found to be flexibly reconstructed in a context-sensitive way rather than retrieved from memory as they were stored—like items buried in a time capsule—as assumed by many current symbolic theories in social psychology.

Irreversibility. In a distributed connectionist network, the learning process incrementally changes the connection weights after each stimulus is processed. The changes are likely to be subtle and global, and in particular there is no computationally feasible way to undo the changes that result from processing a stimulus. For example, processing an "opposite" will not restore the network to its previous state. Thus, if someone learns that George is a wimp, later learning that George is not a wimp will not return the network to its initial state. The difficulty of "unbelieving" information that was once believed is a theme in recent work by Gilbert (1991). In contrast to the natural prediction of irreversibility from a connectionist model, in a symbolic model it seems strange that one cannot just "erase" a symbol string from memory or "attach a negation tag" to it to effectively unbelieve the original information. A symbolic model would predict irreversibility only if many inferences generated from the new belief had been thoroughly integrated into the structure of existing knowledge, a process that should take both time and thought.

Formation of evaluative impressions from traits. One of the few social psychological applications that has appeared to date is Kashima and Kerekes's (1994) distributed connectionist model of the formation of an evaluative person impression. Pairings between activation patterns representing the target person and the given trait information are computed and stored in a set of connection weights. As additional traits are presented, their representations are added in to the weights. Traits with

similar meanings are assumed to be represented with similar patterns (according to the explicit-similarity property) so that the final superposed impression can be compared with global "good" and "bad" patterns to derive an overall evaluative judgment. In many cases, such judgments are well approximated by a weighted average of the evaluations of the items of input information (N. H. Anderson, 1981), and Kashima and Kerekes's model reproduces that pattern as well as various details of order effects when the information is presented serially.

New and Distinctive Predictions From Connectionist Models

As we have seen, connectionist models offer explanations for several phenomena concerning memory, priming effects, and inference that are also predicted by existing social psychological models. Even if this were all that connectionist models could offer, a common framework that could integrate all these phenomena (as well as many findings regarding nonsocial cognition) would be a conceptual advance. Still, one test of a theoretical framework is its ability to generate new predictions not shared by existing models. Here are some examples of derivations from connectionist models that are intriguing, generally have not been tested as yet, and might conceivably even be true.

Retrieving one representation versus using many constraints. In a connectionist network, because activation flows depend on all the network weights, the output produced by a set of input cues draws on all the network's stored knowledge as sources of constraint rather than reflecting only a single stored pattern. As a demonstration of this point, Rumelhart, Smolensky, et al. (1986) trained a network with the typical features of various types of rooms (living room, bedroom, etc.). Presented with cues that clearly related to only one of the known room types (such as a bed) the network reactivated the entire known bedroom pattern. More important, when cues that typically related to different rooms were presented (e.g., bed and sofa), the network did not decide arbitrarily between bedroom and living room, nor did it break down with an error message about incompatible inputs. Instead, it combined compatible elements of the two relevant knowledge structures to produce a concept of a large, fancy bedroom (complete with floor lamp and fireplace).

There is evidence that people generally combine multiple knowledge structures as well (Carlston & Smith, in press). For example, retrieval of a memory may be influenced by general knowledge as well as by traces laid down on a specific occasion (Loftus, 1979; Ross, 1989), or perceptions and reactions to a person who is a member of multiple categories, such as a Pakistani engineer, may be influenced by knowledge relating to all of the categories. Traditional theories, however, have been built around the assumption that the single best-fitting category, schema, or stereotype is searched for and then used as a basis for inference and judgment. Research on whether and how people draw on multiple knowledge representations to process complex stimuli, such as novel combinations of stereotypes, which might allow tests of these predictions, seems to be virtually nonexistent as yet within social psychology.

Accessibility. In a connectionist network, the recency and frequency with which a pattern has been encountered during the learning process influence the ease with which it can be elicited by a given set of cues. This is because learning involves incremental weight changes. When a stimulus is processed, learning changes weights in a way that makes the current pattern and similar ones slightly easier to reproduce in the future, at the expense of slightly distorting (and worsening performance on) unrelated patterns. In other words, the principle of accessibility is inherent in the network's operation. In traditional theories within social psychology, accessibility is not intrinsic to basic theoretical processes but is explained by special ad hoc mechanisms, such as a storage battery containing time-varying amounts of charge attached to each discrete representation (Higgins, 1989), or a top-down search of a storage bin that holds multiple copies of each representation (Wyer & Srull, 1989).

Connectionist models make the novel prediction that recent and frequent exposure produce two distinct types of accessibility with different properties (Wiles & Humphreys, 1993, p. 159), respectively dependent on current unit activations and on changes in connection weights. First, a pattern of activation across a set of units may persist for a short time after a stimulus is processed, so that if the next pattern is related to the first its processing may be facilitated (Masson, 1991). This type of accessibility may underlie semantic priming, the observation that having just read the word *bread* makes it easier for people to read *butter.* The activation patterns representing *bread* and *butter* will overlap to a greater extent than do representations of unrelated words; this is a property of the representations produced by typical connectionist learning rules (Clark, 1993). The connectionist account predicts that this sort of priming should last only briefly and should be abolished by one or two intervening unrelated words (which would create unrelated patterns of activation).

Second, processing a stimulus leads to incremental changes in the connection weights in a network. This change is long lasting, and its effects diminish not with time but with interference from unrelated patterns. Many people have an intuition that the effects of weight changes caused by processing a stimulus on a single occasion could not be demonstrable over days or even weeks, though priming effects clearly can last that long (e.g., Smith et al., 1992). However, Wiles and Humphreys (1993, pp. 159–162) argued in quantitative detail that this intuition is misleading. If a particular stimulus is processed frequently over months and years, the resulting systematic shifts in connection weights will influence the individual's processing characteristics for years—even a lifetime (a property termed *chronic accessibility* in the social literature).

Though the mechanisms are different, under some circumstances these two forms of priming may have similar effects, such as increasing the probability that people will assimilate an ambiguous stimulus to the primed category. Bargh, Bond, Lombardi, and Tota (1986) argued that the two forms depend on the same underlying mechanism, on the basis of a finding that these two sources of accessibility have additive effects. However, this conclusion can be questioned on logical grounds (see Carlston & Smith, in press, pp. xxx). Cognitive psychologists often interpret such additivity as indicative of distinct and separable processes rather than as evidence for process equivalence (Sternberg, 1969). Moreover, other evidence suggests that

the two types of accessibility can have somewhat different properties (Bargh, Lombardi, & Higgins, 1988; Higgins, Bargh, & Lombardi, 1985; Smith & Branscombe, 1987). Further, more focused empirical tests of possible differences between two forms of accessibility, hypothesized by this type of connectionist account but not by existing models, would be of value.

