Personality and Social Psychology Review
2005, Vol. 9, No. 3, 231–274

A Connectionist Model of Attitude Formation and Change

Frank Van Overwalle
Department of Psychology
Vrije Universiteit Brussel, Belgium

Frank Siebler
Department of Psychology
University of Bielefeld, Germany

This article discusses a recurrent connectionist network, simulating empirical phenomena usually explained by current dual-process approaches of attitudes, thereby focusing on the processing mechanisms that may underlie both central and peripheral routes of persuasion. Major findings in attitude formation and change involving both processing modes are reviewed and modeled from a connectionist perspective. We use an autoassociative network architecture with a linear activation update and the delta learning algorithm for adjusting the connection weights. The network is applied to well-known experiments involving deliberative attitude formation, as well as the use of heuristics of length, consensus, expertise, and mood. All these empirical phenomena are successfully reproduced in the simulations. Moreover, the proposed model is shown to be consistent with algebraic models of attitude formation (Fishbein & Ajzen, 1975). The discussion centers on how the proposed network model may be used to unite and formalize current ideas and hypotheses on the processes underlying attitude acquisition and how it can be deployed to develop novel hypotheses in the attitude domain.

Evaluations of our environment are a ubiquitous aspect of human life. Attitudes pervade our thinking because they provide valenced summaries of favorable and unfavorable objects and organisms and so serve as a behavioral guide to approach or avoid them. Without such spontaneous guidance by our evaluations, survival in a complex and, sometimes, threatening world would be impossible.

Social psychologists have made substantial progress in the understanding of attitudes. Most definitions proposed in the literature point to the notion that an attitude involves the categorization of an object along an evaluative dimension. In an extensive overview of theorizing and research, Eagly and Chaiken (1993, p. 1) defined an attitude as "a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor." Attitudes are stored in memory, where they persist over time and from where they "become active automatically on the mere presence or mention of the object in the environment" (Bargh, Chaiken, Govender, & Pratto, 1992, p. 893). After being activated, they provide a ready aid for interaction while at the same time freeing the person from deliberative processes. Furthermore, they aid in a coherent interpretation of the environment by biasing our preferences in a congruent manner (Schuette & Fazio, 1995).

How do attitudes reside in memory? Perhaps the most prominent view is that attitudes are stored in memory in the form of object–evaluation associations. Fazio (1990) noted:

> An attitude is viewed as an association in memory between a given object and one's evaluation of that object. This definition implies that the strength of an attitude, like any construct based on associative learning, can vary. That is, the strength of the association between the object and the evaluation can vary. It is this associative strength that is postulated to determine the chronic accessibility of the attitude and, hence, the likelihood that the attitude will be activated automatically when the individual encounters the attitude object. (p. 81)

Empirical tests of this view of attitudes as object—evaluation associations have yielded confirming results. For instance, participants who had been induced to express their attitudes repeatedly, which should

This research was supported by Grant OZR423 of the Vrije Universiteit Brussel to Frank Van Overwalle. We are grateful to Gerd Bohner and Christophe Labiouse for their suggestions on an earlier draft of this article.

Requests for reprints should be sent to Frank Van Overwalle, Department of Psychology, Vrije Universiteit Brussel, Pleinlaan 2, B 1050 Brussel, Belgium. E-mail: Frank.VanOverwalle@vub.ac.be

strengthen the object—evaluation association, have been found to respond relatively quickly to direct inquiries about their attitudes (for an overview, see Fazio, 1990). However, attitudes are more than evaluations. As stated by Chaiken, Duckworth, and Darke (1999),

> attitudes are represented in memory not only as mere object—evaluation linkages [e.g., Fazio, 1990], but also in a more complex, structural form wherein cognitive, affective and behavioral associations also appear as object—association linkages …. When such linkages are many …, or when such linkages are evaluatively consistent …, attitudes are stronger and thus manifest greater persistence, resistance to change and predictability over behavior. (p. 121)

Until now, however, there has been little theoretical advancement in our understanding of the storage and strengthening of attitude–object associations in human memory. We concur with Eiser, Fazio, Stafford, and Prescott (2003) that "attitude theorists … have tended to make relatively little use of paradigms developed in other areas of learning research. …The time is ripe for a renewed analysis of the learning processes underlying the acquisition of attitudes" (pp. 1221–1222). In particular, we will present a recurrent connectionist model (McClelland & Rumelhart, 1985) that describes how attitude associations are developed, strengthened, and maintained in memory. Connectionist approaches have enjoyed an increasing interest in psychology during the last decade, and in particular in social psychology, because they offer a new perspective on diverse social psychological phenomena, including person impression formation (Smith & DeCoster, 1998; Van Overwalle & Labiouse, 2004), causal attribution (Read & Montoya, 1999; Van Overwalle, 1998), group biases (Kashima, Woolcock, & Kashima, 2000; Queller & Smith, 2002; Van Rooy, Van Overwalle, Vanhoomissen, Labiouse, & French, 2003), and many other social judgments (for a review, see Read & Miller, 1998).

There are several important characteristics that make connectionist approaches superior to earlier attitude models (for an accessible introduction to connectionist networks, see McLeod, Plunkett, & Rolls, 1998). First, a key difference is that the connectionist architecture and processing mechanisms are based on analogies with properties of the human brain. This allows a view of the mind as an adaptive learning mechanism that develops accurate mental representations of the world. Learning is modeled as a process of online adaptation of existing knowledge to novel information provided by the environment. Specifically, the network changes the weights of the connections with the attitude object so as to better represent the accumulated history of co-occurrences between objects and their attributes and evaluations. Most traditional algebraic and associative models in social psychology (for an overview, see Fishbein & Ajzen, 1975), in contrast, are incapable of learning. In many algebraic models, attitudes are not stored somewhere in memory so that, in principle, they need to be reconstructed from their constituent components (i.e., attributes) every time an attitude is accessed (but see Anderson, 1971). Earlier associative models proposed in social psychology, can only spread activation along associations but provide no mechanism to update the weights of these associations. This lack of a learning mechanism in earlier models is a significant restriction. In connectionist models, retrieval and judgment is also reconstructive in the sense that activation spreading along the object–valence association is needed to reactivate the evaluation associated with the attitude. However, this involves dramatically less computational steps than retrieving all constituent attributes and their evaluations, and computing some sort of algebraic integration of it (Fishbein & Ajzen, 1975). The present approach is consistent with the idea that strong attitudes are stored in object–valence associations that are easily accessible, whereas weak attitudes are stored in weaker associations and are therefore more susceptible to salient temporary information and context effects. Interestingly, the ability to learn incrementally puts connectionist models in broad agreement with developmental and evolutionary constraints.

Second, connectionist models assume that the development of internal representations and the processing of these representations are done in parallel by simple and highly interconnected units, contrary to traditional models where the processing is inherently sequential. As a result, these systems do not need a central executive, which eliminates the requirement of central and deliberative processing of attitude information. Although many attitude theories assume that simple object associations are learned implicitly through conditioning (e.g., Fishbein & Ajzen, 1975) or heuristic processing (e.g., Chaiken, 1987; Petty & Cacioppo, 1981, 1986), the process by which the different evaluative reactions to a stimulus are integrated in an overall attitude is often left vague and couched in verbal terms only. At most, the outcome of this integration is described in an algebraic formula, without specifying the underlying mental mechanism. Given that a supervisory executive is superfluous in a connectionist approach, this suggests that much of the information processing in attitude formation is often implicit and automatic without recourse to explicit conscious reasoning. This does not, of course, prevent people from being aware of the outcome of these preconscious processes. In addition, based on the principle that activation in a network spreads automatically to related concepts and so influences their processing, connectionist models exhibit emergent properties such as pattern completion and generalization, (for a review, see

Smith, 1996), which are potentially useful mechanisms for an account of the biasing effect of attitudes on the interpretation of the environment.

Finally, connectionist networks have a degree of neurological plausibility that is generally absent in previous algebraic approaches to information integration and storage (e.g., Anderson, 1971, 1981a, 1981b; Fishbein & Ajzen, 1975). They provide insight into lower levels of human mental processes beyond what is immediately perceptible or intuitively plausible although they go not so deep as to describe real neural functioning. Drawing on Marr's (1982) notion of levels of information processing (see also Kashima & Kerekes, 1994), algebraic models are regarded the computational level of human reasoning, which simply describes input–output relationships; connectionist models attempt to mimic psychological processes, and therefore are considered the algorithmic level; and models that describe neural circuitry and processing that implements mental processes are regarded as the implementational level. Thus, although it is true that connectionist models are highly simplified versions of real neural functioning and only describe the algorithmic level of mental thinking, it is commonly assumed that they reveal a number of emergent processing properties that real human brains also exhibit. One of these emergent properties is that there is no clear separation between memory and processing as there is in traditional models. Connectionist models naturally integrate long-term memory (i.e., connection weights) and short-term memory (i.e., internal activation) with outside information (i.e., external activation). In addition, recent concerns of the biological implementation of evaluative reactions have started to emerge (for reviews, see Adolphs & Damasio, 2001; Ito & Cacioppo, 2001; Ochsner & Lieberman, 2001) and this implementational level of analysis will certainly help to improve our understanding of the cognitive and emotional mechanisms underlying attitude formation.

A Connectionist Account of Attitude Processes

The main goal of this article is to take current attitude models couched either in verbal descriptions, such as dual-process models (e.g., Chaiken, 1987; Petty & Cacioppo, 1981, 1986), or in computational formulations, such as algebraic models (e.g., Fishbein & Ajzen, 1975), as a starting point and to develop a connectionist model at the algorithmic level that is consistent with these earlier theories. This endeavor may have significant benefits. First, it may strengthen the foundation of these earlier theories because it provides a lower-level description of some of their major theoretical postulates. By providing a more formal description of these underlying processes, it may perhaps weed out some confusion about the nature of attitude formation. For example, a connectionist approach might underscore the growing realization among social psychologists that many processes in social cognition, and attitude formation in particular, are implicit and nonconscious. The model can, among other things, specify which aspects of information integration might be largely outside awareness and how attitude heuristics (Chaiken, 1987) have an impact on an attitude.

Second, it may provide a theoretical framework that enables us to integrate various theoretical positions that have until now resided more or less alongside each other. By doing so, it may potentially explain a larger set of empirical data than did earlier formal theories of attitude formation (e.g., Fishbein & Ajzen, 1975). In fact, many of the assumptions in our connectionist model are drawn from previous attitude theories. We will highlight the main sources of inspiration of the proposed connectionist model and explain very briefly how these notions are implemented in the model.

The model defines attitudes primarily as object—evaluation associations (Fazio, 1990) and adopts additional cognitive and behavioral components of attitudes (Katz & Stotland, 1959; Rosenberg & Hovland, 1960) as elements that shape the object—evaluation association. In line with Fazio (1990), the model predicts that when people encounter a novel attitude object, they will develop object—evaluation associations in memory in accordance with the information that is currently accessible. Once people are confronted again with the object, their stored evaluation comes to mind automatically and guides behavior and thoughts. Consistent with Anderson's (1971) information integration theory, the attitude is further updated if warranted by the novel information that is provided. In so doing, the connectionist mechanisms specifying the underlying information processes will end up making the same formal predictions as the algebraic model of Fishbein and Ajzen (1975). These connectionist updating mechanisms are almost identical to earlier formal theories of classical conditioning (Rescorla & Wagner, 1972), which have been quite popular in attitude research (Olson & Fazio, 2001; Staats & Staats, 1958).

In addition, in line with dual-process models of attitude (Chaiken, 1987; Chen & Chaiken, 1999; Petty & Cacioppo, 1981, 1986; Petty & Wegener, 1999), the connectionist model draws on the notion of elaboration likelihood (Petty & Cacioppo, 1981, 1986) or depth of processing in assuming that sources of information may vary in strength as well as in the depth by which they are considered and thus receive little or substantial weight. This leads to variation in the strength of the attitudes as well as to different modes of processing that comprise the cornerstone of dual-process models. Dual-process models conceive essentially two routes by which people use accessible information to change their attitudes. When capacity and motivation are relatively high, people are assumed to carefully consider and weight the available information (central or systematic route). When capacity or motivation are low, people process the information more shallowly and rely on simple heuristics or peripheral cues that give rise, automatically, to stored decision rules, such as "experts can be trusted," "majority opinion is correct," and "long messages are valid messages" (peripheral or heuristic route; see Chaiken, 1987; Chen & Chaiken, 1999; Petty & Cacioppo, 1981, 1986; Petty & Wegener, 1999).[1] In the connectionist model, the notion of elaboration likelihood or depth of processing is simulated by changing a single parameter in the network, the general activation (i.e., attention) to persuasive information that is assumed to be generally lower under peripheral than under central processing. Although our connectionist network is based on a single form of knowledge representation, acquisition, and processing, this single parametric change together with the inclusion of prior learning of heuristic knowledge structures allows simulating these major differences between the processing modes.

This article is organized as follows: First, we describe the proposed connectionist model in some detail, giving the precise architecture, the general learning algorithm, and the specific details of how the model processes information. We then present a series of simulations, using the same network architecture applied to a number of significantly different phenomena, including central and heuristic processing in attitude formation. Previous models of attitudes are reviewed and compared with the present approach. Finally, we discuss the limitations of the proposed connectionist model and discuss areas where further theoretical developments are under way or are needed.

A Recurrent Model: Basic Characteristics

For all simulations reported in this article, we will use the same basic network model, namely, the recurrent autoassociator developed by McClelland and Rumelhart (1985). This model has already gained some familiarity among social psychologists and has been used to study person and group impression (Smith & DeCoster, 1998; Van Overwalle & Labiouse, 2004; Van Rooy et al., 2003) and causal attribution (Read & Montoya, 1999). We decided to apply this model to emphasize the theoretical similarities that underlie attitude formation and change with a great variety of other processes in social cognition. In particular, we chose this model because it is able to reproduce a wider range of social cognitive phenomena, including attitude formation, than other connectionist models, like feedforward networks (Van Overwalle & Jordens, 2002; see Read & Montoya, 1999) or constraint satisfaction models (Shultz & Lepper, 1996; Siebler, 2002; for a critique see Van Overwalle, 1998).

The autoassociative network can be distinguished from other connectionist models on the basis of its architecture (how information is represented in the model) and its learning algorithm (how information is processed and consolidated in the model). We will discuss these points in turn.

