Attitude Representation: Attitudes as Patterns in a Distributed, Connectionist Representational System

Frederica R. Conrey and Eliot R. Smith
Indiana University, Bloomington

In press, Social Cognition
July 2006

Preparation of this manuscript was supported by NSF grants BCS-0091807 and BCS-0527249. Address correspondence to Eliot R. Smith, Department of Psychology and Brain Sciences, 1101 E. Tenth St., Indiana University, Bloomington, IN 47405 or esmith4@indiana.edu.
Attitude Representation: Attitudes as Patterns in a Distributed, Connectionist Representational System

Frederica R. Conrey and Eliot R. Smith
Indiana University, Bloomington

Abstract

How we think about the representation of attitudes has a profound impact on how we think about attitudes themselves, attitude change, and the attitude-behavior relationship. In this paper, we briefly review the model of attribute representation in a distributed, connectionist memory system, which portrays attitudes as time-dependent states of the system rather than as static “things” that are “stored” in memory. This model is particularly well-suited to addressing some of the field’s most pressing questions about the multiplicity of attitudes and their stability (or instability) over time. We address several of these questions from the distributed, connectionist perspective, concluding that the new model renders some questions meaningless, suggests straightforward answers to others, and hints at exciting new hypotheses about the answers to still others.
An attitude seems like a fairly straightforward thing; it is often viewed as simply an association in the mind between an object such as the Bush administration, a dog, or a piece of cheese, and a positive or negative evaluation (Fazio, 1986). Attitudes are nothing more than people’s likes and dislikes, but they have been the subject of intensive research in social psychology for more than fifty years because it turns out that people’s likes and dislikes make up a large part of who they are and what they do. For social psychologists, understanding attitudes is the first step to understanding human behavior.

In order to understand what people like and dislike, we need to understand how they think about the objects of their attitudes. Understanding how an attitude toward dogs is represented can tell us something about what the attitude toward puppies will be, whether the attitude toward dogs will change, and whether a positive dog attitude will lead to the purchase of a dog. Here, we briefly summarize the way we prefer to think about representation and point out some implications of this model for common questions about attitudes and how attitudes relate to behavior.

Connectionist Models of Attitude Representation

In our view, attitudes (as well as beliefs, stereotypes, the self-concept, and all other types of mental representation) can best be understood as being represented as particular states of a connectionist network (Smith, 1998). A connectionist network, intended as a loose analog of a biological brain, is a large number of simple processing units (often termed nodes) that are richly interconnected and send signals to each other depending on the nodes’ individual levels of activation (Smith, 1996, 1998). With appropriate assumptions about the behaviors of the individual units and the patterns of
connectivity, such a network can perform important information-processing tasks, including memory, categorization, and pattern transformation.

One key distinction, to which we will return later, is that “connectionist” models can be assumed to make use of either localist or distributed representations (Thorpe, 1995; Touretzky, 1995). In this article we emphasize the latter. In a distributed representation system no individual node or unit has any specific meaning such as “dog” or “fur” or “positive evaluation.” In contrast, semantically meaningful concepts are represented by distinct patterns of activation across large numbers of units. A useful analogy is the set of individual pixels on a TV screen. No individual pixel means anything in particular, so its specific value of brightness or color has no distinct interpretation. Yet by taking on different patterns of illumination, the entire set of pixels can represent a huge number of discrete concepts or objects.

A distributed, connectionist network processes information in the form of flows of activation across the connections between the units. Each unit receives input over incoming connections from other units, and integrates those signals to determine its own activation at the next moment in time. In turn, it sends output signals that are dependent on its activation level over its outgoing connections, to influence other units’ activity. Connections are weighted, so that unit j receives an incoming signal that is the outgoing signal from unit i multiplied by the weight on the connection from unit i to unit j. While each unit’s activation level may change rapidly from moment to moment, connection weights are assumed to change only slowly over time. The entire set of connection weights in the network has two important properties. Along with the external inputs to the network, the set of connection weights fully determines the network’s processing
patterns and capabilities. The set of weights also serves as the network’s long-term memory or source of stability in its operation.