Evaluative priming effects. *Evaluative priming* (Fazio, Sanbonmatsu, Powell, & Kardes, 1986) is the effect of reading an evaluatively laden prime word (such as *cockroach*) in facilitating processing of an evaluatively congruent target (*crash*) while inhibiting the processing of an incongruent word (*beautiful*). The extent of these effects is controversial; some evidence suggests they occur for virtually all evaluatively non-neutral prime words, and other evidence suggests that they are limited to words for which the person holds a relatively strong attitude (Bargh, Chaiken, Raymond, & Hymes, in press; Fazio, 1993). Still, there is agreement on the robustness of the evaluative priming effect itself. Evaluative priming has usually been attributed to the spread of activation along associative links connecting concepts in memory. However, it is implausible that all positive and all negative concepts are interconnected with strong links as are semantically related words (e.g., *doctor–nurse*). Even if people have often thought about *cockroach* together with some negative concepts such as *disease* and *filth,* many other negative concepts, such as *crash*, seem entirely unrelated. An alternative explanation is that the patterns representing positive concepts, and also those representing negative concepts, overlap to a nontrivial extent. As noted earlier in discussion of the explicit-similarity principle, learning rules in connectionist networks create distributed representations that match conceptual similarity with pattern similarity. In this case, evaluative priming can be explained as resulting from pattern overlap in exactly the same way as semantic priming (as described earlier). This account makes the novel prediction, which apparently has not yet been tested with evaluative priming, that an intervening neutral word or two would abolish the effect. Spreading-activation accounts do not make this prediction, for activation in some remote part of the network (caused by presentation of an unrelated word) should not wipe out activation spreading from one positive word to another one.

Bidirectional causation. Most explicit theories within social psychology postulate unidirectional causal models. For example, Fishbein and Ajzen (1975) held that beliefs cause attitudes, which in turn cause behavioral intentions and behavior. In reality, research has shown that such influences are rarely unidirectional; attitudes can cause beliefs by processes of rationalization, and behavior can cause attitudes through dissonance reduction or self-perception. In fact, Smith (1982) proposed as a general principle that if Cognition A affects B, then manipulating B will also be found to affect A. Our theories rarely reflect such complexities, though theorists often pay lip service to the possibility of feedback and bidirectional causation. In a connectionist framework, mutual causation among a number of cognitions (beliefs, attitudes, goals) is an inevitable consequence of the process of mutual adjustment and constraint satisfaction in recurrent networks. The attractor pattern to which a network converges reflects all the constraints encoded in connections among units and so will be influenced by all existing representations. Once the system has reached such a state, changing any belief, attitude, or goal may change all others as the system adjusts to the perturbation and perhaps switches to a different attractor state—so any element can influence all the others. This appears to be a more promising general description of the human mind than theories postulating unidirectional causation. Further empirical exploration of patterns of mutual causation among beliefs, attitudes, moods, goals, and other cognitive elements might permit the discrimination of this type of connectionist account from traditional unidirectional theories.

Integration of motivation and cognition. Not only input information and learned expectations but also a person's goals and motives (such as self-esteem enhancement or mood regulation) are among the constraints that affect processing and influence the particular attractor into which the network settles. For example, a transient system state that fits current input well but has negative implications for self-esteem may change—as part of the simultaneous constraint satisfaction process—to one that fits the input slightly worse but has a much more positive implication for the self. This dynamic conceptualization offers the potential for an integration of cognition and motivation in a single theoretical framework (see Dorman & Gaudiano, 1995). Social psychologists have investigated many types of cognition–motivation interactions: For example, the tendency of an activated goal to increase the accessibility of goal-related concepts (e.g., Higgins & King, 1981), the ability of an external stimulus or situation to activate a motive to which it is relevant (e.g., Bargh, 1994), or the effects of motives on judgment and memory that persist despite efforts to be accurate (e.g., Sanitioso, Kunda, & Fong, 1990). In fact, social psychologists' theoretical conceptualizations and empirical findings on such interactions may be among our most important contributions to the development of integrative models of motivation and cognition. Such models might be built around the idea of motives as well as cognitive representations as encoded in distributed networks, subject to the law of accessibility and functioning simultaneously as constraints that influence the system's convergence into an attractor state that becomes the basis for further processing.

Separate memory systems for expected and novel information. As reviewed earlier, McClelland et al. (1995) proposed a connectionist model that includes one module analogous to the hippocampal system, which rapidly learns new information, and another analogous to the neocortex, which shows only slow weight changes and stores stably structured general knowledge. New knowledge is transferred from the former system into the latter in a process analogous to consolidation, which takes a long time (up to years) in humans and other animals.

This proposal of independent fast- and slow-learning systems has important though as yet untested implications for us in social psychology. One implication is that people have separate mechanisms that perform the functions that our theories currently attribute to associative networks and schemas. The schematic function of interpreting input information in terms of stable, general world knowledge (e.g., Markus & Zajonc, 1985, p. 145) may be performed by neocortical systems that learn slowly, extracting regularities in the environment and using them in the course of processing further inputs. Though the

learning is slow, specific recent experiences as well as frequently repeated ones may leave traces that have observable effects such as repetition priming or implicit memory (Schacter, 1994). In contrast, the rapid construction of new associative structures that bind together information about different aspects of an object or experience in its context (Wiles & Humphreys, 1993) seems to take place in hippocampal systems that exhibit one-shot learning and mediate conscious, explicit recollection. One implication is that people's verbal reports about what they know may well rest on different representations than those they use in preconscious interpretation, so as a methodological principle it may not be wise to rely on verbal reports to assess the contents of people's schemas.

In addition to these differences in learning speed and conscious accessibility, these two systems are predicted to differ in the type of information to which they attend. Schematic learning is chiefly concerned with regularities, so it records primarily what is typical and expected. In contrast, episodic memories should record the details of events that are novel and interesting; in other words, this system should attend more to the unexpected and unpredicted. Social psychological studies have shown that people attend to and recall mostly expectancy-inconsistent information when forming a new impression but recall mostly expectation-consistent information when working with a well-formed and solid expectation (Higgins & Bargh, 1987). This empirical finding may reflect the more basic differences between two underlying memory systems: one that learns quickly and emphasizes novelty and one that accumulates information slowly and emphasizes regularities.

The independence of these two systems implies that the explicit episodic memories that are our conscious link to our autobiographical past are not the only residue that the past has left in us. Implicit learning in nonconscious systems also affects the way we see and interpret the world. This view may offer theoretical leverage for interpreting many seemingly puzzling observations. For example, "intuitive" emotional reactions such as a fear of flying are often stubbornly independent of our conscious knowledge (Kirkpatrick & Epstein, 1992), and associations of social groups with stereotypic traits may endure even in people who consciously and sincerely reject those stereotypes (Devine, 1989). A commonsense assumption is embodied in many social psychological theories: that all our knowledge and beliefs are represented in a single memory system—so that, for example, the beliefs we can consciously access and verbally report are the same ones that guide our preconscious interpretation of our experiences and reconstruction of our explicit memories. This assumption now seems highly questionable (McClelland et al., 1995; Schacter & Tulving, 1994). Tests of the assumption within social psychology may be spurred by the derivation of distinctive predictions regarding separate memory systems in a connectionist framework.