Architecture

The generic architecture of an autoassociative network is illustrated in Figure 1. Its most salient property is that all nodes are interconnected with all of the other nodes (unlike, for instance, feedforward networks where connections exist in only one direction). Thus, all nodes send out and receive activation. The nodes in the network represent an attitude object as well as various attributes of the object that are linked to favorable or unfavorable valences. In addition, the network includes general features that represent most attitude objects (e.g., the notion that these objects are the focus of someone's evaluation) and contextual variables (e.g., sources) that are important in shaping heuristic processing. The nodes in the network can represent these concepts in basically two ways. In a localist representation, each node represents a single symbolic

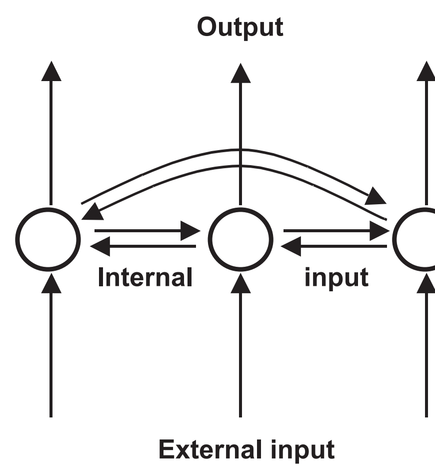

Figure 1. Generic architecture of an autoassociative recurrent network.

[1]In this article, we use the broad terms *central* and *peripheral* (Petty & Cacioppo, 1981, 1986; Petty & Wegener, 1999) to denote the two routes of processing in general, and we reserve the term *heuristic* (Chaiken, 1987; Chen & Chaiken, 1999) for cases where persuasive heuristics are clearly involved in the peripheral process. Although central and systematic are almost synonymous concepts, peripheral is a broader term that includes not only heuristics but also conditioning and other implicit processes.

concept as in earlier automatic spreading activation networks (e.g., Fazio, 1990). In contrast, in a distributed representation, each concept is represented by a pattern of activation across a set of nodes that each represent some subsymbolic microfeature of the conept (Thorpe, 1994). Although a distributed representation is a more realistic neural code, for ease of presentation, we will illustrate the basic workings of the model with a localist representation.

Information Processing

In a recurrent network, processing information takes place in two phases. During the first activation phase, each node in the network receives activation from external sources. Because the nodes are interconnected, this activation is spread throughout the network in proportion to the weights of the connections to the other nodes. The activation coming from the other nodes is called the internal activation (for each node, it is calculated by summing all activations arriving at that node). This activation is further updated during one or more cycles through the network. Together with the external activation, this internal activation determines the final pattern of activation of the nodes, which reflects the short-term memory of the network. Typically, activations and weights have lower and upper bounds of approximately –1 and +1.

In the linear version of the autoassociator that we use here, the final activation is the linear sum of the external and internal activations after two updating cycles through the network. Two cycles are sufficient to spread activation from the attitude object to the attributes and from these to the valences. In nonlinear versions used by other researchers (Read & Montoya, 1999; Smith & DeCoster, 1998), the final activation is determined by a nonlinear combination of external and internal inputs updated during a number of internal cycles (for mathematical details, see Appendix A). During our simulations, however, we found that the linear version with two internal cycles often reproduced the observed data at least as well. Therefore, we used this simpler linear variant of the autoassociator for all the reported simulations.

Memory Storage

After the first activation phase, the recurrent model enters the second learning phase in which the short-term activations are stored in long-term weight changes of the connections. Basically, these weight changes are driven by the difference between the internal activation received from other nodes in the network and the external activation received from outside sources. This difference, also called the *error*, is reduced in proportion to the learning rate that determines how fast the network changes its weights and learns. This error-reducing mechanism is known as the *delta algorithm* (McClelland & Rumelhart, 1988; McLeod et al., 1998; see also Appendix A).

For instance, if the external activation is underestimated because the internal activation suggests that a given valence node should not be activated (e.g., the person did not develop a clear evaluation yet) although the external activation actually does activate it (e.g., the person learns about a favorable aspect of a newly encountered object), the connection weights with the attitude object are increased to reduce this discrepancy. Conversely, if the external activation is overestimated because the internal activation suggests that a given valence node should be activated (e.g., the person is strongly in favor) although the external activation actually does not activate it (e.g., the person learns about an unexpectedly unfavorable aspect of an object), the weights are decreased. These weight changes allow the network to better approximate the external activation and to develop internal representations that accurately describe the environment. Thus, the delta algorithm strives to match the internal predictions of the network as closely as possible to the actual state of the external environment and stores this information in the connection weights.

At this point, it is interesting to note that the delta learning algorithm is formally identical to the Rescorla–Wagner (1972) model of associative conditioning (see Van Overwalle & Van Rooy, 1998, pp. 149–151 for mathematical details). Early attitude theories around 1950 assumed that attitudes are developed through conditional learning and that affective experiences determine the attitude or evaluative response (Olson & Fazio, 2001; Staats & Staats, 1958). According to classical conditioning theory, an attitude is an evaluative response (conditioned response) established by the temporal association of a stimulus (unconditioned stimulus) eliciting an affective reaction with the judgmental target or attitude object (conditioned stimulus). For instance, in one of their first experiments, Staats and Staats (1958) presented Swedish or Dutch names paired with words having a positive (e.g., *pretty*) or negative value (e.g., *failure*). They reported a positive attitude toward names associated with positive words and a negative attitude toward names associated with negative words (see also Zanna, Kiesler, & Pilkonis, 1970).

Hence, by using the delta learning algorithm, the present connectionist model incorporates these earlier conditioning models. This is consistent with the dual-process model of Petty and colleagues (Petty & Cacioppo, 1981, 1986; Petty & Wegener, 1999) although they did not include conditioning processes in their theorizing at such a formal level of analysis as in the present connectionist approach. Hence, an important advantage of our connectionist model using the delta algorithm is that conditioning is an intrinsic part of it, based on the same learning principles. In the next section, we will further describe how these learning principles work.

A Recurrent Implementation of Attitude Formation

To provide some background to our specific implementation of attitude formation, we illustrate its major characteristics with an example that represents processing persuasive information via the central route as outlined in the theory of reasoned action by Fishbein and Ajzen (1975; see also Ajzen, 1991; Ajzen & Madden, 1986).

According to this model, an attitude is a function of

- the expectation or belief that the behavior will lead to a certain consequence or outcome (e.g., using a car is a fast and dry mode of transportation, but also causes air pollution) and
- the person's evaluation of these outcomes (e.g., fast and dry is good, pollution is bad).

According to Fishbein and Ajzen (1975), multiplying the expectancy and value components associated with each outcome and summing up these products determines an attitude (see Appendix B). Many social psychologists have interpreted this summed multiplication as indicating that the integration of this information typically occurs in a conscious, rational, and deliberative manner. For example, Fazio (1990, p. 89) stated that "the Ajzen and Fishbein model is clearly based upon deliberative processing." Furthermore, he argued, "deliberative processing is characterized by considerable cognitive work. It involves the scrutiny of available information and an analysis of positive and negative attributes, of costs and benefits. The specific attributes of the attitude object and the potential consequences of engaging in a particular course of action may be considered and weighted" (p. 88-89).

However, this characterization is not in line with more current views. Although attitude researchers agree that people may pay attention to available information, they do not assume that the process involving the integration of this information is necessarily open to introspection or that this process needs to be repeated once attitudes have been established. As Ajzen (2002) claimed,

> the theory of planned behavior does not propose that individuals review their behavioral, normative, and control beliefs prior to every enactment of a frequently performed behavior. Instead, attitudes and intentions—once formed and well-established—are assumed to be activated automatically and to guide behavior without the necessity of conscious supervision. (p. 108)

Research confirmed that preferences are automatically activated on the mere presence or mention of the attitude object without explicit instruction or environmental cues to evaluate the object (Bargh et al., 1992; Fazio, 1990; Fazio, Sanbonmatsu, Powell, & Kardes, 1986) and that they facilitate decision making (Fazio & Powell, 1997) and attitude-consistent behavior (Fazio, 1990). Nevertheless, current researchers have focused on the indicators and outcomes of attitude processes, leaving unspecified the underlying implicit mechanisms involved in these outcomes. Although these mechanisms were intuitively seen as nonsymbolic and nonconscious, they were presumably not spelled out because researchers lacked the necessary theoretical framework to articulate this process.

Fortunately, connectionism may provide a more appropriate theoretical framework for these implicit processes. As we will argue, it only requires a conscious encoding of attributes or persuasive arguments, whereas the integration of this information and resulting evaluations can occur implicitly by means of connectionist learning principles. We will demonstrate this with the example of a car as a transport vehicle.

Representing Expectations and Values

Figure 2 depicts a recurrent architecture of someone's attitude toward cars as means of transportation that illustrates how Fishbein and Ajzen's (1975) expectancy-value theory is implemented by a connectionist framework. As can be seen, car is the focal attitude object linked by modifiable connection weights to various cognitive and evaluative outcomes. The cognitive attitude component, the belief that the use of a car will result in certain outcomes, can be represented as expectations linking the car with likely outcomes or attributes such as how fast and how polluting a car will be and how dry the trip will be. The likelihood of these outcomes is expressed in the weight of the car→attribute connections, acquired during prior observations. These observations are based on one's own direct experiences as well as on indirect communication or observational modeling (e.g., persuasive messages via the media, witnessing other people's experiences) although indirect information might potentially have less impact. Specifically, as determined by the delta algorithm, the more often a particular consequence or a specific attribute of the attitude object is observed, the stronger the relevant object→attribute connection weight becomes. Conversely, the less often a particular consequence is observed, the weaker this weight will be. Psychologically, this is reflected in an increased or decreased expectation or perceived likelihood that this outcome will occur.

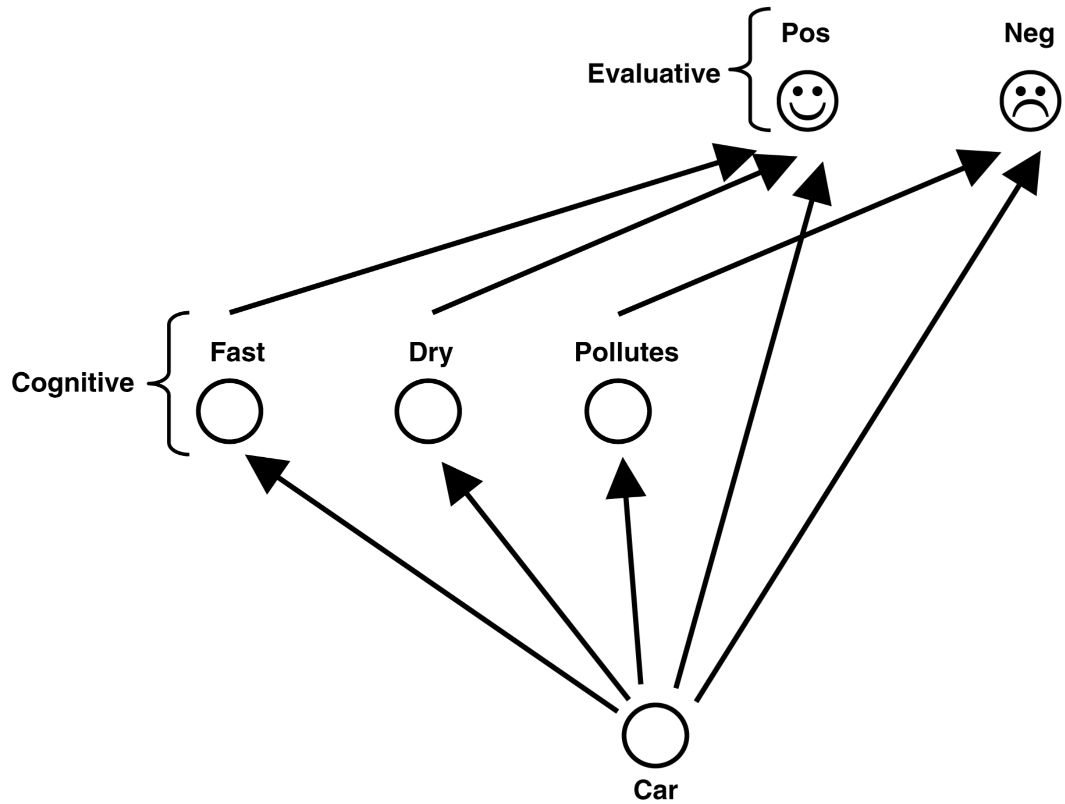

Figure 2. Network architecture with one attitude object (car) connected to three cognitive nodes (fast, dry, and pollutes) and two valence nodes. All nodes are interconnected to all other nodes, but for a clear understanding of the major mechanisms underlying attitude formation, only the most important (feedforward) connections are shown.

The evaluative attitude component can be represented by affective or evaluative responses to the cognitive attributes, such as how much the person likes or dislikes using a fast, dry, and polluting car, again acquired during prior experiences. Thus, in line with Ajzen (1991, p. 191), we assume that the outcomes linked with a behavior or attitude object are "valued positively or negatively." As can be seen, the evaluative responses are represented by two separate unipolar valence nodes, one reflecting a positive evaluation and the other a negative evaluation. We assume that positive and negative evaluations are represented by two separate affective systems because recent neurological evidence suggests that the "neurocognitive system for positive affective associations … serves different functions and can be described without references to neurocognitive systems for negative affect" (Ochsner & Lieberman, 2001, p. 727; see also Canli, Desmond, Zhao, Glover, & Gabrieli, 1998; Ito & Cacioppo, 2001; Lane et al., 1997). These evaluations are expressed in the connections from the cognitive outcomes, or attributes, to the evaluative reactions and are acquired and modified during direct or indirect experiences on the basis of the delta algorithm. Specifically, the more often an evaluation is experienced as a consequence of an attribute, the stronger the relevant attribute→valence connection weight becomes.