A pattern of activation that the network tends to take on in response to a particular configuration of external inputs is termed an attractor (Hertz, 1995). A set of connection weights that leads to the formation of an attractor is not ordinarily put in place by some unspecified external mechanism, but is built up gradually over time by the network itself using a learning rule. After each of many patterns is input to the network, a specific learning rule is used to make small, incremental changes in the connection weights. The effect of the learning process is to make the network better able to reconstruct patterns of activation that it frequently entered in the past, based on the external input patterns. These patterns, then, become attractors of the network, which can be regarded as implementing a form of memory storage. For example if cues representing “barks” and “fur” are input to the network, it may settle into an attractor pattern that represents “dog” – if those inputs have frequently been present along with that overall learned pattern in the past. We could describe this process as “categorizing” the input as a dog, or as “recalling” the dog concept in response to those cues. But a better term is “reconstructing” a previously experienced, meaningful pattern of activation (i.e., an attractor).

**Attitude Retrieval Versus Context-Sensitive Reconstruction**

One important question in the attitude literature is whether attitudes are constructed on the spot in ways that vary from time to time, or reflect more stable, trait-like properties. The functional properties of distributed, connectionist memory systems differ from more familiar types of memory storage on which our intuitions generally rely. For example, in a file cabinet one might “retrieve” a file related to a particular concept
and find therein the information that was “stored” for that concept at an earlier time. The
same information would be there on each occasion when the file was retrieved. This
metaphor underlies conventional symbolic associative models of memory. Such models
suggest that there is a representation of “dog” that can be retrieved and its contents
(including evaluative information) examined. In contrast, in a distributed connectionist
system, there is not a single, unchanging representation for each concept. Instead, infinite
potential representations can be represented by the nodes of the connectionist network.
Representations in the connectionist network don’t exist, they occur; they are
reconstructed from the unique configuration of inputs each time they come to mind. Can
these assumptions account for the properties of human memory that have been revealed
by experimental research? Yes; in fact, numerous specific connectionist models have
been proposed that are versions of the general framework described here, and have been
shown to fit experimental data (e.g., Chappell & Humphreys, 1994; McClelland,
McNaughton, & O’Reilly, 1995). In the social psychological literature as well,
connectionist models of this sort have been shown to display many key properties such as
temporary accessibility (priming), chronic accessibility, and stereotype learning and
application (Smith, 1996, 1998; Queller & Smith, 2002).

Understanding that representations in a distributed connectionist network are
states rather than things makes it easier to understand that exactly what is reconstructed
depends on the entire configuration of features that are currently present, not just the
focal retrieval cues. While a connectionist memory system is excellent at extracting
commonalities in patterns across multiple exposures, each individual reconstruction of a
representation is a context-sensitive version of the concept (Clark, 1993). “Barks” and
“fur” in the context of an aggressive-sounding bark and bristling fur may lead to the reconstruction of a fearsome German Shepherd watchdog representation, while a yippy bark and short, curly fur may result in the reconstruction of a quite different representation of a small lapdog. This type of context sensitivity is a natural, indeed inevitable consequence of distributed connectionist representational assumptions, and it is pervasively found in human representations of concepts (e.g., Yeh & Barsalou, in press).

This discussion has emphasized context sensitivity of representations, raising the question of whether a distributed representational system has any way of reconstructing more or less the same representation of a concept across multiple occasions – that is, of accounting for a certain degree of stability (as well as context-sensitivity). The answer is that given sufficient experience with a domain, someone may learn to activate roughly the same pattern in many different contexts. This is accomplished by focusing on the key inputs that trigger that particular attractor (while ignoring other inputs, even highly salient ones, as irrelevant), and this ability to focus is precisely what constitutes domain expertise. For example, while a young child may represent a Shih Tzu and a Doberman with entirely different activation patterns, maturation brings the ability to understand that these are both instances of the same “dog” concept (Keil, 1989; see discussion in Smith & DeCoster, 2000, p. 115). In this sense, the ability to activate a context-free representation despite wide variability in superficial features is a significant cognitive accomplishment (rather than being the default or basic state), as discussed by Clark (1993) and Elman (1995). As another example, a political novice may activate different attitudes for different political issues such as immigration restrictions or cuts in estate taxes, while a political expert may activate a single “liberal” or “conservative”
representational pattern and corresponding attitude for each issue, despite their superficial differences. As these examples suggest, a distributed connectionist system generates context-sensitive representations as a fundamental property, but can also account for stability (relative freedom from contextual influences), as domain-specific learning and experience allow a focus on the most important dimensions and neglect of inessential ones.