Summary. Connectionist models have been developed to account for many observations that are familiar in social psychology (e.g., explicit recall, schematic interpretation of input information) and can also make novel predictions. The same has been true of connectionist models in nonsocial cognition. For instance, as Nosofsky's (1986) model of exemplar-based categorization was translated into a connectionist model by Kruschke (1992), several novel predictions emerged, which were eventually supported (at the expense of the original version of the model) by empirical test. Another example is provided by Hinton and Shallice's (1991) work, which showed that "damage" to a connectionist network can simulate various properties of aphasia and other neurological disorders, including counterintuitive properties that appear quite difficult to account for with models of other types. Thus, connectionist models often generate new predictions about aspects of observable behavior that derive less naturally from theories of other types. Other examples will no doubt emerge as connectionist models are developed in more detail. Of course, given the newness of the connectionist framework, particularly in social psychology, most of these predictions have not yet been tested, but their mere existence suggests that connectionist models will spur new empirical insights as well as offer integrative accounts for existing findings.

Critiques and Open Questions

As connectionist models have been developed for various phenomena of nonsocial cognition, several important critiques and questions have been raised in the literature. In this section I sketch some of the issues involved and attempt to anticipate some questions that the reader may have.

Explanations for Language Use and Conscious, Explicit Processing

Among the connectionist models developed in the mid-1980s were several efforts to model linguistic phenomena (e.g., Rumelhart & McClelland, 1986). These models were the subject of vigorous critiques, particularly by Fodor and Pylyshyn (1988) and Pinker and Prince (1988). The critics argued that language has special properties, which in principle cannot be reproduced in a connectionist framework (unless such a framework is considered simply as the underlying hardware that implements a classical symbol-manipulating system). One such property is *systematicity,* the idea that if an organism can use a particular concept in one context (e.g., the concept of "green" in "green grass") it must be able to use it in any relevant context (e.g., "green eggs and ham"). This and other linguistic and logical properties were argued to be fundamentally incompatible with connectionist networks that learn from experience and hence, encountering green grass but never green eggs and ham, might be able to represent the former but not the latter concept.

Numerous responses to these critiques have been made by connectionist theorists over the years. (See Barnden, 1995a; Clark, 1993; Elman, 1995; Shastri, 1995.) Responses have acknowledged that the earlier models of linguistic phenomena were in many cases naive and unrealistic and that language does have special properties that theoretical models must respect. However, many have argued that the conclusion that connectionist models are in principle incapable of showing these properties is at best premature. A variety of active research programs, which cannot be summarized here, are taking various approaches to modeling language (e.g., Henderson, 1994; Plunkett & Marchman, 1991). A full connectionist account of systematicity and other linguistic properties is clearly not yet at

hand, but it is equally clear that much progress has been made beyond the inadequate early models.

Even as the question of the possibility of adequate connectionist accounts of language (and linguistically encoded thought) remains open, it is important to consider the place of language in an overall psychological model. In recent years, social psychologists have advanced many related dual-process models emphasizing the distinction between controlled (conscious, systematic) and automatic (nonconscious, heuristic) processing (see Smith, 1994, for a review). These ideas have strong parallels with Smolensky's (1988) approach to connectionism, which also holds that people have two separate processors. The "top-level conscious processor" uses linguistically encoded and culturally derived knowledge as its "program;" this is the processor that people use when they follow explicit step-by-step instructions or engage in conscious, effortful reasoning. It is based on the same cognitive capacities that underlie public language use, such as the ability to parse sentences into their components and to combine words to form sentences following grammatical rules. This system can recombine known linguistic symbols into new patterns, giving rise to the property of systematicity, and can quickly formulate and store symbolic expressions representing newly learned knowledge. (Ultimately, of course, all these capacities must rest on computations carried out by connectionist networks, the only type of hardware available in the brain. For example, linguistic expressions must be encoded as distributed patterns of activation and stored in connectionist memories; Smolensky, 1988, pp. 12–14.)

In contrast, in Smolensky's (1988) model the "intuitive processor" is responsible for most human behavior (and all animal behavior), including perception, skilled motor behavior, and intuitive problem solving and pattern matching. This processor does not rely on language but directly rests on properties of subsymbolic connectionist networks. Learning in this system is slow, occurring only with repeated experience. Processing in this system can be described in rational, symbolic terms, but they will always be imprecise approximations. In Smolensky's summary: "The intuitive processor is a subconceptual connectionist dynamical system that does not admit a complete, formal, and precise conceptual-level description" (p. 7).

Smolensky's (1988) approach seems quite compatible with social psychological dual-process models, though the latter often incorporate important points that Smolensky failed to consider, such as the fact that both cognitive capacity and motivation are typically required for people to use the top-level conscious processor rather than the heuristically based intuitive processor. Thus, it is significant to note that social psychologists in recent years have emphasized the importance of preconscious and implicit processes (see Bargh, 1994; Greenwald & Banaji, 1995; Higgins, 1989). The assumption is that such processes ordinarily determine our conscious experience and therefore direct our thoughts, feelings, and behavior. Only when we are specially motivated to look beneath the surface of things do we apply systematic reasoning and question the results of our preconscious processing (e.g., Martin, Seta, & Crelia, 1990). Arguably, even if connectionist models should prove totally incapable of handling language, so that they are useful in understanding only low-level, nonconscious mental processes, they would still have great relevance to many issues of concern to social psychologists.

Catastrophic Interference in Memory

Another critique of the psychological applicability of connectionist models stems from demonstrations of what has been termed *catastrophic interference* (McCloskey & Cohen, 1989; Ratcliff, 1990). A feed-forward network was trained with the back-propagation procedure to simulate human performance in a paired-associate learning paradigm. First the network learned a series of A–B pairs, such that when Item A was presented at the input the network learned to generate B at the output. Then the same network was trained with a different set of outputs for the same input patterns, A–C. When the network was retested on the A–B set its performance was essentially zero. Humans show considerable interference caused by the A–C learning when tested in the same paradigm but far less than the total forgetting exhibited by this network. These demonstrations have been used to argue that connectionist models are unlikely to provide adequate accounts for human memory performance.

In reality, these demonstrations may simply suggest that the use of a feed-forward network with A as input and B as output is a poor way to model paired-associate learning. Several alternatives are available. One is to learn novel stimuli (e.g., the A–C pairs) in a separate memory system so that the new knowledge can be gradually and nondestructively incorporated into the framework of existing knowledge (the A–B pairs). This is the approach taken by McClelland et al. (1995), as described above. Another approach is to combine the A and B stimuli into a composite pattern to be memorized using a recurrent network, rather than using A as input and B as output from a feed-forward network. Re-presentation of the A component would allow retrieval of B by the network's pattern-completion property. McClelland and Chappell (1995) and others have advanced models of memory that use these approaches and have shown, with detailed comparisons against psychological data, that they do not share the empirical failings of the approach that McCloskey and Cohen (1989) and Ratcliff (1990) found to be inadequate.