In line with Fazio (1990), we suggest that a person's attitude is reflected in the connection between a given object and one's evaluation of that object. Specifically, we define an attitude as the activation of the valence nodes after the attitude object (e.g., car) was activated. This spreading mechanism implements Fazio's (1990) idea that the role of attitudes depends on the extent that "encountering the attitude object [will] automatically activate the evaluation from memory" (pp. 93–94). In this definition, an attitude depends solely on the connections between objects and evaluations that are accessible in memory at the time of judgment. Outcomes reflecting cognitions related to the attitude object (e.g., fast, dry, and polluting attributes) can be computed in the network and "retrieved" later, but they are actually not taken into account for constructing an attitude. This is consistent with the dominant view in the attitude literature that takes attitudes primarily as evaluative responses.

Integrating Expectations and Values: The Attitude

How are the cognitive and evaluative components integrated to create a novel attitude? The core idea of the expectancy-value model is that any object is associated with an evaluation, so that when some objects form a (first-order) connection with an evaluation, these connections mediate the formation of second-order connections and this is repeated for higher-order connections. Thus, what we have called attribute→valence connections in the previous section are basically also object→valence connections. From this perspective, the attribute→valence (e.g., fast→positive) connections are first-order connections that mediate the formation of second-order object→valence (e.g., car→positive) connections. This will, of course, go on recursively so that these newly formed object→valence connections will mediate the formation of still further higher-order connections. This expectancy-value logic has been adopted in the implementation of our connectionist framework and has several consequences.

One implication is that the evaluative reactions to some attributes are learned relatively early in human life (first-order), whereas others are learned relatively late (second-order). We presume that what is learned relatively early are evaluative responses to attributes such as dry, fast, and polluting because these are consequences of direct experiences as well as responses to substantive arguments. These arguments often rely on simple persuasive phrases such as "improved," "better," "advantageous," "do's and don'ts," and so on that children and adolescents are repeatedly exposed to via advertisements, school, and family. In contrast, people are continuously faced with new objects—products and social agents—so that these are unfamiliar and their constituent attributes are learned only much later. Consequently, for the model depicted in Figure 2, we assume that the attribute→valence connections are typically developed early and constitute first-order connections, whereas the object→attribute connections are developed relatively later.

Now comes the integration of cognitive and evaluative components (see Figure 2). In the connectionist model we implemented two activation spreading cycles, so that activation sent out by the attitude object to the attributes (along the object→attribute connections) is further spread to the evaluative reactions (along the attribute→valence connections). These connections were shaped during earlier learning, as explained previously, and may be further updated given novel information. However, what is most crucial is that this activation spreading leads to the co-activation of the attitude object and valence nodes, resulting in the development of novel second-order connections from the object to the evaluative responses. These second-order object→valence connections reflect the formation of a novel attitude.

It is interesting to note that our definition of an attitude is mathematically very close to the multiplicative function of expectations and values in Fishbein and Ajzen's (1975) theory of reasoned action. To see this, replace "expectations" in their model by object→attribute connection weights and "values" by attribute→valence connection weights. If the activation from the attitude object is turned on, it spreads to the valence nodes in proportion to the combined weight of these two connections. Mathematically, this is accomplished by multiplying these connection weights, which is very similar to Fishbein and Ajzen's proposed algorithm (see Appendix B for a formal proof).

A second implication of the recursive expectancy-value mechanism is that we augmented the standard recurrent approach with additional features to enable the network to generate and use evaluative reactions in the formation of higher-order connections. In particular, after a standard network has learned the attribute→valence connections, it cannot use these first-order connections to generate these same evaluative reactions again by means of spreading of activation. This is because in the absence of an explicit coding of the valenced outcome, the network will recognize the internal activation spread to the valence nodes as mere internal predictions by the system. This will cause a decrease in the existing connections (e.g., the delta algorithm interprets the absence of an explicit evaluation at a valence node while receiving high internal activation, as an error of overestimation, to which it reacts by reducing the connection weight). One way to overcome this limitation of standard recurrent networks is by coding the evaluations generated by automatic spreading of internal activation as genuine or external input (denoted by "*i*" to represent "internal input," see Table 2). However, this does not yet allow second-order learning as there is no error in the learning algorithm because the internal and external activation match. To allow the development of second-order connections, we created error by "boosting" these external activations beyond the internal activation. Specifically, we forced them to approach the extremes of –1 and +1, using a standard nonlinear updating algorithm with 10 internal cycles and decay = .15 (see Appendix A for more mathematical details). In effect, these two mechanisms give the valence nodes a special status, as compared to all other nodes. This can be seen as a limitation of the current implementation. On the other hand, with evaluation being a ubiquitous aspect of everyday life, people may actually have learned to use their evaluative responses in a somewhat different way than other information (for a similar argument, see De Houwer, Thomas, & Baeyens, 2001). If so, then our connectionist implementation would model a psychological tendency to rely particularly strongly on one's evaluation as a basis for judgment and learning.

Developing Expectations and Values

Before moving on, we will illustrate how connection weights with attribute nodes (or expectations) and with valence node (or values) are developed through experiences with the attitude object. We will illustrate this with a small simulation example. To make the understanding of this example as easy as possible, the reader should focus mainly on the most important upward connections drawn in Figure 2 although all interconnections between all nodes play a role in an autoassociative network.

In general, according to the delta algorithm, the more often an attitude object is experienced with the same pleasant or unpleasant evaluations, the stronger the connection between the corresponding (un)favorable valence nodes and the attitude object node becomes. This illustrates an important property of the delta learning algorithm, namely that as more confirmatory information is received, the connections gradually grow in strength. We call this the *acquisition property* (Van Overwalle & Labiouse, 2004; Van Rooy et al., 2003).

The sensitivity to sample size of the delta algorithm has been exploited in the earlier associative learning models that preceded connectionism, such as the popular Rescorla–Wagner (1972) model of animal conditioning and human contingency judgments. As noted earlier, conditioning has often been used in past attitude research to explain attitude change (e.g., Berkowitz & Knurek, 1969; Staats & Staats, 1958; for a review see Petty & Cacioppo, 1981). Consistent with the sample size prediction, it has been found that the more often an initially neutral cue (i.e., conditioned stimulus) is paired with another stimulus that strongly evokes an evaluative response (i.e., unconditioned stimulus), the stronger the cue value association becomes, resulting in more vigorous evaluative or affective responses when the cue is present.

Sample size effects have also been documented in many areas of social cognition. For instance, when receiving more supportive information, people tend to agree more with persuasive messages (Eagly & Chaiken, 1993), to hold more extreme impressions about other persons (Anderson, 1967, 1981a), to make more extreme causal judgments (Baker, Berbier, & ValléeTourangeau, 1989; Försterling, 1989; Shanks, 1985, 1987; Shanks, Lopez, Darby, & Dickinson, 1996), to make more polarized group decisions (Ebbesen & Bowers, 1974; Fiedler, 1996;), to endorse more firmly a hypothesis (Fiedler, Walther, & Nickel, 1999), and to make more extreme predictions (Manis, Dovalina, Avis, & Cardoze, 1980).

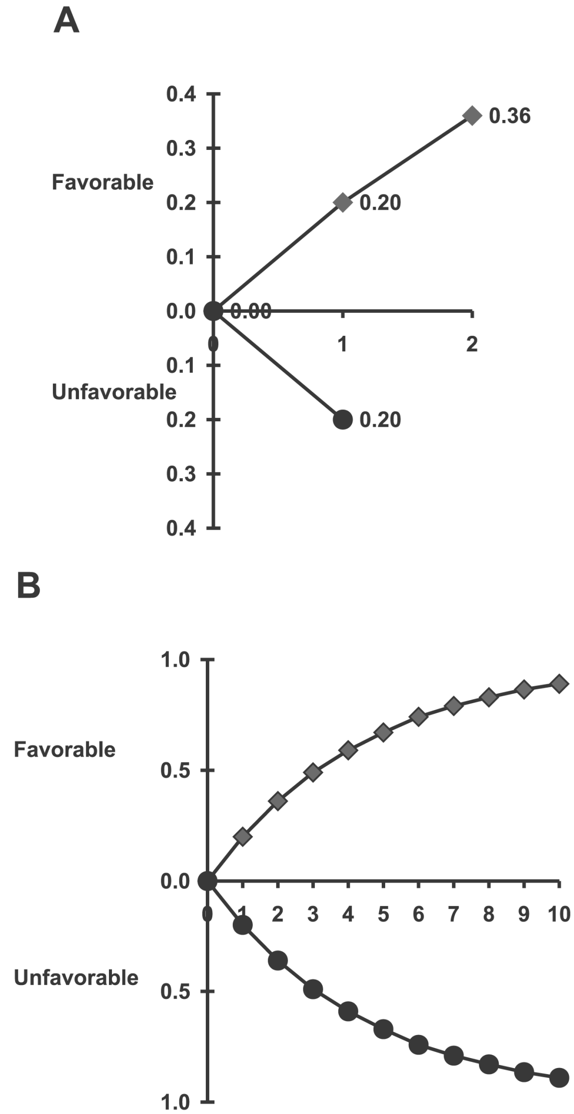

Figure 3. Graphical illustration of the principle of acquisition with learning rate 0.20. The *Y* axis represents the weight of the connection linking the attitude node with the favorable and unfavorable valence nodes. (A) Weights after two favorable and one unfavorable experiences. (B) Favorable and unfavorable weights growing to asymptote after multiple experiences.

Figure 3 depicts an idealized example of the acquisition process in attitude formation. Consider first the simplest case in which a person experiences once each cognitive consequence, or attribute, of car driving. First, activation is spread to the attribute of pollution (e.g., the person realizes that cars pollute the air). This activation is then spread to the unfavorable valence node where it generates a negative evaluative reaction (e.g., the person feels uncomfortable about using a polluting car). For simplicity, let's assume that the unfavorable node is activated to its maximum value of +1, but the favorable node remains inactive. The concurrent activation of the car and the unfavorable reaction leads to an increase of the connection between these two nodes. With a learning rate of .20, the connection weight increases to .20 (see bottom half of Figure 3A). Similarly, following the same mechanism as described earlier, for each of the two positive consequences (i.e., fast and dry), the favorable valence node is activated (e.g., the person feels good about this), resulting in an increase of the connection between the car and the favorable evaluation. Because this increase occurs two times, the resulting weight is .36 given the same learning rate (see top half of Figure 3A). Taken together, the weights of the favorable evaluations exceed those of the unfavorable evaluations. When testing for the attitude response, that is, after activating the car attitude object, this results in a stronger activation of the favorable node than the unfavorable node or a differential activation of .16, leading to a positive attitude in favor of car driving.

In general, however, an attitude depends not only on the direction of the evaluative responses for each of the cognitive outcomes as illustrated in the preceding example, but also on the perceived likelihood of these consequences. As noted earlier, the perceived likelihood is represented in the connection weights of the attitude object and the cognitive consequences, or attributes. Each of these weights, as well as the weights of the attitude object with the evaluative responses, increases as a consequence of the number of experiences with each outcome. For example, if we repeated the favorable information of the previous example, this would result in an increase of connection weight (see Figure 3B, top). After many positive experiences, the acquisition property of the delta algorithm dictates that the weight of the favorable valence node becomes much stronger than that of the unfavorable node, resulting in an overall positive attitude in favor of car driving. Conversely, imagine that the unfavorable consequences of pollution are experienced more often because of recent media coverage. By the property of acquisition of the delta algorithm, this should lead to an increase of the connection weights with the unfavrable valence node as illustrated in Figure 3B (bottom).

It is important to note that according to the delta algoithm when getting closer to the external environment, the learning error shrinks and learning slows down, resulting in a negatively accelerating learning curve. Thus, acquisition is fast and steep at the beginning but then gradually gets slower and flat toward an asymptote (+1 or –1 in this example, see Figure 3B). Stated more generally, during the first phases of learning, the connection weights reflect the amount of evidence, that is, the network is sensitive to sample size. However, the error decreases as more information is processed in the network, so that after some time, learning reaches asymptote and the weight of the connections reflects the average of the favorable versus unfavorable evidence.

In sum, the delta learning algorithm shapes the connections between attitude object nodes with the cognitive and evaluative responses. The acquisition property describes how the connections from the attitude object grow stronger in function of a growing sample size, and so result in the preponderance of favorable or unfavorable evaluations in proportion to the number of experiences with each of these consequences.

General Methodology of the Simulations

We basically used the same methodology throughout all simulations. We applied the connectionist processing principles, including the property of acquisition and sample size, to a number of classic findings in the attitude literature. For explanatory purposes, we most often replicated a well-known representative experiment that illustrates a particular phenomenon although we also simulated a theoretical prediction. Table 1 lists the topics of the simulations we will report shortly, the relevant empirical study or theory that we attempted to replicate, and the major underlying processing principle responsible for reproducing the data. Although not all relevant data in the vast attitude literature can be addressed in a single article, we are confident that we have included some of the most relevant phenomena from the current literature of dual-process models.

We first describe the successive learning phases in the simulations, the general parameters of the model, how cognitions and evaluations were coded, how often they were presented to the network, and how attitudes and attribute-relevant thoughts were measured.

Table 1. *Overview of the Simulations*

Number	Topic	Major Processing Principle	Empirical Evidence / Theoretical Prediction
1	Reasoned Action	Acquisition of valued information	Fishbein and Ajzen, 1975
2	Length Heuristic	Prior acquisition of few versus many values	Petty and Cacioppo, 1984
3	Consensus Heuristic	Prior acquisition of few versus many values	Maheswaran and Chaiken, 1991, exp. 1
4	Expertise Heuristic	Prior acquisition of high (expert) versus low (non-expert) values	Chaiken and Maheswaran, 1994
5	Mood Heuristic	Prior acquisition of high (positive mood) versus mixed (neutral mood) values	Petty, Schumann, Richman, and Strathman, 1993, exp. 2
6	Ease of Retrieval	Competition with valences acquired earlier	Tormala, Petty, and Briñol, 2002, exp. 2

Learning Phases

In all simulations, we assumed that participants brought with them learning experiences taking place before the experiment. We argued before that evaluative responses to attributes and outcomes typically develop early. This was simulated by inserting a Prior Valence Learning phase during which the connections between the object's attributes and their evaluations were developed. When appropriate, we also inserted a Prior Heuristic Learning phase during which connections were established between attitude objects and the heuristic cues of the environment in which they were generated. These two prior learning phases are based on earlier direct experiences or observations of similar situations or indirect experiences through communication or observation of others' experiences. Thus, the connection weights established during these pre-experimental phases reflect the beliefs and evaluations that participants bring with them into the experimental situation.