Localist versus distributed representations

As we have said, representational models that involve nodes connected by links and can therefore be labeled “connectionist” are of two fundamentally different types: distributed (which we emphasize in this article) and localist. The distinction involves the question of whether a meaningful concept is represented by a pattern of activation over an entire set of nodes, or by a single node (or small local group of nodes that are strongly interconnected so they essentially act as a unit). The properties of these two types of models are extensively reviewed elsewhere (Clark, 1993; Barnden, 1995; Thorpe, 1995; Touretzky, 1995). For our purposes the key differences are two.

1. Learning new concepts. In a localist network, when a new concept is learned, some process external to the network must be assumed to create the new node that is needed to represent the concept. Learning in a localist network can operate to tune the strength of the connection between two existing nodes (e.g., increasing connection strength when the two concepts appear together). However, this process does not account for the creation or assignment of a new node – so an entirely distinct form of learning has to be postulated for learning a new concept. In contrast, a distributed network can readily learn a new pattern (attractor) to represent a new concept; the
standard operation of the learning rules accomplishes this without the need for any external mechanism or homunculus.

2. **Context sensitive versions of concepts.** A distributed network automatically learns and reconstructs context-specific versions of a concept, as described above. A localist representational model can account for context sensitivity only by postulating a later (and often mysterious) processing stage. That is, a localist model may assume that the nodes representing “dog” and “fierce” (or “dog,” “small,” and “yippy”) are concurrently activated, but some external process must be postulated that can take those activations and combine them into a context-specific version of the “dog” concept (such as a fierce watchdog or a small lapdog). Because in a localist network a node represents a fixed symbolic meaning, the meaning of the “dog” node itself does not vary from occasion to occasion. People can certainly construct context-specific versions of concepts (e.g., Barsalou, 1987), but the processes involved in conceptual combination are complex and often go unmodeled in localist theories.

On both of these counts, we believe that distributed connectionist models offer more parsimonious explanations compared to localist models: both the learning of new concepts and the construction of context-specific variants of concepts occur through the normal operation of a distributed network’s learning rules. Context-specificity is intrinsic to the representational system itself, in other words. In contrast, in a localist system new concept learning and context sensitivity need to be accomplished by external processes, which have not been explicitly modeled in most theories.

**Summary: Attitudes as States in a Distributed, Connectionist System**
A connectionist model holds that an attitude (or any other mental representation) is a functional state, not a “thing” that is “stored” statically in memory (Smith, 1998). Specifically, a representation is an attractor state of a distributed connectionist network, which can take on patterns of activation that amount to context-sensitive reconstructions of concepts that have been encountered in the past in conjunction with the currently present cues. Concretely, this means that the reconstructed attributes and the evaluation of a given concept (such as “dog”) that becomes active on one occasion will differ from that on another occasion, depending on contextual cues that may include external features of the situation, concepts that were active in the recent past, internal states such as goals or moods, etc.

Existing Attitude Models Involving Connectionist Principles

Earlier we elaborated on the fundamental distinction between localist and distributed models, both of which can be termed connectionist. Virtually all existing connectionist models of representation in social psychology use localist assumptions (e.g., Read & Miller, 2002; Van Overwalle & Jordens, 2002). This is clear in graphical presentations of the models, when theorists label each node with a specific symbolic concept. Furthermore, most such models are aimed at person perception, stereotyping, or other representational issues rather than at attitude representation per se.