The Question of Levels

Some readers may wonder whether connectionist models are at too low a level to make predictions for social psychological variables such as beliefs, attitudes, social judgments, and behaviors. Here are three arguments for a negative answer.

First, as noted above, the history of social psychology over the last generation can be read as a story of an ongoing shift from the study of conscious judgmental and inferential processes to an increasing emphasis on preconscious or heuristic processes and the cognitive representations that underlie them (Devine et al., 1994; Greenwald & Banaji, 1995). Today, influential dual-process models hold that effortful conscious reasoning takes place only under relatively rare circumstances, when people possess both cognitive capacity and strong motivation (Smith, 1994). This shift of theoretical focus has been accompanied by a shift in research methodology, from heavy reliance on questionnaires requiring more or less thoughtful verbal responses to an increased use of process-oriented measures (such as response latencies and memory) that are better able to tap non-

verbal processes. If this view is accurate, the core issues of process and representation about which many social psychologists care are exactly those that are mediated by the connectionist networks of Smolensky's (1988) subsymbolic "intuitive processor." Thus, application of connectionist models to social psychological phenomena would continue the approach of social cognition researchers, who have long assumed that modeling the details of memory representations and cognitive processes will help shed light on social thoughts, feelings, and actions (e.g., Hastie & Kumar, 1979; Wyer & Srull, 1989). Virtually all theories of representation and process in social cognition have been advanced at the symbolic level, because theorists failed to recognize any alternative. Today an alternative is available, and if subsymbolic connectionist theories provide a better account of the details of memory and cognitive processes, we should expect that they will shed light on social behavior as well.

Second, it seems likely that connectionist theories will be better able to model the complexity, flexibility, and dynamic qualities of social behavior—in contrast to the more rigid, static account that most naturally flows from theories based on symbolic rules. To explain the complexity of social behavior, symbolic theories must postulate that perceivers use general rules (or general knowledge structures such as schemas or prototypes) but also process myriad exceptions, qualifications, and special cases. The result is often systems of almost Ptolemaic complexity. In contrast, many connectionist models can handle general rules and exceptions within a common framework (see Seidenberg, 1993) and therefore deal more naturally with the complexity of social behavior. As Smolensky (1988) emphasized, behaviors that are actually generated by a subsymbolic connectionist system can sometimes be globally described by symbolic rules. However, these rule-based descriptions are inevitably approximate and will fail under difficult conditions such as limited cognitive capacity, mixed or inconsistent input information, unclear task demands, and the like.

As an example, Smolensky (1986) described a connectionist network that learned to predict various features such as voltage or current in simple electrical circuits. When tested with well-structured problems and given unlimited time to answer, the network's responses matched those that would be given by exact symbolic rules such as Ohm's Law. Under these conditions an observer who treated the network as a "black box" would be tempted to say that the network possessed explicit representations of such laws and used them to compute its answers. However, this is an illusion, which breaks down when the network is given an ill-posed problem (e.g., one in which some of the given values for the circuit are mutually incompatible) or limited time for consideration. Under these conditions the network gives a sensible performance: It satisfies as many constraints as possible, and it gives an approximate answer quickly and refines it if permitted more time. Such behaviors reflect the performance of a system that satisfies multiple soft constraints as well as possible rather than one that computes using explicit, symbolically represented hard rules. Yet within a core subset of the domain (well-posed problems, plenty of time), the rules are perfectly adequate approximate descriptions of the network's answers. It seems plausible that many "laws" describing human social behavior (such as "people maximize subjective expected utility," or "people make attributions based on observed covariations between potential causes and effects") are similarly rough-and-ready approximate generalizations that may adequately characterize the outcomes of processing under ideal conditions. However, such explicit rules may play no role in the processes people use to make judgments or choose behaviors and may not even describe the outcomes under conditions of limited time, degraded information, or divided attention. In sum, connectionist models promise unified and parsimonious explanations for performance under varying conditions, whereas a symbolic model may characterize performance under ideal conditions but will typically have to be supplemented in an ad hoc fashion with extra heuristic mechanisms and qualifications for less-than-ideal circumstances.

Finally, I argued above that connectionist models yield predictions that are familiar in many ways. They can embody the principle of accessibility and can act like schemas in fleshing out input information based on past experiences, for instance. However, these similarities go hand in hand with several new predictions, some of which were detailed above. The schema that emerges from an underlying connectionist representation will have somewhat different properties from the schema that social psychologists currently postulate (Rumelhart, Smolensky, et al., 1986; Smolensky, 1986). Novel predictions might be expected to be especially prevalent concerning behavior under cognitive load, with divided attention, or with vague, contradictory, quickly presented, or ill-formulated information. Of course, these are conditions that characterize much of social life!

Role of Computer Simulation in an Empirical Science

Social psychology is, and should remain, an empirically based science. In this context, some people question the value of computer simulation methods—in general, not just of connectionist models—based on the belief that running computer simulations replaces running human participants through social psychological experiments. This belief reflects a misconception of the role of computer simulations in the overall logic of the research enterprise. In a traditional empirical research program, a theory is the starting point; specific hypotheses are logically derived from the theory. Then a study is run, and the results are compared against the hypotheses to draw conclusions about the viability of the theory. What is important to understand is that the use of simulation does not change this logic in any way. Its role in the process is not to replace running the study but to facilitate deriving the hypotheses. When a theory exceeds a certain degree of complexity (as with virtually all connectionist models as well as many traditional theories, such as that of Wyer & Srull, 1989), it becomes essentially impossible for unaided human intelligence to accurately and uncontroversially derive the theory's implications—the research hypotheses. Thus, simulation does not replace gathering data from human participants but enhances the logical process of drawing theoretical conclusions from data by making the derivation of hypotheses from theory more precise and reliable (Hastie, 1988).

It is true that many sciences, such as physics, have evolved a rough division of labor between theorists and experimentalists. Though such a division has not historically been prominent in social psychology, it may emerge in the future to the extent that theory development (and data gathering as well) come to re-

quire more and more sophisticated and specialized skills. For the foreseeable future, probably only a small percentage of social psychologists will wish to work with connectionist models at the level of units, connections, and activations, to investigate (for example) the implications of novel network architectures or learning rules for social psychological phenomena. Other social psychologists, however, may well wish to apply connectionist models at the higher level—at which most of the discussion in this article is presented—of distributed representations and their properties such as context sensitivity and multiple constraint satisfaction. They can do this by applying standard models with well-understood properties (such as back-propagation pattern transformation networks or recurrent memory networks) without involving themselves in the low-level details of learning rules or activation equations. An analogy can be made with statistics. A few individuals with special expertise develop the mathematics behind new data analytic procedures, which other researchers can then apply without necessarily understanding all the mathematical details. Statistical computer packages, and now connectionist simulation packages (see below), permit wide access to standard procedures by nonspecialist researchers. Perhaps in an ideal world we would all understand the mathematical details of our connectionist models as well as our statistical procedures from the ground up. However, life is short, and the division of labor within science means that we ordinarily are content to use the results of others' research without knowing all the low-level details.