We then simulated specific experiments. The particular conditions and trial orders of the focused experiments were reproduced as faithfully as possible although minor changes were introduced to simplify the presentation (e.g., fewer trials or arguments than in the actual experiments). Nevertheless, the major results hold across a wide range of stimulus distributions.

Model Parameters

For all simulations, we used the linear autoassociative recurrent network described earlier, with parameters for decay and excitation (for internal and external activation) all set to 1 and with two internal activation cycles. This means that activation is propagated to neighboring nodes and cycled two times through the system, so that nodes linked via one or two connections receive activation from an external source. The activation of a node is computed as the linear sum of all internal and external activations received by this node (McClelland & Rumelhart, 1988; McLeod et al., 1998; for technical details, see Appendix A). Assuming that the major experiments to be simulated used very similar stimulus materials, measures, and procedures, the general learning rate that determines the speed by which the weights of the connections are allowed to change was set to 0.35. This learning rate was chosen because it accommodated all simulations to be reported, although any value between 0.33 and 0.37 typically yielded the same general pattern. All connection weights were initialized at zero. To ensure that prior learning would not overshadow the learning of the experimental information, external activation during prior learning was set to 0.5 or –0.5 instead of the standard level of +1 or –1, whereas learning during the experimental phase was set to the standard level.

Trial Frequencies

For simulating prior valance learning, in all simulations we first ran a number of trials in which positive and negative attributes, or strong and weak arguments, were paired with the favorable and unfavorable valence nodes, respectively. The rationale for this is that strong arguments elicit primarily favorable evaluations about the attitude object, whereas weak arguments elicit primarily unfavorable evaluations (see Petty & Cacioppo, 1984, p. 73). The number of trials for each attribute or argument was set to 15 to ensure that the connection weights approached asymptote. For simulating prior heuristic learning, we first ran a number of pre-experimental trials that varied between 1 and 12 (to be discussed later). These pre-experimental phases and one condition in the experimental phase were run till completion by going once through all trials. To generalize across a range of presentation orders, each network run was repeated for 50 different random orders, thus simulating 50 different participants. Because of the random ordering of trials, the results for each run (or participant) were somewhat different, reflecting the variable conditions of human perception in the actual experiments.

Given that a critical experimental manipulation usually lasts about 1 min (the time to read the information), it seemed reasonable to assume that participants would think at least once about each piece of information and that this would generate their evaluations about it. This was implemented by using one trial for each attribute presented in the experimental condition. It is important to note that the frequencies in the experimental phase were intentionally kept low for two reasons. First, it seemed to us that individuals do not exert extreme effort in thinking about attributes or in interpreting substantive arguments, so that a single trial for each attribute or argument seemed appropriate. Second, having a limited number of trials avoids the destruction of the connection weights learned previously—a result that is known as catastrophic interference (French, 1997; McCloskey & Cohen, 1989). It is implausible that a novel piece of information would totally reverse long-term background knowledge, and this indeed never occurred in the simulations (more on this later).

Measuring Attitudes and Relevant Thoughts

At the end of each simulated experimental condition, test trials were run in which certain nodes of interest were turned on and the resulting activation in other nodes was recorded to evaluate our predictions or to compare with observed experimental data. For measuring the attitude, we turned on the attitude object and recorded the resulting activation at the favorable and unfavorable valence nodes. The unfavorable activation was subtracted from the favorable activation to arrive at an overall attitude measure. For measuring attribute-relevant thoughts, we also activated the attitude object and recorded not only the differential activation of the valence nodes, but also the activation of the nodes reflecting the attributes or arguments presented. Hence, this measure reflects a combination of valences and thoughts about the object's attributes, which seems most appropriate to measure valenced thought or the degree of positively versus negatively valenced thinking. This will be explained in more detail for each simulation. These obtained test activations were averaged across all participants and then projected onto the observed data using linear regression (with intercept and a positive slope) to visually demonstrate the fit between the simulations and experimental data because only the pattern of test activations is of interest, not the exact values.

Central Processing

Dual-process models of persuasive communication assume that given sufficient motivation and capacity, an audience will process incoming arguments extensively via the central or systematic route of persuasion. Perhaps the most influential model of attitude formation that describes this sort of deliberative weighting of all salient alternatives and consequences is Fishbein and Ajzen's (1975) theory of reasoned action (see also Ajzen, 1991; Ajzen & Madden, 1986). As noted earlier, according to this theory, an attitude is a function of the expectation that the behavior will lead to certain consequences or outcomes (e.g., a car is fast and dry, but also pollutes the air), and the person's evaluation of these outcomes (e.g., fast and dry is good, pollution is bad). The attitude is the outcome of this weighting process and is computed by multiplying the expectancy and value components associated with each outcome and summing up these products. This formula of attitude formation has received considerable empirical support in many studies (see Ajzen, 1991; Ajzen & Madden, 1986; Fishbein & Ajzen, 1975) although it has been found that other factors besides attitudes may also exert an influence on behavior. However, a limitation of the theory is that it remained vague about the underlying integration process.

Current views on the attitude process assume that many aspects of deliberative attitude formation and change through the central route to persuasion is implicit. For instance, Chen and Chaiken (1999) claimed that "although perceivers are clearly aware when they are systematically processing information, they are by no means necessarily aware of the precise form of this processing or of the factors that may influence it" (p. 86). Our connectionist model is consistent with this position and assumes that perceivers must be minimally aware only of the information provided at the time of encoding. In contrast, the generation of evaluative reactions to this information and the integration of these evaluations depend on automatic spreading of activation and weight updating, which proceed at an automatic and implicit level as we have seen earlier.

Recent evidence supports this view (e.g., Betsch, Plessner, Schwieren, & Gütig, 2001; Lieberman, Ochsner, Gilbert, & Schacter, 2001; Olson & Fazio, 2001). A very convincing neuropsychological study by Lieberman et al. (2001) documented that amnesic patients could form and change their attitudes for various pictures at different moments in time, despite a severe impairment in their ability to consciously remember the pictures they had encountered earlier, in contrast to control participants who were able to recollect their earlier preferences. Thus, although the amnesic patients might have been aware of the pictures at the time of exposure and probably also of their preferences for some pictures, the online evaluative integration over time of these preferences occurred largely outside awareness.

Simulation 1: Central Processing of Expectations and Valences

We will demonstrate the integration of persuasive information using the central route by mimicking the predictions of the theory of reasoned action (Fishbein & Ajzen, 1975). As assumed by Ajzen (1988), this integration will follow "reasonably from the beliefs people hold about the object of the attitude" so that "we learn to like objects we believe have largely desirable characteristics, and we form unfavorable attitudes toward objects we associate with mostly undesirable characteristics" (p. 32). To illustrate this "reasoned" integration of persuasive information, we extend the car example used in the introduction with other transportation modes such as public buses or bicycles. Table 2 lists a simplified simulated learning history of this example.

Simulation. Each line in the top panel of Table 2 represents a pattern of external activation at a trial that corresponds to either a direct personal encounter or an indirect persuasive statement. The first three cells of each line represent the attitude object presented in each trial. The next three cells reflect the attributes paired with an attitude object. The last two cells denote the evaluation of these attributes, which is either favorable or unfavorable. As can be seen, each node was turned on (activation level of +1, 0.5, –0.5, or –1) or turned off (activation level 0).

As argued earlier, the likelihood variable in the Fishbein and Ajzen (1975) formula is determined by the frequency that an attitude object is paired with an attribute, and the evaluation variable is determined by the degree of satisfaction or dissatisfaction experi-

Table 2. *Learning Experiences During Reasoned Behavior (Simulation 1)*

	Objects			Attributes			Valence	
	Car	**Bicycle**	**Bus**	**Fast**	**Dry**	**Pollutes**		
				Prior Valence Learning				
#15	0	0	0	+	0	0	+	0
#15	0	0	0	0	+	0	+	0
#15	0	0	0	0	0	+	0	+
				Car				
#1	1	0	0	1	0	0	*i*	*i*
#1	1	0	0	0	1	0	*i*	*i*
#1	1	0	0	0	0	1	*i*	*i*
				Bicycle				
#1	0	1	0	1	0	0	*i*	*i*
#2	0	1	0	0	–1	0	*i*	*i*
				Bus				
#2	0	0	1	–1	0	0	*i*	*i*
#1	0	0	1	0	1	0	*i*	*i*
#2	0	0	1	0	0	1	*i*	*i*
				Test				
Attitude Toward Car	1	0	0	0	0	0	?	–?
Attitude Toward Bicycle	0	1	0	0	0	0	?	–?
Attitude Toward Bus	0	0	1	0	0	0	?	–?

Note. Schematic version of learning experiences in attitude formation along Fishbein and Ajzen (1975). J = favorable; L = unfavorable; # = frequency of trial, + = external activation of 0.5, *i* = internal activation (generated mainly by the attributes) is taken as external activation. Each of the transportation means was trained separately, and was always preceded by the Prior Valence Learning phase and followed by the Test phase. Trial order was randomized in each phase and condition.

enced when that attribute is present. We further assumed that learning the attribute→valence connections occurs relatively early, whereas the object→attribute connections develop relatively late. Therefore, in the simulation history, first a prior learning phase is inserted, during which valences are developed, and then the main learning phase, in which the attributes of each transportation vehicles are learned.

These learning phases do not only determine the weight of the object→attribute and attribute→valence connections, but they also shape the object→valence connections that reflect the attitude. Thus, an attitude is stored in the network in the connections from each attitude object with the favorable and unfavorable evaluations. Consequently, testing or measuring an attitude in the network is accomplished by activating the attitude object and reading off the resulting activation of the favorable and unfavorable valence nodes (denoted by ? in the bottom panel of Table 2). In particular, we tested the differential activation of the favorable and unfavorable nodes.

We compared the predictions of the recurrent network model with those of the theory of reasoned action using the summed multiplicative equation outlined earlier (Fishbein & Ajzen, 1975; see also Appendix B). For computing this equation, we used the trial frequencies as estimates of the likelihood of outcomes, and we used a value of +1 for favorable attributes and a value of –1 for unfavorable attributes. Assuming that higher frequencies lead to stronger beliefs, we took the raw trial frequencies rather than proportions or probabilities. Hence, the predictions reflect the relative attitude toward each object; and the pattern is identical if taken proportional to the total number of trials.

Results. The statements listed in Table 2 were processed by the network for 50 participants with different random orders. In Figure 4, the simulated values

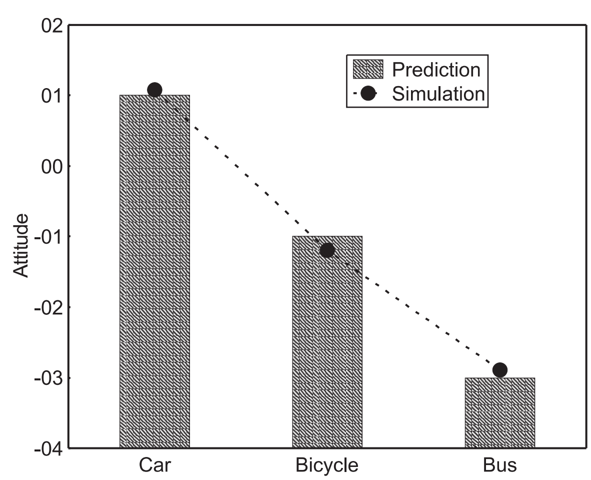

Figure 4. Attitude formation: Predicted data from Fishbein and Ajzen (1975) and simulation results. Theoretically predicted data are denoted by bars, simulated values by broken lines.

(broken lines) are compared to the predictions from the theory of reasoned action (striped bars). As can be seen, the simulated and predicted data match almost perfectly. Furthermore, an ANOVA on the simulated attitude revealed that the main effect of transportation modes was significant, $F(2,147) = 53.72$, $p < .0001$, and further t tests indicated that all transportation modes differed significantly from each other, $t(98) = 3.67 – 10.85$, ps $< .001$.

It is important to note that as soon as the external activation was turned on for the attitude objects and their attributes during the main learning phase, all the remaining mechanisms were implicit as they involved only activation spreading and weight updating. Thus, this simulation presents a formal mechanism that explicates that perceivers need only be aware of the information provided (i.e., the attitude object and its attributes, either perceived directly or communicated through persuasive arguments) and that the rest of deliberative attitude processing may proceed outside awareness.

Heuristic Processing

Although people may prefer to systematically scrutinize all relevant information for forming an opinion about an important issue (Gollwitzer, 1990), in many cases attitudes are created or changed in a more shallow or heuristic manner. This distinction is crucial to dual-process models like the elaboration likelihood model (Petty & Cacioppo, 1981, 1986; Petty & Wegener, 1999) and the heuristic-systematic model (Chaiken, 1980, 1987; Chen & Chaiken, 1999). According to these dual-process models, when motivation or capacity for systematic scrutiny of information is low, such as when the issue is of low personal relevance or when time is limited, people use a heuristic processing strategy. Heuristic processing implies that people form or change their attitudes by using situational cues that automatically give rise to stored decision rules such as “experts can be trusted,” “majority opinion is correct,” and “long messages are valid messages.”

The Nature of Heuristics

What are heuristics and how do they work? According to Chaiken, Liberman, and Eagly (1989, p. 213), “rules or heuristics that define heuristic processing are learned knowledge structures … perceivers sometimes use heuristics in a highly deliberate, self-conscious fashion, but at other times they may use heuristics more spontaneously, with relatively little awareness of having done so.” They argued that heuristics are abstracted on the basis of past experiences and observations or via direct instruction from socializing agents. Consequently, they can vary in their strength or perceived reliability, depending on the statistical relationship between situational cues and agreement with messages during prior learning. As Chaiken et al. (1989, p. 218) put it,

> a person whose past experience with likable and unlikable persons has yielded many confirmations and few disconfirmations of the liking-agreement rule, should perceive a stronger association between the concept of liking and interpersonal agreement than a person whose experience has yielded proportionally more disconfirmations.

Heuristics will only exert an impact on the attitude to the extent that they are reliable and available in memory and to the extent that the situation provides cues that can be processed heuristically.