The one existing model that uses distributed assumptions akin to ours is by Bassili and Brown (2005). The authors assume that presentation of an attitude object elicits the activation of a number of “microconcepts.” These are defined as “molecular elements of knowledge that yield meaning when assembled into networks with other microconcepts” (2005, p. 552). This definition seems to qualify the model’s representational assumption
as distributed, for a microconcept does not involve a specific, fixed meaning on its own
as a node in a list representation does. Instead, a configuration of microconcepts (i.e.,
a distributed pattern of activation) is said to elicit an evaluation – the attitude toward the
stimulus object. As we do, Bassili and Brown (2005) emphasize context sensitivity and
flexibility, the idea that patterns of microconcept activation are fluid and that
microconcepts can be recruited into many different combinations resulting in different
evaluations, depending on context. Thus, we generally share their assumptions on the
underlying representational issues.

Our approach differs from that of Bassili and Brown in one major way. They
assume that the key difference between explicit and implicit measures is that only the
former involve information entering into working memory and therefore into conscious
awareness (see their Figure 13.1). This implies that people are (by definition) unaware of
their implicit evaluations (see Greenwald & Banaji, 1995). We believe, in contrast, that
people are often aware of their automatic evaluations of objects (Fazio & Olson, 2003),
which may be subjectively experienced as “immediate gut reactions.”

Connectionist Representation and Automaticity

Many aspects of attitudes have received research attention in social psychology in
the last two decades, but none has been as prominent, perhaps, as automaticity. The
connectionist model of representation informs our thinking about automaticity and
attitudes in general, and also about how controlled and automatic processes interact to
construct and reconstruct representations.

Aspects of Automaticity: Efficiency and Spontaneity
Bargh (1994) outlines four features associated with the automatic-controlled distinction: awareness, efficiency, intentionality, and controllability. At its heart, the idea of automaticity implies spontaneity—activation occurs both efficiently (without effort) and uncontrollably. In fact, if representations are states rather than things, activation patterns must be seen as continuously occurring and changing over time. The connectionist representational system is always on; it is constantly entering different representational states in response to all the attended inputs in our internal and external environments. Attractors are spontaneously activated in a connectionist system; they do not require cognitive resources to appear, so in the sense that representations come to mind efficiently and uncontrollably, attitude reconstruction is automatic.

**Other aspects of automaticity**

Other aspects of automaticity, such as intentionality and awareness, are outside the scope of models of attitude representation (our topic in this paper). They depend on the overall architecture of a cognitive system, including assumptions about such things as executive control (needed for the formation and activation of intentions) and the nature and determinants of conscious experience (required for any discussion of awareness). Although distributed connectionist representations can be part of an overall model of cognitive architecture (e.g., Smolensky, 1988), addressing these larger issues fall outside the scope of this paper, which focuses on the distinctive properties of distributed representations and their relations to the properties of attitudes as investigated in social psychology.

“Automatic” and “Controlled” Evaluations
If attitudes are not necessarily automatic in all these senses, the question of what makes attitudes more automatic or controlled naturally arises. Social psychology has placed heavy emphasis on the distinction between automatic and controlled processes to such an extent that some believe that automatic and controlled attitudes tap separate internal representations, “stored” in separate “places.” To draw a distinction between an automatic system and a controlled system that operate independently, however, is to misrepresent the role of control in processing. Control and automaticity do not define processes; they characterize them. Rather than thinking of highly controlled self-report measures as measuring “the explicit representation” and measures tapping more automatic processes as measuring a distinct “implicit representation,” we believe it is more useful to consider how multiple controlled and automatic processes interact to produce responses on both types of measures (Conrey et al., 2005).

Specifically, we have suggested that the production of any behavior, whether a key press on a reaction time measure or a circle on a Likert scale, is the product of a sequence of stages like the following (Smith & Conrey, in press). We claim no novelty for this general set of ideas; rather, we believe that it represents a general consensus among many researchers in the area (e.g., Strack & Deutsch, 2004; Bassili & Brown, 2005; Gawronski & Bodenhausen, in press).

1. To make a judgment, such as an attitude judgment, people start with what is spontaneously activated by the target stimulus in the given context. This is the same material that is assumed to be tapped by so-called “implicit” measures such as evaluative priming or the IAT.
2. If the perceiver has plenty of time, the initial spontaneous representation may evolve based on changing features of the cognitive and social context. New representations may be activated; currently active representations may decay.