How to Learn More About Connectionism

A final potential question is how an interested social psychologist can learn more about connectionism. Acquiring any new body of knowledge takes some investment of time and effort. However, connectionism is not particularly difficult to learn or to understand. The fundamentals of connectionist models are no more abstract or hard to grasp than many statistical topics that every working social psychologist presumably commands, such as the analysis of variance.[2] A number of excellent tutorial introductions to connectionism have been published, including many chapters in Volume 1 of Rumelhart, McClelland, et al. (1986). (The reader is cautioned that most of the illustrative psychological applications in volume 2 of the 1986 "PDP Bible" are far from the current state of the art; many applications can be found in the current journal literature.) Once the basic principles are understood, Smolensky (1988) and recent books by Churchland and Sejnowski (1992) and Clark (1993) offer outstanding discussions of connectionist models treated as psychological theories. It is important to understand that the overall connectionist literature is broad and multidisciplinary. Many published articles and chapters focus on philosophical, engineering, computational, or neurobiological aspects of connectionist modeling and so are less likely to be directly relevant to readers interested in psychological issues. (One possible reason that many social psychologists have received the impression that connectionism is boringly irrelevant to their work is that they have unluckily happened to pick up an article of one of these types.)

For those who wish to go beyond understanding other investigators' applications of connectionist models to actually applying models themselves, computer simulation software is essential. Caudill and Butler (1992a, 1992b) offer software for constructing small-scale models, ideal for tutorial purposes and perhaps for some actual research applications. This software is inexpensive and runs on widely available IBM-compatible PCs under DOS. Several systems that run under many varieties of the UNIX operating system are available for running large-scale connectionist simulations. My laboratory uses the Stuttgart Neural Network Simulator (SNNS; Zell et al., 1994; URL: ftp://ftp.informatik.uni-stuttgart.de/pub/SNNS). An alternative is the PDP++ system (URL: ftp://hydra.psy.cmu.edu/pub/pdp++). Both systems are available free of charge. However, as with SAS or any comparably massive software system, a significant investment of time is required to learn how to use a large-scale connectionist simulation package. This time, however, is well spent; there is no real substitute for playing around with models to develop intuitions about their behavior.

Summary and Conclusions

What Can Connectionism Offer Social Psychology?

In this article I have pointed out that connectionist models can be described at two levels. Properties at the higher level—the level at which one can speak of representations, processes, memory, pattern completion, and constraint satisfaction—offer powerful new tools for theoretical development in social psychology. They will be useful both for constructing new and integrative accounts of known findings and for generating new predictions and inspiring empirical studies to test these predictions. Among these higher level properties are the following:

1. Processing has the character of simultaneously satisfying many soft constraints as well as possible rather than applying hard, exceptionless rules.

2. Symbolic rules may approximately describe the results of processing under ideal circumstances but are not necessarily part of the process itself (unless they are explicitly used by the conscious symbolic processor). Under less-than-ideal circumstances, such rules will become less and less adequate as descriptions of the outcome.

3. Constraint-satisfaction processing implies that changes in any type of cognition (belief, attitude, or motive) will generally lead to corresponding adjustments in others as the overall system state shifts from one attractor to another; psychological causation is not unidirectional.

4. Representations emerge (as reconstructions) from activation patterns set up by input cues rather than being located (and retrieved from memory unchanged) by a search process.

5. Representations can be flexibly recombined and are intrinsically context sensitive rather than being retrieved in invariant form each time they are used.

[2] In fact, many types of connectionist models can be understood as statistical models. For example, an unsupervised network that operates as a "feature detector," finding covariations among features in the input patterns, is effectively performing a factor analysis. Some of the relationships between connectionist models and statistics, which have been outlined by Sarle (1994), may help some readers in understanding properties of the connectionist networks.

6. Representation and process are inextricably intertwined; the weights on the connections serve as both the network's representational structure and as the determinants of processing.

7. Representation (and hence processing) are changed incrementally by experience, giving rise to the principle of accessibility.

8. Representations reflect learned regularities in the environment and can be used to help interpret new experiences by filling in typical values for unobserved features, for instance. In some complex, multiple-module connectionist systems, other memory systems may specifically look for and quickly learn about deviations from typical or expected properties.

Though these properties generally characterize most distributed connectionist models, widely varying models have been proposed, and not all possess all these properties to the same degree. The properties actually flow from the workings of a model at the lower level—the level of units, connection strengths, learning rules, and activation flows—and simulations are required to derive the higher level properties from a low-level description of a network. Not all social psychologists need to engage themselves in this work, but some probably should; leaving it entirely to our colleagues in cognitive psychology may mean that important issues such as affect and motivation remain unconsidered.

What Can Social Psychology Offer Connectionism?

Part of the promise of the connectionist approach is that a common theoretical language will result in increased integration across different areas of psychology, particularly cognitive, developmental, and social. Social psychology stands to be a contributor to, as well as a beneficiary of, this increased integration. Our contributions may be especially important in areas such as these:

1. Accessibility. Though the principle of accessibility is acknowledged in some cognitive theories (e.g., J. R. Anderson, 1993), most theoretical and empirical development to date has been by social psychologists (see Higgins, in press). Analyses of the ways in which recent and frequent use of a concept affect its potential for further use, and the role of accessibility in understanding individual differences in perception, motivation, and social behavior are important and unique contributions that social psychologists can bring to an overall integrative psychology.

2. Social interaction. With the rise of connectionist models, a common theoretical framework—based on the idea of rich informational connections among processing units—can be applied both within and between individuals. The dynamic models of opinion structure developed by Nowak, Szamrej, and Latané (1990), which assume that people interact and change their opinions toward the majority of their interaction partners, show roughly how the interpersonal component of such a model might look. Hutchins's (1991) model is the single example of which I am aware that simultaneously models individual belief structures (though using localist constraint-satisfaction networks rather than distributed connectionist models) and communication between individuals.

3. Affect and motivation. Connectionist workers in cognitive psychology have developed theories of such important phenomena as memory, categorization, and language comprehension. The theories include multiple modules to represent visual, orthographic, and semantic information. Adding modules for self-regulatory, motivational, and affective systems will permit understanding of additional phenomena, which have been much more intensively studied within social psychology. For example, we might assume that some concepts (such as "cockroach," represented in a semantic network) are linked to affect or motives (such as disgust or avoidance, represented in other networks). Perhaps most important, connectionist models may capture many types of cognition–motivation interaction by assuming that accessible motives act as constraints which, along with other types of representations of past experiences and the current input information, influence the system's dynamic evolution over time and the particular attractor to which it finally converges. This picture offers the promise of an integrated account of motivation and cognition.