However, except for the notion that heuristics are knowledge structures reactivated from memory before they can take effect, as far as we know, dual-process theories did not spell out in much detail how these heuristics are applied and integrated into an attitude judgment. Heuristic processing in the attitude literature is often equated with inferential rules, schemas, and procedural knowledge. Hence, one interpretation is that heuristics consist of well-learned abstracted rules like “I agree with people I like” that are applied in some way or another on the current attitude. Another interpretation is that heuristics consist of summarized past knowledge people have about similar situations and the statistical relation between situational cues and agreement with messages. This latter interpretation does not involve the development of abstract, rule-based knowledge and is compatible with a connectionist perspective (see also Smith & DeCoster, 2000; Strack & Deutsch, 2004).

Indeed, numerous simulations in connectionist research have suggested that sensitivity to an abstract rule need not involve the acquisition of a corresponding abstract rule and that a single connectionist network is quite capable of processing both rule-abiding and rule-deviating behaviors and judgments (cf. Pacton, Perruchet, Fayol, & Cleeremans, 2001). Many instances of seemingly rule-like behavior need not necessarily depend on explicit rule knowledge and instead may be based on the processing of exemplars and subsymbolic properties of connectionist models (McLeod et al., 1998). A very well known example is the Rumelhart and McClelland (1986) model of the acquisition of past tense morphology. In that model, not only are regular verbs processed in just the same way as exceptions, but neither are learned through anything like processes of abstract rule acquisition. In the domain of social cognition, Smith and DeCoster (1998) demonstrated that a connectionist network can learn a schema from exposure to exemplars of a category (e.g., learn a stereotype about a social group from exposure to its members) and apply this knowledge to make inferences about unobserved attributes of the category.

Heuristics as Learned Connections

Our core idea on heuristics is that they consist of summarized exemplar knowledge embedded in connection weights reflecting past co-occurrences of heuristic cues and attitude agreement. Hence, there is no storage of explicit abstracted or symbolic rules (see also Smith & DeCoster, 2000). In an early phase, when heuristic learning occurs, the communicated message establishes a cue→valence connection such as that between a likeable source and attitude agreement (Chaiken et al., 1989). This heuristic learning phase can result in different heuristic "rules" by associations of the valences with different cues such as message length, consensus, and expertise and by variations in the association frequencies or valences. In a later application phase, whatever heuristics is operative at the time directs the system at reusing the heuristic knowledge on the basis of the cue→valence connection, with little input from the object→attribute→valence connections (used in central processing). We believe that the preferential use of one of these connections or "routes" depends on differences in attention to either heuristic cue information or substantive arguments (i.e., attributes). We assume that these differences in attentional focus (and thus of heuristic vs. central processing) is governed by the same factors of motivation and capacity that determine elaboration likelihood as proposed by dual-models of attitude (Chen & Chaiken, 1999; Petty & Wegener, 1999).

Let us first elaborate on the prior heuristic learning phase. Our perspective on heuristic learning posits that heuristic knowledge is principally built from earlier persuasive messages in which some arguments drive the valences in a positive or negative direction (e.g., strong arguments producing a favorable valence, weak arguments generating an unfavorable valence). What results from this learning process is an association between the heuristic cue and the valence. The specific arguments in this process are of no further substance and are easily forgotten later because they typically differ between situations and so become random noise that drops out. Thus, what is stored in memory is a direct association between a cue, such as message source, and attitude agreement or disagreement (without arguments) that reflects the statistical relationship during prior learning of messages varying in source expertise and acceptance of opinions. For instance, many confirmations that experts' strong arguments in favor of a certain position leads to attitude agreement will create a strong connection between trustworthy pro-arguing sources and positive valences. Conversely, one can also develop source knowledge indicating that trustworthy experts arguing against a certain position most often leads to attitude disagreement. (For ease of presentation, however, we tacitly assume a favoring expert in most examples.) People can also abstract the functional realm of application of a heuristic cue. For instance, an "expert" source is defined as trustworthy only in a limited range of domains. Thus, a doctor is an expert in diagnosing diseases but not in advising how to enjoy rock-and-roll.

Second, once this cue→valence connection is formed, it resides in memory so that any heuristic processing of novel messages that contains information about the cue (e.g., source) will not start from scratch, but instead will start from a nonzero connection. This nonzero connection will facilitate all further activation spreading and learning on similar issues. For example, an expert's message will be received favorably if processed heuristically because the only thing that matters is the reactivation of the valence associated with the expert cue. This is how our model conceives the spontaneous application of heuristics without depending on symbolic abstraction of rules. Of course, it is possible that under some circumstances "people can reflect on their own past experiences and summarize them, perhaps in the form of a symbolically represented rule" (Smith & DeCoster, 2000, p. 116). In this case, perceivers consciously decide "whether it is appropriate to use activated mental constructs as guides to judgment" (Chen & Chaiken, 1999, p. 83). In sum, we typically see heuristics as implicit knowledge based on past exemplars although they may sometimes be abstracted as symbolic rules. As we will argue later, this view may have important consequences on how heuristics influence attitude formation.

Simulating Heuristic Learning

In the simulations, we implemented heuristic processing in two steps. First, we assumed that alongside attributes of an attitude, the system also stores old experiences with heuristic cues (see Figure 5). Specifically, as shown on the right side of Figure 5, we programmed a pre-experimental heuristic learning phase in which cue→valence connections were built up (via arguments that shape the valence but which are of no further importance and therefore not shown). Second, after this heuristic learning phase, a novel attitude is developed under heuristic processing by a generalization of the cue's valence to the attitude object. With generalization we mean that the prior cue→valence connection creates a similar object→valence connection during the co-occurrence of the cue and the attitude object. This object→valence connection stands on its own. For instance, although the sheer number of favorable reviews in the media shaped our high respect for an artist and created a cue→valence link, because

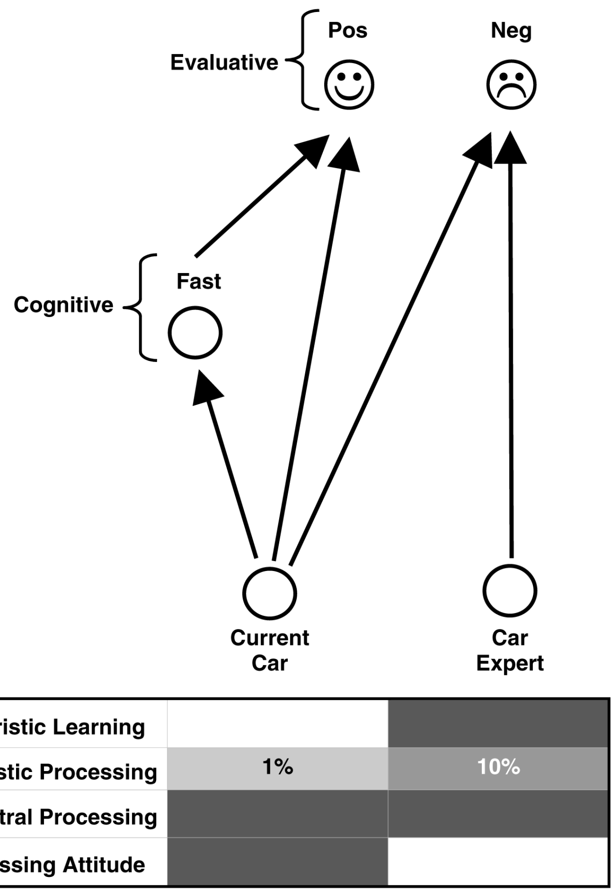

Figure 5. Network architecture with one attitude object (current car) and a heuristic cue (car expert). The example illustrates prior learning when the expert is negatively disposed toward cars. For a clear understanding of the major mechanisms underlying heuristic processing, only one attribute is listed and only the most important connections are shown. The table on the bottom reflects the activation of the network involved in different phases of learning and testing; a lighter shade of gray reflects less activation than the default (activation level indicated by %).

of this generalization to the artist himself we may be able to report later that we evaluated this artist very highly but do not remember that is was because of a great number of flattering reviews.

The fact that the attitude is directly based on an object→valence link without strong supporting links to attributes or environmental cues may explain why attitudes formed by the heuristic route are not very enduring. As soon as a novel generalization is established between the attitude object and another situational cue (or new links with substantive arguments under central processing), the earlier object→valence link changes. For instance, if a friendly expert tells us that she read negative reviews, we might change our mind about this artist more quickly than if we had developed our own arguments under central processing.

We also explored an alternative process in which heuristic processing is not based on a direct generalization of the cue's valence to the attitude object, but rather on an indirect generalization through an object→cue→valence link. This alternative implies that the reactivation of the cue is a crucial mediator in attitude activation. To take our previous artist example, we would be able to report later that the reason we evaluated this artist very highly was the great number of flattering reviews (without remembering the content of it). This approach converges on very similar simulation results. However, because it seems to us that most people tend to forget the heuristic cue under which they developed their attitude, this alternative seemed less intuitively plausible and is therefore not reported. However, future research should establish convincingly that none of the heuristic cues is remembered at a later stage or only very little of it.

Returning to the heuristic learning phase, how often need heuristic events be repeated before the statistical relationship between the cue and message acceptance is stored in memory? For instance, if a perceiver agrees with a message that was advocated by a trustworthy source, how many times should this event (and the opposite event involving an untrustworthy source and disagreement) be repeated before a strong connection is established with the valence nodes? Obviously, perceivers need to be exposed to multiple events or episodes before the heuristic cue is summarized and its application automated. Moreover, there is a growing realization that memory for specific episodic events and the extraction of more generalized statistical knowledge resides in different memory systems in the brain (the hippocampus and neocortex, respectively; see McClelland, McNaughton, & O'Reilly, 1995; Smith & DeCoster, 2000). Prior knowledge—such as schemas and stereotypes—is often built up slowly and is often resistant to change in the presence of new contradictory information. Various modelers have proposed modular connectionist architectures mimicking this dual-memory system with one subsystem dedicated to the rapid learning of unexpected and novel information and the building of episodic memory traces and another subsystem responsible for slow incremental learning of statistical regularities of the environment and gradual consolidation of information learned in the first subsystem (Ans & Rousset, 1997, 2000; French, 1997, 1999; McClelland et al., 1995). Efforts to improve these models are still under way.

Therefore, a full-blown dual-memory system is beyond the scope of this article. However, all these approaches agree that learning in semantic memory is much slower than in episodic memory. Based on this central idea, in the present simulations, we used a single system in which the heuristic episodes were repeated 10 times with a learning rate that was only 10% of the learning rate of the other material. This implementation conserves the basic tenet that learning statistical regularities embedded in earlier messages is much slower than learning a specific (i.e., current) message.

Heuristic Versus Central Processing: Differences in Attention

As noted earlier, to explain the difference between heuristic and central modes of processing, we borrow the elaboration likelihood assumption from dual-process models (Petty & Wegener, 1999), which implies that

> central route attitude changes are those that are based on relatively extensive and effortful information-processing activity, aimed at scrutinizing and uncovering the central merits of the issue or advocacy. Peripheral-route attitude changes are based on a variety of attitude change processes that typically require less cognitive effort. (p. 42)

Thus, during central processing, people assess the relevance and favorability of the persuasive arguments, which requires a lot of mental effort and attention, whereas during heuristic processing, due to lack of cognitive resources or motivation, current persuasive arguments are less well attended to. Because of the shallower encoding of novel persuasive arguments during heuristic processing, they will have less effect on the final attitude than prior heuristic knowledge, so that heuristic knowledge will prevail. In contrast, during central processing, the persuasive arguments are more carefully attended to and scrutinized, so that they largely override prior knowledge and dominate the final attitude.

These differences in attentional focus put our approach in large agreement with dual-process models. Heuristic and central processing differ both quantitatively and qualitatively (Petty & Wegener, 1999). They differ quantitatively because they presume differences in elaboration likelihood of information that is learned by the same fundamental delta algorithm, whether it was learned previously or currently. However, they also differ qualitatively because the information base underlying previous heuristic learning and current central processing differs completely.

Simulating Differences in Attention

We have already seen how the model implements the different learning histories of heuristic and central processing (the qualitative difference). How do we simulate differences in elaboration likelihood (the quantitative difference)? We propose that the degree of elaboration essentially depends on differences in attention to earlier versus recent information and that this determines what type of information will have the most impact on attitude formation. Specifically, we argue that under heuristic processing, there is a general reduction in attention so that it becomes negligible for novel arguments but remains influential for the cue. This remaining attention for the cue allows the generalization of the cue's valence to the attitude. In connectionist models, variation in attention is typically implemented by differences in external activation.

Research has demonstrated that differences in attention can originate from modulations in basic arousal and behavioral activation (e.g., the sleep–wake cycle), responsiveness to salient stimuli, or to task-specific attentional focus and voluntary control of exploring, scanning, and encoding information. Research on the neurological underpinnings of attention suggests that general arousal is driven by lower-level nuclei and pathways from the brain stem, whereas basic features of the stimuli are detected by the thalamus and related subcortical nuclei (e.g., amygdala, basal ganglia). In contrast, task-specific attention and voluntary control are most likely modulated by supervisory executive centers in the prefrontal neocortex (LaBerge, 1997, 2000; Posner, 1992). Variations in elaboration likelihood stem mainly from such conscious control over one's attentional focus. Some connectionist researchers have attempted to model these higher-level voluntary attention processes (e.g., O'Reilly & Munakata, 2000, pp. 305–312, 379–410). The basic idea of their approach is that motivation or task instructions maintain activation in the frontal areas of the brain (through dopamine-based modulation) and that this "attentional" activation spreads to other internal representation in the brain where it results in greater accessibility and activation of other internal representations.

In line with the basic ideas of O'Reilly and Munakata's (2000) model, but somewhat more simplified, we incorporated a general attentional module in our model that served as a gateway to all other areas of our network and that modulated the activation of the nodes in some areas. By varying the attention level in this general attentional module during heuristic processing, all activation levels in some areas were changed to the same degree (see also bottom scheme in Figure 5). Of course, this is again a simplification of real life. It may well be that during heuristic processing, some members of the audience evaluate some arguments with high motivation and attention and then stop upon realizing that it is of little personal relevance. This could be built in the simulations. On average, however, all arguments will be processed less extensively, and therefore we kept the simpler simulation procedure. Note that changes in the activation level do not violate the locality principle of connectionism, which says that each connection weight should be updated using only locally available information from associated nodes. That is because the activation level only affects the general speed of learning, not how much and in which direction weight adaptation should occur, which is uniquely determined by local information according to the delta learning algorithm.