3. Based on an intuitive sufficiency criterion, the perceiver may proceed to intentionally activate additional representations that are considered relevant to evaluating the object, even if those are not spontaneously activated.

4. The perceiver may engage in complex attributional inferences or propositional/logical reasoning, e.g., if there are apparent inconsistencies or contradictions among the material that is active.

5. The perceiver’s controlled responses are in line with the final representation. From this perspective, both automatically accessed and explicitly reported (controlled) attitudes are based on representations elicited from the distributed connectionist memory system. The difference is that explicit attitudes might, given awareness, ability, and motivation, also be based on additional representations added in the service of meeting the sufficiency criterion or complex reasoning (e.g., Gawronski & Bodenhausen, in press).

For example, someone may experience the automatic activation of negative affect and negative stereotypic attributes immediately on encountering an image of a racial outgroup member. That activated information might largely control responses on a speeded task such as an IAT. Implicit measures of attitudes limit participants’ time, capacity, and awareness in various ways, increasing participants’ reliance on such spontaneously activated (step 1) representations. In contrast, given more time, the person may intentionally activate additional representations that he/she considers relevant to an
attitude judgment about the outgroup, such as sympathetic feelings about victims of
discrimination or even feelings of collective guilt at the role of his/her own group in
creating and maintaining the outgroup’s disadvantaged situation. The person might use
these intentionally activated representations to control an explicit attitude response. Thus,
the distinction between implicit and explicit measures is not that they access two
completely independent and distinct representations. Nor is it equivalent to a distinction
between processing systems—both types of measures implicate the same systems. The
difference is the extent to which the measures are based mainly on the spontaneous
affective response to a stimulus, versus allowing other representations to be added to the
mix.

Distributed Representations and Other Issues Regarding Attitudes

The most common models of representation in social psychology are localist and
symbolic (Smith, 1998; Barnden, 1995). They assume that each object or category
corresponds to a node that stands for it in memory. There is a node for dogs, for instance,
and another for cheese. To have an attitude about cheese, we associatively link the cheese
node to the “good” node with a given degree of strength. A distributed connectionist
system, on the other hand, assumes that objects or concepts are represented by patterns of
activated nodes. Different activation patterns across the same set of nodes can represent
a dog, cheese, or anything else. Like symbolic systems, distributed, connectionist systems
link nodes together to construct representations, but, unlike localist/symbolic systems, no
individual node (or link) in a distributed network carries any distinct meaning.

Localist/symbolic representational architectures give rise to questions such as “is
there only one attitude, or might there be two?”, “what happens to an old attitude when a
new attitude is formed?”, and “how is ambivalence represented?” The flexibility of the distributed, connectionist framework provides simple answers to some such questions and renders others meaningless. In this section, we illustrate some of the important distinctions between connectionist and localist/symbolic models of representation by addressing issues that concern attitude researchers.

How many attitudes are there?

Implicit attitude measures were first developed at least in part to tap what researchers thought of as the true attitude that participants were reluctant to report or of which they were unaware (e.g., Fazio et al., 1995). More recently, researchers have proposed that there are two true attitudes: old attitudes and new attitudes (Wilson, Lindsey, & Schooler, 2000); or societal attitudes and personally endorsed attitudes (Olson & Fazio, 2004). Within a localist, symbolic representational system, it is meaningful to ask how many different attitudes toward the same object exist, because each attitude requires a different dedicated node. The current and former attitudes (or the societal and endorsed attitudes) toward brie must be represented by distinct links or distinct nodes, which could in principle be separately identified. In contrast, a connectionist system uses the same set of nodes to represent all objects so there is no dedicated spot for storing attributes of brie, or cheddar, or cheese in general. Instead, each of these representations is a slightly modified version of the same pattern of activation. The overlap in patterns used to represent objects means that, in the distributed connectionist system, there could be two answers to the question “how many attitudes toward the same object exist?” Depending on the definition of a true attitude, the
connectionist system can be said to represent no true attitudes, or it can be said to represent an infinite number of true attitudes.