Connections to Our History?

The idea of cognitive dynamics has a long and illustrious history within social psychology. In the 1985 *Handbook of Social Psychology*, Markus and Zajonc wrote

> The . . . chapter [on cognitive approaches] in the second edition of the *Handbook* (Lindzey and Aronson, 1968, p. 391) ended with the expectation that the emphasis on cognitive dynamics prevalent during the sixties, with its particular focus on cognitive dissonance and balance, would soon be combined with the earlier descriptive approaches that focused on the structural and substantive properties of cognitions. This expectation for an integrated approach to social cognition was definitely not realized. Not only have the seventies and the early eighties failed to achieve a synthesis of the dynamic and descriptive approaches, but for the most part they have abandoned cognitive dynamics altogether. Today's cognitive approaches in social psychology show little concern with the dynamic properties of cognitions—those that posit forces and interdependence among cognitions and produce changes over time. (p. 139)

As this quotation suggests, the types of theories that were prevalent in the mid-1980s did not offer ready accounts for many dynamic properties of cognition, despite occasional vague, generally unelaborated claims (e.g., the idea that schemas may include "processing elements" as well as knowledge structures).

Today we can perhaps see how to bring dynamics back into our conceptions of mental representation. In social psychology, the advent of connectionist models would not be a revolution so much as a continuation of an ongoing transition from static to dynamic conceptions of mind (Kruglanski, 1994). In the 1950s, social psychologists described separate categories of cognitions (such as beliefs, attributions, person impressions, or attitudes) and developed distinct processing laws for each. Later, with the rise of social cognition, theorists tended to treat all types of cognitive content as stored in common representational formats (e.g., in storage bins) and subject to common processing principles (e.g., accessibility). Distributed connectionist models carry this trend toward cognitive dynamics still further. Not only do we no longer have distinct categories of cognitions, but we also do not have discrete representations at all. Not only do common processing principles apply to all types of representation, but representations also do not even exist independent

of process: They are elicited on the spot when needed rather than being "stored" and later "retrieved" unchanged.

Although they arguably represent the continuation into the future of ongoing theoretical trends, connectionist models also have deep resonances with the history of social psychology. Symbolic models grew out of research on solving well-formulated problems (e.g., Newell & Simon, 1972), but connectionist models have much in common with the Gestalt psychologists' view of problem solving, as others also have noted (Holyoak & Spellman, 1993; Read, Vanman, & Miller, 1994). The Gestaltists focused on ill-defined problems in which there may be no hard rules defining a single solution but instead multiple soft constraints to be satisfied as well as possible. Gestalt theorists emphasized that the "whole" (the resulting interpretation) may have emergent properties that make it different from the "sum of its parts." Similarly, the variability and context-sensitivity of concepts and cognitive processes is more consistent with the Gestalt view—and the connectionist view—than with the traditional symbolic approach in which concepts are hard, atomic entities. Finally, the Gestaltists emphasized what they termed *perceptual* and we would now call *implicit* or *automatic* processes. Our modern dual-process theories hold that this type of processing is the way we operate most of the time.

Of course, our present day knowledge has advanced beyond the assumptions of Gestalt theory in many ways. With the heyday of dissonance theory long past, we now realize that consistency or perceptual coherence cannot be assumed to be a single, all-powerful social motive. We also know much more about the cognitive mechanisms underlying perceptual interpretation, judgment, and behavior. The Gestalt approach of the 1930s relied more on metaphor than on the postulation of concrete theoretical mechanisms. Now, within the new connectionist approach we can see the outlines of mechanisms that can capture the Gestalt insights—in fact, mechanisms of the very simplest sort: units that receive activations, sum them, and pass them on to other units. Fuzzy talk about "holism" and "Gestalts," though it was the only resource available in the 1930s, will not lead to theoretical progress in the 1990s; we need instead to come to grips with explicit mechanisms and models that offer those properties. This process will require focused theoretical and empirical development over the coming years. Achieving this goal may also impose the cost of giving up some familiar and therefore appealing ideas, such as the notion that all representations are symbolic in nature, explicit encodings assembled from elements representing concepts—constituting almost a language of thought (Bickhard & Terveen, 1995; Clark, 1993).

I conclude with one additional quotation concerning the promise of connectionism:

> Might [connectionism] turn out to be a seductive blind alley? Yes. Might it be the beginning of a revolution in the study of the mind? Yes. Let us find out which, by getting on with the building and testing of the models. For the fact is that connectionist models actually do surprising things, and if they did not, they would not have sustained [the interest they have]. (Dennett, 1991, pp. 28–29)

References

Anderson, J. A. (1995). Associative networks. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 102–107). Cambridge, MA: MIT Press.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, J. R. (Ed.). (1993). *Rules of the mind.* Hillsdale, NJ: Erlbaum.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory.* Washington, DC: Winston & Sons.

Anderson, N. H. (1981). *Foundations of information integration theory.* New York: Academic Press.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., Vol. 1, pp. 1–40). Hillsdale, NJ: Erlbaum.

Bargh, J. A., Bond, R. N., Lombardi, W. J., & Tota, M. E. (1986). The additive nature of chronic and temporary sources of construct accessibility. *Journal of Personality and Social Psychology, 50,* 869–878.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (in press). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology.*

Bargh, J. A., Lombardi, W. J., & Higgins, E. T. (1988). Automaticity of chronically accessible constructs in person × situation effects on person perception: It's just a matter of time. *Journal of Personality and Social Psychology, 55,* 599–605.

Barnden, J. A. (1995a). Artificial intelligence and neural networks. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 98–102). Cambridge, MA: MIT Press.

Barnden, J. A. (1995b). Semantic networks. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 854–857). Cambridge, MA: MIT Press.

Bickhard, M. H., & Terveen, L. (1995). *Foundational issues in artificial intelligence and cognitive science: Impasse and solution.* Amsterdam: Elsevier Scientific.

Carlston, D. E. (1994). Associated systems theory: A systematic approach to cognitive representations of persons. In R. S. Wyer & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 7, pp. 1–78). Hillsdale, NJ: Erlbaum.

Carlston, D. E., & Smith, E. R. (in press). Principles of mental representation. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles.* New York: Guilford Press.

Caudill, M., & Butler, C. (1992a). *Understanding neural networks: Computer explorations Volume 1: Basic networks.* Cambridge, MA: MIT Press.

Caudill, M., & Butler, C. (1992b). *Understanding neural networks: Computer explorations Volume 2: Advanced networks.* Cambridge, MA: MIT Press.

Chappell, M., & Humphreys, M. S. (1994). An auto-associative neural network for sparse representations: Analysis and application to models of recognition and cued recall. *Psychological Review, 101,* 103–128.

Churchland, P. S., & Sejnowski, T. J. (1989). Neural representation and neural computation. In L. Nadel, L. A. Cooper, P. Culicover, & R. M. Harnish (Eds.), *Neural connections, mental computation* (pp. 15–48). Cambridge, MA: MIT Press.