In the next simulations, we implemented four heuristic "rules" of length, consensus, expertise, and mood that were largely based on this assumption although there were some essential differences between each heuristic which prompted us to consider each of them separately. We are focusing here on the effects of heuristics given low and high elaboration conditions and do not address how the heuristics themselves can sometimes determine the extent of thinking (see the section "Quantitative and qualitative processing differences").

Simulation 2: Length Heuristic

We begin the demonstration of our connectionist approach to heuristic reasoning with the length heuristic. In empirical research, lengthy messages are typically manipulated by providing more arguments or by repeating the same arguments in different words with more detail (Haugtvedt, Schumann, Schneier, & Warren, 1994; Petty & Cacioppo, 1984; Schumann, Petty, & Clemons, 1990; Wood, Kallgren, & Preisler, 1985). According to the principle of acquisition, greater sample size of arguments should result in stronger effects on attitude judgments. Thus, the more often an argument that an attitude object possesses a positive or negative attribute is repeated, the stronger the object→valence connection will grow.

Sometimes, people are misled by the apparent length of a message even if it does not include more persuasive arguments but only contains cosmetic alterations such as the use of larger fonts and margins (Wood et al., 1985) or minor rewording (Haugtvedt et al., 1994; but see Schumann et al., 1990). This seems to suggest that superficial characteristics of a message can sometimes influence processing rather than the arguments themselves. A connectionist framework can account for this effect by assuming that such superficial characteristics of length are often correlated with actual differences in message length and so may influence attitudes also.

Key experiment. Petty and Cacioppo (1984) provided a well-known demonstration of the length heu-

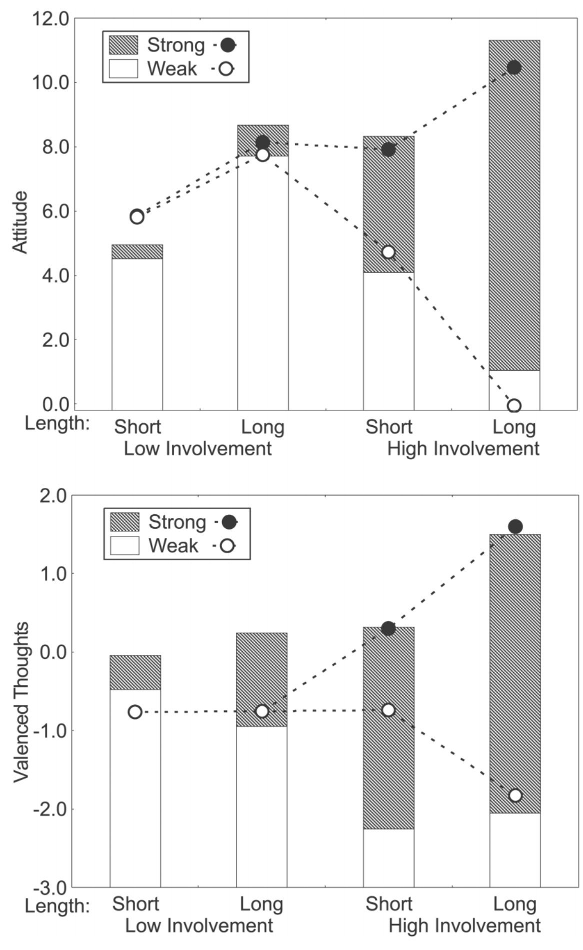

Figure 6. Length heuristic: Observed data from Petty and Cacioppo (1984) and simulation results of attitudes (top panel) and valenced thoughts (bottom panel). Human data are denoted by bars, simulated values by broken lines. The human data are from Table 1 in "The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion" by R. E. Petty & J. T. Cacioppo, 1984, *Journal of Personality and Social Psychology, 46*, p. 75. Copyright 1984 by the American Psychological Association.

ristic. Participants read about a committee that would advise a change in academic examination policy. Involvement in the issue was manipulated by telling participants that the recommendations would be initiated the following year (high involvement) or after 10 years (low involvement). Next, participants read either three or nine arguments in favor of the proposed changes that were all strong or weak. As can be seen in Figure 6 (top panel), under low involvement the number of arguments had a strong impact, suggesting that the length heuristic was applied. In contrast, under high involvement, the quality of the arguments had a greater impact so that nine strong arguments lead to more attitude change than three strong arguments and, similarly, that nine weak arguments lead to less attitude change than three weak arguments.

It is important to note that the length heuristic in the Petty and Cacioppo (1984) study revealed only more agreement with the advocated position although in principle the length heuristic might result in less agreement if weak or unconvincing arguments had been considered. This seems to suggest that under heuristic processing, perceivers primarily reactivate knowledge indicating that lengthy arguments were strong and convincing. This may reflect statistical regularities in perceivers' past experiences, in that among all naturally encountered messages in the past, the longer ones might usually have been the more convincing ones. This idea is reproduced in the simulations by using only strong arguments in the Prior Heuristic Learning phase, and this assumption is crucial in simulating the results from the experimental data.

Dual-process models assume that the cognitive responses while receiving a persuasive message are crucial mediators in forming an attitude. The greater the proportion of favorable responses and the smaller the proportion of unfavorable responses elicited by a message, the greater is the attitude change in the direction advocated. To measure these cognitive responses, Petty and Cacioppo (1984) gave their participants a thought-listing task in which they had to "try to remember the thoughts that crossed your mind while you were reading the material" (p. 74). The results of this thought-listing task, taking into account the favorable or unfavorable valence of the thoughts, are depicted in Figure 6 (bottom panel). Under conditions of high involvement, they show a similar pattern of increased agreement with the message given more arguments, but under conditions of low involvement, they reveal little explicit thinking. These results are consistent with the dual-process model's hypothesis that valenced attribute-relevant thoughts mediate attitude change given central processing but that heuristic processing elicits little thought about the attitude-related arguments.

Table 3. *Learning Experiences and the Length Heuristic (Simulation 2)*

	Object & Cue		**Arguments**[a]						**Valence**	
	Exam	**Length**	**Str1**	**Str2**	**Str3**	**Wk1**	**Wk2**	**Wk3**		
				#10 Prior Heuristic Learning: Short (*Long*) Strong Message						
#1 *(4)*	0	+	+	0	0	0	0	0	+	0
#1 *(4)*	0	+	0	+	0	0	0	0	+	0
#1 *(4)*	0	+	0	0	+	0	0	0	+	0
				Short Strong (*Weak*) Message						
#1	1	1	1(*0*)	0	0	0(*1*)	0	0	*i*	*i*
				Long Strong (*Weak*) Message						
#1	1	1	1(*0*)	0	0	0(*1*)	0	0	*i*	*i*
#1	1	1	0	1(*0*)	0	0	0(*1*)	0	*i*	*i*
#1	1	1	0	0	1(*0*)	0	0	0(*1*)	*i*	*i*
				Test						
Attitude Toward Exam	1	0	0	0	0	0	0	0	?	–?
Valenced Thoughts	1	0	?	?	?	?	?	?	6?	–6?

Note. Simplified version of the experimental design by Petty and Cacioppo (1984). Str = strong, Wk = weak, J = favorable; L = unfavorable; # = frequency of trial or condition, + = external activation of 0.5, *i* = internal activation (generated mainly by the arguments) is taken as external activation. Each experimental condition was run separately, and always preceded by a Prior Valence Learning phase (not shown) and Prior Heuristic Learning phase, followed by the Test phase. Trial order was randomized in each phase and condition. During Prior Valence Learning (not shown), all strong and weak argument nodes were paired with the favorable or unfavorable valence nodes respectively for 15 trials (see also Simulation 1). During Prior Heuristic Learning, each condition was repeated 10 times with 10% of the default learning rate. During heuristic processing of the experimental phase, activation was reduced to 10% for the cue and to 1% for the arguments during acquisition of novel information & testing of attribute-relevant thoughts.

[a]The arguments during prior learning are completely different from those in the experimental and test conditions, but are shown in the same columns to conserve space. The arguments during prior heuristic learning serve to drive the cue's valence into a positive or negative direction, but are of no further importance.

Based on these findings, Petty and Cacioppo (1984) concluded that the number of arguments in a message leads to agreement with a message by serving as a simple heuristic cue when personal involvement was low, whereas it increases issue-relevant thinking when personal involvement was high.

Simulation. Table 3 represents a simplified simulated learning history of the experiment by Petty and Cacioppo (1984). As can be seen, the network consists of a current attitude object (exam) and a contextual cue (message length) and six strong and six weak arguments and two (favorable and unfavorable) valences, each represented by a node. The table lists three strong and weak arguments during prior heuristic learning and three strong and weak arguments during the experimental phase. Note, however, that these arguments are totally different between the two phases but are listed in the same columns to conserve space. As in the first simulation, each line or trial in the top panel of Table 3 corresponds to an argument presented to participants. Because this is the first of four simulations on heuristic processing, we will describe it in somewhat more detail.

First, during the Prior Valence Learning phase, argument quality was paired with the valences. This aspect of the simulation is not listed in the table but is similar to the previous simulation (see top panel of Table 2). Specifically, all six strong arguments and all six weak arguments were paired 15 times with favorable or unfavorable valences, respectively. The rationale for this pairing is that, in most empirical studies, strong arguments are represented by descriptions of attributes that are predominantly superior to alternative attitude objects, whereas weak arguments are represented by predominantly inferior attributes. Consequently, strong arguments elicit primarily favorable thoughts and evaluations about the attitude object, and weak arguments elicit primarily unfavorable thoughts and evaluations (see Petty & Cacioppo, 1984, p. 73). Note that this implementation ignores the fact that arguments may be sensitive to the context or attitude for which they are used (e.g., “improved colors” are crucial for a TV-screen but irrelevant for an answering machine; see also Barden, Maddux, Petty, & Brewer, 2004). This sensitivity could be built in by incorporating a more distributed representation in which the arguments are bound with an attitudinal context or category (by so-called configural nodes) and so develop weak or strong connections with the valence nodes depending on the attitudinal context (also see the section Alternative Implementations).

Second, during the Prior Heuristic Learning phase, the acquisition of the length heuristic was simulated. In particular, to reflect prior learning of short messages, a single trial was presented for each of the three strong arguments, whereas to reflect prior learning of long messages, four trials were presented for each argument or 3 versus 12 arguments overall (we used these frequencies in most of the subsequent simulations). Recall that we used different subsets of arguments for Prior Learning versus the Experimental phase. To simplify the simulation, during Prior Learning, we repeated the arguments from the Prior-Learning subset. This is admissible because the specific arguments themselves do not matter and only serve to activate a positive valence and so increase the connection from the length cue to the positive valence node. As noted earlier, given that research indicates that the length heuristic typically increases endorsement of a message (although in principle it might also decrease endorsement if the audience assumes a weak or unconvincing argumentation), we assumed here that lengthy messages are retrieved mainly from strong and convincing arguments, leading to a link with positive valence. To simulate the idea that multiple heuristic experiences are necessary to detect statistical regularities and consolidate them in long-term memory, the entire heuristic episode of this phase was repeated 10 times with a learning rate that was only 10% of the learning rate for the other material.

Next, during the Experimental phase, the simulation ran through each experimental condition that consisted of a simplified replication of Petty and Cacioppo’s (1984) experiment. In one condition, the attitude object (the exam) was paired with strong arguments that elicited favorable evaluations, whereas in the other condition, the attitude object was paired with weak arguments that elicited unfavorable evaluations. To represent short versus long messages, the simulation was run through either one or three arguments, which reflects the same proportion of arguments as in Petty and Cacioppo’s (1984) empirical study. During the high involvement conditions, central processing was reproduced by setting the activation levels of all nodes in the experimental phase to standard levels. During low involvement, heuristic processing was implemented by setting the activation levels of the heuristic cue to one 10th of these standard levels (= .10), and activation for the arguments was reduced even further one 10th (= .01). Other activation values are also possible, but increasing the activation much above the levels described here may result in an indirect generalization of the cue’s valence rather than a direct generalization, as discussed earlier.

Finally, in the Test phase, measuring the attitude was accomplished in the same manner as the previous simulation (see bottom panel of Table 3). We additionally measured post-message attribute-relevant thoughts, taking into account their favorable or unfavorable valence. Like our measures of attitude, this involves activating the attitude object node and then reading off the activation of the valence nodes and the activation of the cognitive nodes representing the arguments. Specifically, as shown in the last line of Table 3, the activation of the attitude object node was turned on, and the activation of the arguments and valence nodes was measured and then averaged. To balance the output activation of the arguments and valences, the activation of the valences was multiplied by the number of arguments before all output activations were averaged. (In this network model, this procedure is analogous to testing each argument and its associated evaluative activation one after the other, and then averaging the results). During heuristic processing, all testing activation levels of the attribute-relevant thoughts (including those indicated by ?) were reduced to 1% of the standard activation level (in the same manner as for the activation of the arguments in the Experimental phase).

Results. The statements of each condition listed in Table 3 were processed by the network for 50 participants in each condition with different random orders for each phase. Figure 6 depicts the mean test activation for all simulated attitude measures (top panel) and thought measures (bottom panel) projected on top of the empirical data from Petty and Cacioppo (1984). As can be seen, the simulation matched the attitude data reasonably well. Under low involvement, the length heuristic had the strongest impact on the simulated attitude, whereas under high involvement, the quality and number of the arguments had a greater impact. As can be seen, under high involvement, the simulation reproduced a significant increase and decrease of attitude given more strong or weak arguments, respectively. The thought data were also replicated although to a somewhat lesser degree. Under low involvement, there were few simulated thoughts, and under high involvement, thought favorability revealed the same general pattern as the simulated attitudes.

These observations were verified with an ANOVA with three between-subjects factors: Involvement (low and high), Quality of Arguments (strong and weak), and Number of Arguments (few and many). The analysis on the simulated attitudes revealed the expected three-way interaction, $F(1,392) = 1225.20, p < .0001$. Two interactions were of special interest and were also observed in the empirical data (Petty & Cacioppo, 1984). First, there was a significant interaction between Involvement and Number of Arguments, $F(1,392) = 1043.71, p < .0001$. As expected, increasing the number of arguments produced significantly more agreement under low involvement, $t(396) = 5.33, p < .0001$, and less so under high involvement, $t(396) = 2.81, p < .01$. Second, there was a significant interaction between Involvement and Quality of Arguments, $F(1,392) = 4385.29, p < .0001$. As predicted, strong arguments produced significantly more agreement than did weak arguments under high involvement, $t(396) = 29.60, p < .0001$, but not under low involvement, $t < 1$, ns.