If the “truth” of a “true” attitude lies in its stability and consistency across situations, in the constancy of the strength of the connection between the symbol representing cheese and the symbol representing positivity, then a connectionist representational system does not represent any true attitudes at all. Representations in the connectionist system are different every single time they are activated. Each time a given pattern is activated, all the weights between nodes in the pattern are strengthened by the learning rule. That is, every time I think of how much I love cheese, the system associates love and cheese more strongly. Importantly, however, each time I think of how much I love dogs or how much dogs love cheese, the same nodes are activated in a different pattern, and the same weights are altered in a slightly different way. The same attitude toward cheese, then, is never activated twice because intervening processing, even about unrelated concepts, is constantly changing the attitude toward cheese. In this sense, the system holds no true attitudes.

The other sense in which implicitly measured attitudes have been thought of as “true” is that they are products of the associative system rather than of deliberate machinations on the part of subjects. Researchers have long concerned themselves with the possibility that people might lie when they are asked about sensitive attitudes. Measures of implicit attitudes are widely interpreted as reflecting the attitude free of artificial correction. If the truth of a true attitude is in its being the genuine product of the associative system and not something that the perceiver has made up, then our distributed
representational system holds infinite true attitudes—as many attitudes as there are activation patterns cued by the attitude object, in different contexts or at different times.

From a distributed connectionist perspective it is not particularly useful to count the number of attitudes, or to question the extent to which any of those attitudes are “true”. Instead, the key issue is the extent to which each attitude influences the person’s cognitive and affective reactions as well as overt behavior. Sometimes the spontaneously activated evaluation drives behavior directly, while at other times the attitude produced by careful consideration and adjustment of the spontaneous attitude predicts behavior (Fazio, 1990). After all, if a subject is willing to lie to an experimenter about his attitude toward African-Americans, he or she might equally be willing to perform similar self-presentational behaviors in the world outside the laboratory. We believe that the most fruitful approach for attitude research is not to try to identify which representations are “real” and which are manufactured, but to assume that all representations are manufactured (actually: reconstructed) and to examine the specific processes by which attitude representations are transformed into different types of overt behavior.

**Attitudinal Ambivalence**

A connectionist perspective holds that there is nothing special from a representational viewpoint about an ambivalent response to an object. You might learn that a new type of cheese is both tasty and low-fat – so if both of those attributes are activated when you again encounter the cheese, you experience a clearly positive attitude. If instead the cheese is tasty but also 70% fat, the activation of these attributes (one positive, one negative) could be thought of as an ambivalent attitude. But nothing in the nature of the representation or the activation process itself differs in any meaningful way
between these two possibilities. Ambivalence is an issue not for attitude representation but for response processes farther downstream, which must combine the implications of multiple, potentially conflicting activated representations into a single, coherent behavioral plan (e.g., to approach or avoid an object).

**Attitude Change**

The answers to questions about how many attitudes there are toward the same object shed light on questions about how attitudes are formed and changed. Again, each time any representation is activated in a distributed system, all the weights in the system have the potential to change. Thus, each time I think of my deep and abiding love for cheese, my love deepens. Each time I smell cheese in the real world or taste a delicious slice of cheddar and experience love, my love grows deeper. But it is possible to imagine a situation in which I no longer love cheese—society tells me that cheese is wrong, or I discover that I am lactose intolerant. Anyone who has tried to quit smoking, or drinking, or eating cheese can attest that the original positive attitude never really seems to go away. Certainly for a very long time after I have given up Gouda, the smell of a cheese shop, the sight of good crackers, or an evening spent with my former cheese-buddies can evoke a strongly positive reaction to cheese. And yet, my negative attitude is very real as well. I can still associate the eating of cheese with overwhelming feelings of nausea, for instance. Neither of these in isolation seems to be meaningfully identifiable as my “true” attitude.