Churchland, P. S., & Sejnowski, T. J. (1992). *The computational brain.* Cambridge, MA: MIT Press.

Clark, A. (1993). *Associative engines: Connectionism, concepts, and representational change.* Cambridge, MA: MIT Press.

Damasio, A. R. (1989). Multiregional activation: A systems level model for some neural substrates of cognition. *Cognition, 33,* 25–62.

Dennett, D. (1991). Mother Nature versus the walking encyclopedia: A Western drama. In W. Ramsey, S. P. Stich, & D. E. Rumelhart (Eds.), *Philosophy and connectionist theory* (pp. 21–30). Hillsdale, NJ: Erlbaum.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and

controlled components. *Journal of Personality and Social Psychology, 56,* 5–18.

Devine, P. G., Ostrom, T. M., & Hamilton, D. L. (Eds.). (1994). *Social cognition: Impact on social psychology.* Orlando, FL: Academic Press.

Dorman, C., & Gaudiano, P. (1995). Motivation. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 591–594). Cambridge, MA: MIT Press.

Elman, J. L. (1995). Language processing. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 508–513). Cambridge, MA: MIT Press.

Estes, W. K. (1994). *Classification and cognition.* New York: Oxford University Press.

Fazio, R. H. (1993). Variability in the likelihood of automatic attitude activation: Data reanalysis and commentary on Bargh, Chaiken, Govender, and Pratto (1992). *Journal of Personality and Social Psychology, 64,* 753–758.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50,* 229–238.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior.* Reading, MA: Addison-Wesley.

Fodor, J. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind.* Cambridge, MA: MIT Press.

Fodor, J., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. In S. Pinker & J. Mehler (Eds.), *Connections and symbols* (pp. 3–71). Cambridge, MA: MIT Press.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46,* 107–119.

Gillund, G., & Shiffrin, R. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–67.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding: Part I. Parallel development and coding of neural feature detectors. *Biological Cybernetics, 23,* 121–134.

Hamilton, D. L., Katz, L. B., & Leirer, V. (1980). Organizational processes in impression formation. In R. Hastie, T. M. Ostrom, E. B. Ebbesen, R. S. Wyer, D. L. Hamilton, & D. E. Carlston (Eds.), *Person memory* (pp. 121–153). Hillsdale, NJ: Erlbaum.

Hastie, R. (1988). A computer simulation model of person memory. *Journal of Experimental Social Psychology, 24,* 423–447.

Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology, 37,* 25–38.

Henderson, J. B. (1994). *Description based parsing in a connectionist network.* Unpublished doctoral dissertation, University of Pennsylvania.

Higgins, E. T. (1989). Knowledge accessibility and activation: Subjectivity and suffering from unconscious sources. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 75–123). New York: Guilford Press.

Higgins, E. T. (in press). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles.* New York: Guilford Press.

Higgins, E. T., & Bargh, J. A. (1987). Social cognition and social perception. *Annual Review of Psychology, 38,* 369–426.

Higgins, E. T., Bargh, J. A., & Lombardi, W. (1985). The nature of priming effects on categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 59–69.

Higgins, E. T., & King, G. A. (1981). Accessibility of social constructs: Information-processing consequences of individual and contextual variability. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality, cognition, and social interaction* (pp. 69–122). Hillsdale, NJ: Erlbaum.

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review, 98,* 74–95.

Holyoak, K. J., & Spellman, B. A. (1993). Thinking. *Annual Review of Psychology, 44,* 265–315.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review, 96,* 208–233.

Hutchins, E. (1991). The social organization of distributed cognition. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 283–307). Washington, DC: American Psychological Association.

Jordan, M. I. (1989). Serial order: A parallel distributed processing approach. In J. L. Elman & D. E. Rumelhart (Eds.), *Advances in connectionist theory.* Hillsdale, NJ: Erlbaum.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93,* 136–153.

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

Kirkpatrick, L. A., & Epstein, S. (1992). Cognitive-experiential self-theory and subjective probability: Further evidence for two conceptual systems. *Journal of Personality and Social Psychology, 63,* 534–544.

Kruglanski, A. (1994, October). *Commentary at symposium on "Second thoughts on basic assumptions in social psychology."* Society of Experimental Social Psychology convention, Lake Tahoe, NV.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22–44.

Loftus, E. F. (1979). *Eyewitness testimony.* Cambridge, MA: Harvard University Press.

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

Markus, H., & Zajonc, R. B. (1985). The cognitive perspective in social psychology. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 1, pp. 137–230). New York: Random House.

Martin, L. L., Seta, J. J., & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. *Journal of Personality and Social Psychology, 59,* 38–49.

Masson, M. E. J. (1991). A distributed memory model of context effects in word identification. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 233–263). Hillsdale, NJ: Erlbaum.

McClelland, J. L., & Chappell, M. (1995). *Familiarity breeds differentiation: A Bayesian approach to the effects of experience in recognition memory.* Unpublished manuscript, Carnegie-Mellon University.

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,* 419–457.

McClelland, J. L., & Rumelhart, D. E. (1986). A distributed model of human learning and memory. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2, pp. 170–215). Cambridge, MA: MIT Press.

McClelland, J. L., Rumelhart, D. E., & The PDP Research Group (Eds.). (1986). *Parallel distributed processing* (Vol. 2). Cambridge, MA: MIT Press.

McClelland, J. L., Rumelhart, D. E., & Hinton, G. E. (1986). The appeal of parallel distributed processing. In D. E. Rumelhart, J. L. McClelland, & The PDP Research Group (Eds.), *Parallel distributed processing* (Vol. 1, pp. 3–44). Cambridge, MA: MIT Press.

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in

connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 24, pp. 109–165). New York: Academic Press.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227–234.

Miikkulainen, R. (1993). *Subsymbolic natural language processing.* Cambridge, MA: MIT Press.

Miller, L. C., & Read, S. J. (1991). On the coherence of mental models of persons and relationships: A knowledge structure approach. In G. J. O. Fletcher & F. Fincham (Eds.), *Cognition in close relationships* (pp. 69–99). Hillsdale, NJ: Erlbaum.

Moll, M., Miikkulainen, R., & Abbey, J. (1994). The capacity of convergence-zone episodic memory. In *Proceedings of the Twelfth National Conference on Artificial Intelligence* (Vol. 1, pp. 68–73). Menlo Park, CA: American Association for Artificial Intelligence.

Moscovitch, M. (1994). Memory and working with memory: Evaluation of a component process model and comparisons with other models. In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 269–310). Cambridge, MA: MIT Press.

Newell, A. (1980). Physical symbol systems. *Cognitive Science, 4,* 135–183.

Newell, A., & Simon, H. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice Hall.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115,* 39–57.

Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review, 97,* 362–376.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. In S. Pinker & J. Mehler (Eds.), *Connections and symbols* (pp. 73–193). Cambridge, MA: MIT Press.