The same ANOVA applied to the valenced thoughts revealed the predicted interaction between Involvement and Quality of Arguments, $F(1,392) = 4526.89$, $p < .0001$. Strong arguments generated significantly more thoughts that were consistent with the valence of the arguments than did weak arguments under high involvement, $t(396) = 34.51, p < .0001$, but not under low involvement, $t < 1$.

Simulation 3: Consensus Heuristic

Let us now turn to the consensus heuristic. Dual-process research has documented that under heuristic processing, a higher consensus or majority endorsement of a message implies correctness of that message (Axsom, Yates, & Chaiken, 1987; Darke et al., 1998; Erb, Bohner, Schmälzle, & Rank, 1998; Maheswaran & Chaiken, 1991). This consensus heuristic works in ways very similar to those of the length heuristic. Instead of the sheer number of arguments, here it is the number of people endorsing the arguments that is crucial. In both cases, however, the arguments are more often repeated and therefore influence our attitudes more strongly. Thus, past encounters with many people providing strong arguments in favor of an attitude position led to links of high consensus with positive valence. Conversely, past encounters with many people providing weak arguments in favor of an attitude position (or strong arguments counter to that position) led to links of low consensus with negative valence. In sum, high consensus builds a link with a positive valence, whereas low consensus builds a link with negative valence. In a connectionist framework, this prior consensus knowledge is “retrieved” under heuristic processing and dominates attitude formation because little attention is given the novel information, whereas under central processing, the novel information is fully attended to and tends to overwrite the effects of prior consensus beliefs.

Key experiment. We explore our connectionist approach to the consensus heuristic by simulating a prominent study by Maheswaran and Chaiken (1991, experiment 1). Participants read about a fictitious “XT100” telephone answering machine. Involvement in the issue was manipulated by telling the participants that their reactions would be used to decide whether to distribute the product in their own state (high involvement) or in another state (low involvement). Next, participants read the results of an ostensible marketing test that revealed that either 81% (high consensus) or less than 3% (low consensus) of the consumers were extremely satisfied with the product. Finally, they read a message that described the answering machine as mainly superior to competing brands (strong arguments) or mainly inferior (weak arguments). As can be seen in Figure 7 (top panel), the results confirmed the dual-process predictions. Under low involvement, the proportion of satisfied customers had a strong impact,

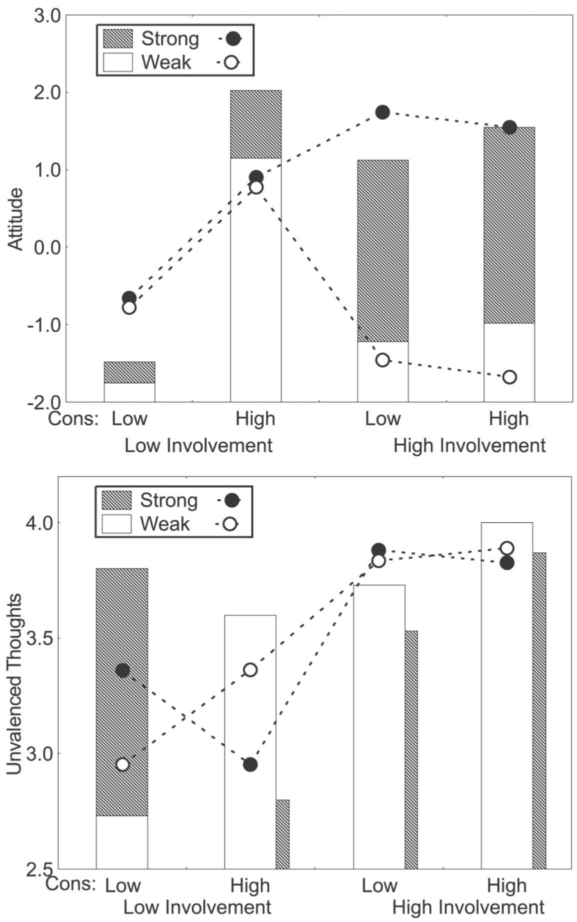

Figure 7. Consensus heuristic: Observed data from Maheswaran and Chaiken (1991, experiment 1) and simulation results of attitudes (top panel) and unvalenced thoughts (bottom panel). Human data are denoted by bars, simulated values by broken lines. The human data are from Table 1 in "Promoting systematic processing in low-motivation settings: Effect of incongruent information on processing and judgment" by D. Maheswaran & S. Chaiken, 1991, *Journal of Personality and Social Psychology, 61*, p. 18. Copyright 1991 by the American Psychological Association.

suggesting that the consensus heuristic was applied. The smaller the proportion of satisfied customers indicating satisfaction with the product, the smaller the agreement with the message, and the greater the proportion, the higher the agreement. In contrast, under high involvement, the quality of the arguments had a greater impact so that strong arguments led to more attitude change than weak arguments.

What about the cognitive responses mediating attitude formation? Unfortunately, Maheswaran and Chaiken (1991) reported only the amount of thoughts without their valence, so that Figure 7 (bottom panel) presents nonvalenced thoughts. They suggested that apart from the influence of the consensus heuristic, incongruency between the majority position and the novel arguments would result in increased central processing by undermining perceivers' confidence in their heuristic-based judgments (see also Erb et al., 1998; Mackie, 1987). That is indeed what they found. As can be seen in Figure 7, the amount of nonvalenced thinking under heuristic processing reveals the expected interaction between congruency and consensus. There was little thinking under low involvement when there was congruency between consensus information and persuasive arguments. In contrast, when there was incongruency, the amount of thinking under low involvement was as high as under high involvement. However, these results are theoretically less informative as they do not reveal to what extent valenced thoughts mediated attitude change although Maheswaran and Chaiken (1991) reported regression analyses revealing a positive relationship that was stronger under high than low involvement.

Although valenced thoughts were not available, we found it interesting to see whether we could simulate nonvalenced thinking instead. To do so, we had to incorporate Maheswaran and Chaiken's (1991) finding that there is increased thinking given incongruency. We presume that the incongruency in the stimulus material alerted the attentional control system so that more activation was devoted to it. Specifically, we implemented the same test procedure as in the previous simulation for the thought-listing task (i.e., with only 1% of the default attention during heuristic processing), with the exception that—in line with Maheswaran and Chaiken (1991)—attention was back at the default level during recall of incongruent thoughts.

Simulation. Table 4 represents a schematic learning history of our simulation of the experiment by Maheswaran and Chaiken (1991, experiment 1). The rationale is similar to that in the simulation of the length heuristic. During the Prior Heuristic Learning phase, to reflect low consensus, four trials were presented for each unfavorable argument, and four trials were presented of favorable arguments to reflect high consensus. These trial frequencies reflect an arbitrary number of other people who expressed their opinions on the issue and were chosen to represent a simple learning history of extreme low consensus (all perceivers disagree) and high consensus (all perceivers agree). As before, this whole phase was repeated 10 times with a learning rate reduced to 10% to mimic consolidation of heuristic information in long-term memory.

Next, the network ran through the Experimental phase. Each argument was presented once, and their quality—strong versus weak—differed according to condition. During the low involvement conditions, heuristic processing was simulated by setting the activation of the cue to 10% of the default and to 1% for the arguments, and during high involvement conditions, central processing was implemented by setting the standard activation levels.

Measuring the attitude and post-message valenced thoughts was accomplished in the same manner as the previous simulation (see bottom panel of Table 4). We additionally measured nonvalenced thought by reading

Table 4. *Learning Experiences and the Consensus Heuristic (Simulation 3)*

	Object & Cue		**Arguments**[a]						**Valence**	
	XT–100	**Consensus**	**Str1**	**Str2**	**Str3**	**Wk1**	**Wk2**	**Wk3**		
				#10 Prior Heuristic Learning: High (*Low*) Consensus						
#4	0	+	+(*0*)	0	0	0(+)	0	0	*i*	*i*
#4	0	+	0	+(*0*)	0	0	0(+)	0	*i*	*i*
#4	0	+	0	0	+(*0*)	0	0	0(+)	*i*	*i*
				Strong (*Weak*) Message						
#1	1	1	1(*0*)	0	0	0(*1*)	0	0	*i*	*i*
#1	1	1	0	1(*0*)	0	0	0(*1*)	0	*i*	*i*
#1	1	1	0	0	1(*0*)	0	0	0(*1*)	*i*	*i*
				Test						
Attitude Toward XT-100	1	0	0	0	0	0	0	0	?	–?
Non-Valenced Thoughts	1	0	?	?	?	?	?	?	6?	6?

Note. Simplified version of the experimental design by Maheswaran and Chaiken (1991, exp. 1). Str = strong, Wk = weak, J = favorable; L = unfavorable; # = frequency of trial or condition, + = external activation of 0.5, *i* = internal activation (generated mainly by the arguments) is taken as external activation. Each experimental condition was run separately, and always preceded by a Prior Valence Learning phase (not shown) and Prior Heuristic Learning phase, followed by the Test phase. Trial order was randomized in each phase and condition. During Prior Valence Learning (not shown), all strong and weak argument nodes were paired with the favorable or unfavorable valence nodes respectively for 15 trials (see also Simulation 1). During Prior Heuristic Learning, each condition was repeated 10 times with 10% of the default learning rate. During heuristic processing of the experimental phase, activation was reduced to 10% for the cue and to 1% for the arguments during acquisition of novel information & testing of attribute-relevant thoughts.

[a]The arguments during prior learning are completely different from those in the experimental and test conditions, but are shown in the same columns to conserve space. The arguments during prior heuristic learning serve to drive the cue's valence into a positive or negative direction, but are of no further importance.

off the total amount of favorable and unfavorable evaluation, instead of the differential activation of the favorable and unfavorable evaluation, using the activation levels as indicated earlier. That is, when there was incongruency between the information implied by the consensus information and the novel arguments, we assumed that the activation levels were restored to the standard levels while measuring these thought (but not when the learning phases were running).

Results. The information listed in Table 4 was processed by the network for 50 participants in each condition with different random orders. Figure 7 depicts the mean test activation for all simulated attitude measures (top panel) and thought measures (bottom panel) on top of the empirical data of Maheswaran and Chaiken (1991). It can be seen that the simulation closely matched the attitude data. Under low involvement, level of consensus had the strongest impact on the simulated attitude, and under high involvement, the quality of the arguments had a greater impact. The simulation also replicated the observed data on the non-valenced attribute-relevant thoughts. Fewer thoughts were found under congruent conditions of low involvement, whereas the largest number of thoughts was observed under incongruent conditions as well as under high involvement.

These observations were tested with an ANOVA with three between-subjects factors: Involvement (low and high), Quality of Arguments (strong and weak), and Consensus (low and high). The analysis on the simulated attitudes revealed two predicted interactions that were also observed in the empirical data (Maheswaran & Chaiken, 1991). First, there was a significant interaction between Involvement and Consensus, $F(1,392) = 2096.25$, $p < .0001$. As expected, increasing the consensus produced significantly more agreement under low involvement, $t(396) = 9.51$, $p < .0001$, but not under high involvement, $t(396) = 1.27$, *ns*. Second, there was a significant interaction between Involvement and Quality of Arguments, $F(1,392) = 6419.85$, $p < .0001$. As predicted, strong arguments produced significantly more agreement than did weak arguments under high involvement, $t(396) = 38.46$, $p < .0001$, but not under low involvement, $t(396) = 1.46$, *ns*.

We applied the same ANOVA on the simulated valenced thoughts, that is, the thoughts including their valence (not in Figure 7). The analysis revealed the predicted interaction between Involvement and Quality of Arguments, $F(1,392) = 9956.85$, $p < .0001$. Strong arguments generated significantly more positive thoughts than weak arguments primarily under high involvement, $t(396) = 68.14$, $p < .0001$, and less so under low involvement, $t(396) = 15.05$, $p < .0001$. In addition, an analysis on the nonvalenced thoughts revealed (see Figure 7), consistent with Maheswaran and Chaiken's (1991) congruency hypothesis, a significant interaction between Consensus and Quality of Arguments under low involvement, $F(1, 392) = 606.19$, $p < .0001$. Congruency between consensus and argument quality led to less nonvalenced thoughts than incongruency, $t(396) = 6.14$, $p < .0001$.

Simulation 4: Expertise Heuristic

Another heuristic cue often explored in the context of dual-process approaches is the expertise heuristic, which says that "experts can be trusted." Again, this heuristic can be viewed as another instance of the acquisition property of the delta learning algorithm, by assuming that the arguments and thoughts compiled from highly regarded experts are more favorable than those compiled from nonexperts. This is consistent with the assumption by Bohner, Ruder, and Erb (2002) that source expertise may lead people to form different expectations about message strength. This can be simulated by a prior heuristic learning phase in which experts or trusted sources are seen as using stronger arguments that elicit more favorable valences than nonexperts or untrustworthy sources.

Several studies revealed different effects of heuristic and central processing given different levels of source credibility, in line with predictions of dual-process models. It was found that source credibility determined attitudes under heuristic processing but not under central processing where argument quality was of major importance (Chaiken & Maheswaran, 1994; Petty, Cacioppo, & Goldman, 1981; Ratneshwar & Chaiken, 1991). Similar results have also been reported for likeable or famous sources (Chaiken, 1980; Petty, Cacioppo, & Schumann, 1983) because such communicators are seen as more expert and trustworthy (Chaiken, 1980; Chaiken & Eagly, 1983).