Wilson et al. (2000) proposed that these two attitudes—the old one and the new one—are actually represented separately. In a localist/symbolic system, such separate representation would require a node marked “old cheese” that is strongly linked to the
positive node and a node marked “new cheese” that is weakly linked to negativity. In a
distributed connectionist system, attitudes are not represented separately, so it is
technically impossible to have independent old and new attitudes toward the same object
(or any completely independent attitudes at all, for that matter). This is not to say,
however, that changing an attitude removes all trace of an old attitude. A distributed
connectionist system never erases an experience; it simply updates connections with each
activation. Each time I have a negative cheese experience, my cheese association gets
slightly more negative. Importantly, though, representations are context-sensitive; my
representation of cheddar is slightly different from my representation of brie, and my
memory of cheese yesterday is a different representation from my imagination of cheese
tomorrow. That means that, while the old attitude is slowly overwritten by new
experiences, it is possible for me to remember how I used to feel, and, given the right
circumstances, features of the old attitude may guide my behavior.

*Attitudes Versus “Cultural Knowledge”*

Researchers have long drawn distinctions between types of attitudes, often
labeling one type the “true” attitude and contrasting it against another type of evaluation
such as knowledge about the evaluations of others or an older attitude. Some researchers
have suggested that personal evaluations have a special status behaviorally and
correspond better to explicit judgments (Olson & Fazio, 2004). We propose that the only
meaningful criterion for saying that a construct is a “true” attitude is its capability to
drive one’s own reactions and overt behavior. To endorse an attitude or deliberatively
choose to hold it as one’s own implies a commitment to acting consistently with the
evaluation. If I truly like blue cheese, I should be willing to eat it under at least some
circumstances. However, our criterion implies that knowledge about other people’s evaluations might also be a true attitude, because such knowledge can also drive behavior. Knowledge about others’ attitudes has an informational impact: knowing that cheddar is a popular cheese is likely to make me try it even if I’m not sure I like it. And knowledge about how others feel has normative impact; along with one’s attitude, social norm information is the other half of the behavior equation in the Theory of Reasoned Action (Fishbein & Ajzen, 1975). Under many circumstances, particularly when the opportunity to intentionally activate desired representations is absent, knowledge of the attitudes of others is as likely to become automatically activated and to contribute to behavior as is one’s personally endorsed attitude. While personal and social evaluations may be different, a careful analysis of the processes that allow activated evaluative representations to shape behavior is more useful than the simple assumption that the explicitly endorsed attitude is more “real.”

Summary

Distributed connectionist models of representation have distinctive properties compared to more conventional, localist and symbolic models that have predominated in social psychology. The key difference that we have emphasized is context sensitivity: distributed connectionist networks end up reconstructing potentially different versions of a concept on each occasion, including different evaluations. In contrast, if a concept is represented by a single localist node, the meaning of that concept is fixed and static, and theorists must look elsewhere (for example, to external processes the deliberatively shape responses) to account for the observed malleability and flexibility of representations and attitudinal responses. The plausibility of the latter assumption has been weakened in
recent years by several demonstrations that even implicit measures, which are hypothesized to allow relatively little deliberative post-processing, show great malleability (e.g., Blair, 2002).

We are not arguing that distributed connectionist representations are the sole currency in which our mental systems conduct their business. The larger debate within psychology and the cognitive sciences more generally on the respective roles of connectionist and symbolic representations and processes (e.g., Clark, 1993) will not be settled anytime soon, and certainly not by this article. In the long term, probably combinations of principles from both localist and distributed models will be required for a full account of human cognitive capabilities (as suggested by Barnden, 1995 and Thorpe, 1995). Our aim in this paper is limited, to lay out the implications of distributed connectionist representations for current issues regarding attitudes.

Understanding these implications makes clear that the distributed connectionist framework offers a distinctive perspective on representation. In this article, we have suggested that seeing attitudes as occurrent patterns of activation as opposed to static connections between localist nodes informs many of the most important questions in attitude research. From this perspective, some of these questions lose their meanings. The question of how many attitudes can be separately represented, for instance, dissolves because a distributed connectionist system does not support any separate representations; all representations are distributed across the same network of nodes. Other questions have clear answers. What happens to an old attitude when new learning occurs is clear from the assumptions of the connectionist model. On still other questions, this perspective offers tentative though intriguing novel hypotheses, such as the idea that a personal
attitude and a representation of others’ attitudes can both be considered “true” attitudes in the sense of being capable of driving behavior. Reformulating fundamental representational assumptions can help us think more clearly about a variety of issues involving attitudes, automatic and controlled processing, and the generation of overt behavior.
References