Plate, T. A. (1994). *Distributed representations and nested compositional structure.* Unpublished doctoral dissertation, University of Toronto.

Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a multilayered perceptron: Implications for child language acquisition. *Cognition, 38,* 43–102.

Ramsey, W., Stich, S. P., & Garon, J. (1991). Connectionism, eliminativism, and the future of folk psychology. In W. Ramsey, S. P. Stich, & D. E. Rumelhart (Eds.), *Philosophy and connectionist theory* (pp. 199–228). Hillsdale, NJ: Erlbaum.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R. (1990). Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review, 97,* 285–308.

Read, S. J., Vanman, E. J., & Miller, L. C. (1994). *Parallel constraint satisfaction processes and Gestalt principles: (Re)introducing cognitive dynamics to social psychology.* Unpublished manuscript, Department of Psychology, University of Southern California.

Rosenblatt, F. (1962). *Principles of neurodynamics.* New York: Spartan.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review, 96,* 341–357.

Rueckl, J. G. (1990). Similarity effects in word and pseudoword repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 374–391.

Rueckl, J. G., & Olds, E. M. (1993). When pseudowords acquire meaning: Effect of semantic associations on pseudoword repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 515–527.

Rumelhart, D. E. (1992). Towards a microstructural account of human reasoning. In S. Davis (Ed.), *Connectionism: Theory and practice* (pp. 69–83). New York: Oxford University Press.

Rumelhart, D. E., & McClelland, J. L. (1986). PDP models and general issues in cognitive science. In D. E. Rumelhart, J. L. McClelland, & The PDP Research Group (Eds.), *Parallel distributed processing* (Vol. 1, pp. 110–146). Cambridge, MA: MIT Press.

Rumelhart, D. E., McClelland, J. L., & The PDP Research Group. (Eds.). (1986). *Parallel distributed processing* (Vol. 1). Cambridge, MA: MIT Press.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland, D. E. Rumelhart, & The PDP Research Group (Eds.), *Parallel distributed processing* (Vol. 2, pp. 7–57). Cambridge, MA: MIT Press.

Rumelhart, D. E., & Todd, P. M. (1990). Learning and connectionist representations. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 3–30). Cambridge, MA: MIT Press.

Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated recruitment of autobiographical memories. *Journal of Personality and Social Psychology, 59,* 229–241.

Sarle, W. S. (1994, April). Neural networks and statistical models. In SAS Institute (Ed.), *Proceedings of the Nineteenth Annual SAS Users Group International Conference* (pp. 1538–1550). Cary, NC: SAS Institute, Inc. [Available at URL ftp://ftp.sas.com/pub/sugi19/neural/neural1.ps]

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 501–518.

Schacter, D. L. (1994). Priming and multiple memory systems: Perceptual mechanisms of implicit memory. In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 233–268). Cambridge, MA: MIT Press.

Schacter, D. L., & Tulving, E. (Eds.). (1994). *Memory systems 1994.* Cambridge, MA: MIT Press.

Schneider, W. (1987). Connectionism: Is it a paradigm shift for psychology? *Behavior Research Methods, Instrumentation, and Computers, 19,* 73–83.

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

Sejnowski, T., & Rosenberg, C. (1987). Parallel networks that learn to pronounce English text. *Complex Systems, 1,* 145–168.

Shastri, L. (1995). Structured connectionist models. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 949–952). Cambridge, MA: MIT Press.

Sloman, S. A., Hayman, C. A. G., Ohta, N., Law, J., & Tulving, E. (1988). Forgetting in primed fragment completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 223–239.

Smith, E. R. (1982). Beliefs, attributions, and evaluations: Nonhierarchical models of mediation in social cognition. *Journal of Personality and Social Psychology, 43,* 248–259.

Smith, E. R. (1994). Procedural knowledge and processing strategies in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., Vol. 1, pp. 99–151). Hillsdale, NJ: Erlbaum.

Smith, E. R., & Branscombe, N. R. (1987). Procedurally mediated social inferences: The case of category accessibility effects. *Journal of Experimental Social Psychology, 23,* 361–382.

Smith, E. R., Stewart, T. L., & Buttram, R. T. (1992). Inferring a trait from a behavior has long-term, highly specific effects. *Journal of Personality and Social Psychology, 62,* 753–759.

Smith, E. R., & Zárate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review, 99*, 3–21.
Smolensky, P. (1986). Information processing in dynamical systems: Foundations of harmony theory. In D. E. Rumelhart, J. L. McClelland, & The PDP Research Group (Eds.), *Parallel distributed processing* (Vol. 1, pp. 194–281). Cambridge, MA: MIT Press.
Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences, 11*, 1–74.
Smolensky, P. (1989). Connectionist modeling: Neural computation/mental connections. In L. Nadel, L. A. Cooper, P. Culicover, & R. M. Harnish (Eds.), *Neural connections, mental computation* (pp. 49–67). Cambridge, MA: MIT Press.
Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. In W. G. Koster (Ed.), *Attention and performance II* (pp. 276–315). Amsterdam: North-Holland.
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.
Touretzky, D. S. (1995). Connectionist and symbolic representation. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 243–247). Cambridge, MA: MIT Press.
van Gelder, T. (1991). What is the "D" in "PDP"? A survey of the concept of distribution. In W. Ramsey, S. P. Stich, & D. E. Rumelhart (Eds.), *Philosophy and connectionist theory* (pp. 33–59). Hillsdale, NJ: Erlbaum.
Weber, E. U., Goldstein, W. M., & Busemeyer, J. R. (1991). Beyond strategies: Implications of memory representation and memory processes for models of judgment and decision making. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory to data: Essays on human memory in honor of Bennet B. Murdock* (pp. 75–101). Hillsdale, NJ: Erlbaum.
Wiles, J., & Humphreys, M. S. (1993). Using artificial neural nets to model implicit and explicit memory test performance. In P. Graf & M. E. J. Masson (Eds.), *Implicit memory: New directions in cognition, development, and neuropsychology* (pp. 141–165). Hillsdale, NJ: Erlbaum.
Wilson, T. D., & Hodges, S. D. (1992). Attitudes as temporary constructions. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 37–65). Hillsdale, NJ: Erlbaum.
Wyer, R. S., & Srull, T. K. (1989). *Memory and cognition in its social context*. Hillsdale, NJ: Erlbaum.
Zell, A., Mamier, G., Vogt, M., Mache, N., Hübner, R., Herrmann, K., Soyez, T., Schmalzl, M., Sommer, T., Hatzigeorgiou, A., Döring, S., Posselt, D., & Schreiner, T. (1994). *Stuttgart neural network simulator User Manual, Version 3.2* (Tech. Rep. No. 3/94). University of Stuttgart, Institute for Parallel and Distributed High Performance Systems.

Received February 6, 1995
Revision received October 15, 1995
Accepted October 24, 1995 ■