Key experiment. One of the studies by Chaiken and Maheswaran (1994) is particularly important because it also demonstrated that central and heuristic processing modes are not mutually exclusive. For instance, when the arguments are too ambiguous to form an opinion by extensive processing alone, heuristic cues may additionally help to form an opinion by biasing the selection and interpretation of ambiguous information. This interaction between central and heuristic processing, referred to as the bias hypothesis (Chaiken et al., 1989; Chen & Chaiken, 1999), was investigated by Chaiken and Maheswaran (1994). They presented a message about a fictitious "XT100" answering machine in which different attributes were described. Involvement was manipulated in the typical manner by telling the respondents that their opinion about the answering machine would have little bearing on the manufacturer's decision to distribute the product in another state (low involvement) or would count heavily on the decision to distribute the product in their own state (high involvement). Expertise was manipulated by taking this information ostensibly from a highly regarded magazine specialized in scientific testing of new products (high expertise) or in a promotional pamphlet prepared by sales personnel (low expertise). The message described the answering machine as superior to competing brands on all important attributes (strong arguments), inferior on all important attributes (weak arguments), or superior on some attributes while inferior on others (ambiguous arguments).

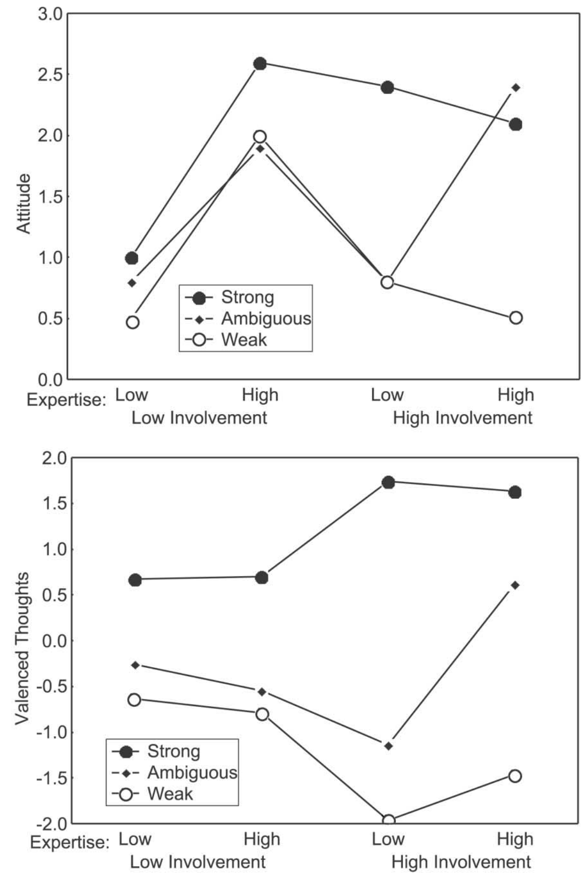

Figure 8. Expertise heuristic: Observed attitude data (top panel) and observed valenced thoughts (bottom panel) from Chaiken and Maheswaran (1994). The human data are from Figure 1 and Table 2, respectively, in "Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment" by S. Chaiken & D. Maheswaran, 1994, *Journal of Personality and Social Psychology, 66*, p. 466. Copyright 1994 by the American Psychological Association. Adapted with permission.

As can be seen in Figure 8 (top panel), consistent with the predictions of Maheswaran and Chaiken (1991), source credibility was the only determinant of people's attitude under low involvement. In contrast, under high involvement, argument quality was the main determining factor, except when the message was ambiguous and source credibility alone influenced the attitude. As might be expected, the valenced thoughts reflected a similar pattern under high involvement and little thought under low involvement (see Figure 8, bottom panel).

Table 5. *Learning Experiences and the Expertise Heuristic (Simulation 4)*

	Object & Cue		Arguments[a]						Valence	
	XT-100	Source	Str1	Str2	Str3	Wk1	Wk2	Wk3		
				#10 Prior Heuristic Learning: High (*Low*) Expertise						
#4	0	+	+(*0*)	0	0	0(+)	0	0	*i*	*i*
#4	0	+	0	+(*0*)	0	0	0(+)	0	*i*	*i*
#4	0	+	0	0	+(*0*)	0	0	0(+)	*i*	*i*
				Strong (*Weak*) Message						
#1	1	1	1(*0*)	0	0	0(*1*)	0	0	*i*	*i*
#1	1	1	0	1(*0*)	0	0	0(*1*)	0	*i*	*i*
#1	1	1	0	0	1(*0*)	0	0	0(*1*)	*i*	*i*
				Ambiguous Message						
#1	1	1	1	0	0	0	0	0	*i*	*i*
#1	1	1	0	0	0	1	0	0	*i*	*i*
				Test						
Attitude Toward XT-100	1	0	0	0	0	0	0	0	?	–?
Valenced Thoughts	1	0	?	?	?	?	?	?	6?	6?

Note. Simplified version of the experimental design by Chaiken and Maheswaran (1994). Str = strong, Wk = weak, J = favorable; L = unfavorable; # = frequency of trial or condition, + = external activation of 0.5, *i* = internal activation (generated mainly by the arguments) is taken as external activation. Each experimental condition was run separately, and always preceded by a Prior Valence Learning phase (not shown) and Prior Heuristic Learning phase, followed by the Test phase. Trial order was randomized in each phase and condition. During Prior Valence Learning (not shown), all strong and weak argument nodes were paired with the favorable or unfavorable valence nodes respectively for 15 trials (see also Simulation 1). During Prior Heuristic Learning, each condition was repeated 10 times with 10% of the default learning rate. During heuristic processing of the experimental phase, activation was reduced to 10% for the cue and to 1% for the arguments during acquisition of novel information & testing of attribute-relevant thoughts.

Simulation. We simulated the biasing nature of heuristic cues on central processing as investigated by Chaiken and Maheswaran (1994). Table 5 presents a simplified learning history with similar network architecture and history as before except for the elements detailed next.

In the Prior Heuristic phase, knowledge about expertise is built up by several experiences of good and bad argumentation by expert and nonexpert sources, respectively. As in the earlier simulations, three arguments were each presented in four trials (or 12 arguments overall), so that strong connections from the expert source to the favorable or unfavorable valences were established for expert and nonexpert sources respectively. This whole phase was repeated 10 times with a learning rate reduced to 10% of the standard rate.

During the Experimental phase, we ran one of three message types, involving strong, weak, and ambiguous arguments. As before, strong messages were represented by three strong arguments associated with a favorable evaluation, whereas weak messages were represented by three weak arguments associated with an unfavorable evaluation. In addition, ambiguous messages were represented by one strong and one weak argument. This reflects—in a simplified manner—the quality and direction of the arguments in Chaiken and Maheswaran's (1994) empirical study. Simulating heuristic versus central processing and measuring the attitude and post-message valenced thoughts was accomplished in the same manner as in the previous simulation (see also Table 5 note).

Results. The statements in each condition listed in Table 5 were processed by the network for 50 participants in each condition with different random orders. Figure 9 depicts the mean test activation for all simulated attitude measures (top panel) and thought measures (bottom panel). When comparing with the empirical results of Chaiken and Maheswaran (1994; see Figure 8), it can be seen that the simulation closely matched the attitude data. Source credibility strongly determined the simulated attitudes under low involvement, whereas under high involvement, argument quality was the main determining factor, except when the message was ambiguous and source credibility alone influenced the attitude. The thought data were also replicated although to a somewhat lesser degree. There were few simulated thoughts under low involvement, and under high involvement, the thoughts revealed the same pattern as the simulated attitudes.

These observations were verified with an ANOVA with three between-subjects factors, Involvement (low and high), Quality of Arguments (strong, ambiguous,

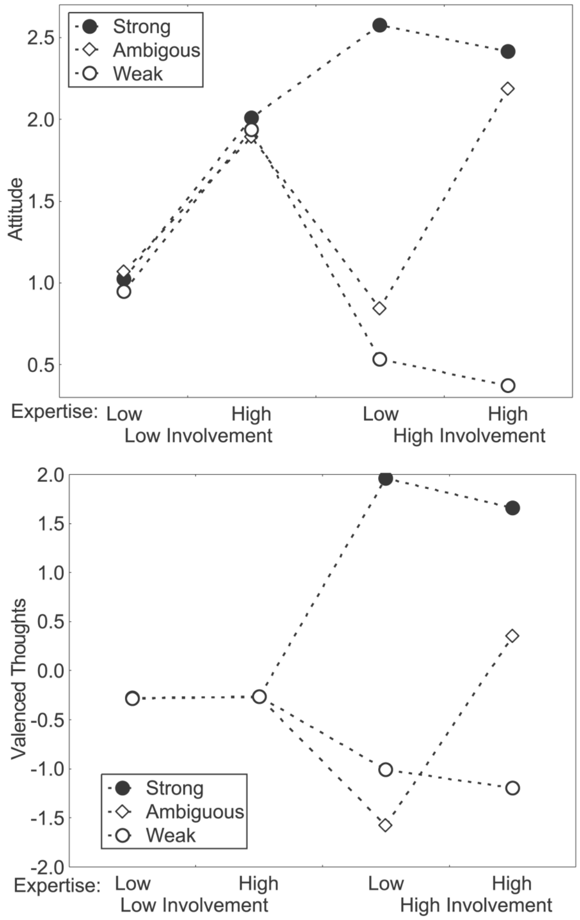

Figure 9. Expertise heuristic: Simulation results of attitudes (top panel) and valenced thoughts (bottom panel) from the simulation of the Chaiken and Maheswaran (1994) study.

and weak) and Expertise (low and high). The analysis on the simulated attitudes revealed the expected three-way interaction, $F(2,588) = 105.48$, $p < .0001$. Two interactions were of special interest and were also observed in the empirical data (Chaiken & Maheswaran, 1994). First, there was a significant interaction between Involvement and Expertise, $F(1,588) = 118.97$, $p < .0001$. As expected, increasing the expertise produced significantly more agreement under low involvement, $t(596) = 11.16$, $p < .0001$, and much less so under high involvement, $t(596) = 4.07$, $p < .0001$. Second, there was a significant interaction between Involvement and Quality of Arguments, $F(2,588) = 438.26$, $p < .0001$. As predicted, strong arguments produced significantly more agreement than did weak arguments under high involvement, $t(594) = 26.44$, $p < .0001$, but not under low involvement, $t < 1$, *ns*. More important, as predicted by the bias hypothesis, when the arguments were ambiguous, higher expertise produced more agreement, $t(594) = 13.27$, $p < .0001$.

The same ANOVA applied on the valenced thoughts revealed the predicted interaction between Involvement and Quality of Arguments, $F(2,588) = 520.74$, $p < .0001$. Strong arguments generated significantly more thoughts that were consistent with the valence of the arguments than did weak arguments under high involvement, $t(594) = 32.85$, $p < .0001$, but not under low involvement, $t < 1$, *ns*. Again, as predicted by the bias hypothesis, when the arguments were ambiguous, high expertise elicited more favorable thoughts about the message than low expertise, $t(594) = 7.96$, $p < .0001$.

Simulation 5: Mood Heuristic

Dual-process research has revealed that mood also operates like a heuristic and that it also influences central processing (e.g., Petty, Schumann, Richman, & Strathman, 1993; for discussion, see Schwarz, Bless, & Bohner, 1991). Our connectionist approach to the heuristic impact of mood on cognition shares many similarities with affect priming theories (e.g., Bower, 1981) and affect-as-information theories (e.g., Schwarz & Clore, 1983) that instigated a lot of research on mood-congruent judgments (for an overview, see Forgas, 2001). According to affect priming theory (Bower, 1981; Isen, 1984), mood biases occur through mood-congruent attention, encoding, and retrieval of information involved in judgmental processes. These biases were explained by the mechanism of activation spreading in an associative memory network. This is obviously consistent with the present connectionist approach that assumes the same mechanism of automatic activation spreading. According to affect-as-information theory (Schwarz, 1990; Schwarz & Clore, 1983), affect has an informational value because people ask themselves "How do I feel about it?" when they evaluate persons or objects. More importantly, this theory posits that mood biases occur when people attribute (erroneously) the source of their affect to the attitude object.

In these simulations, we take this former mood activation spreading approach to simulate the impact of the mood heuristic. This assumes that, unlike the previous heuristic simulations, mood has a direct effect on valence without the aid of arguments. Thus, we learned in the past that a positive mood is favorable and a neutral mood is a mixture of favorable and unfavorable valences. However, if we take the latter misattribution approach, which assumes that perceivers often erroneously associate their feelings with the quality of the arguments or attributes of an attitude object, we obtain similar simulation results. This alternative presupposes that perceivers attribute their positive mood during heuristic processing to attributes and arguments of high quality, which elicits a positive valence, whereas

[2]Research has shown that people can discount their current mood as a source of information when made aware of it (e.g., Sinclair, Mark, & Clore, 1994). Such strategic use of mood is not modeled in our simulation.

they interpret their neutral or negative mood as indicating that the attributes and arguments are of low quality, which elicits a negative valence. In other words, mood acts as if it is equivalent to a piece of positive or negative information.[2]

Key experiment. In a prominent study by Petty et al. (1993, experiment 2), the effects of mood on attitude formation were investigated under low and high personal involvement. After a positive or neutral mood induction, participants were exposed to persuasive communication concerning a fictitious "Maestro" pen. Involvement was manipulated by telling the participants either that the Maestro pen would be marketed soon in their community and they had to make a selection between several brands of writing implements (high involvement) or that the marketing would take place in other cities and they had to make a selection between several brands of instant coffee (low involvement). Quality of argumentation was manipulated in this study but had no significant effects on attitudes (see Petty et al., 1993, for details[3]). This factor will therefore not be further discussed here. As can be seen in Figure 10 (top panel), positive mood produced more positive attitudes in agreement with the persuasive message under both low and high involvement, whereas positive mood influenced the positivity of thoughts only under high involvement (bottom panel).

Simulation. We simulated the effects of mood as explored by Petty et al. (1993, experiment 2). Table 6 presents a schematic learning history that is similar to earlier simulations of heuristic processing, except for the following elements. As mentioned previously, because quality of arguments had no effect, to simplify the simulation we simulated only strong arguments. In the Prior Heuristic phase, we assume that positive mood generates favorable evaluations directly, whereas neutral mood generates favorable and unfavorable evaluations.

Results. The statements listed in Table 6 were processed by the network for 50 participants with different random orders. Figure 10 depicts the mean test activation for all simulated attitude measures (top panel) and thought measures (bottom panel) on top of the empirical data of Petty et al. (1993). As can be seen, the simulation closely matched the attitude data. Positive mood increased the simulated attitude under both low and high involvement. The valenced thoughts were also replicated although to a somewhat lesser degree. There were few simulated thoughts under low involvement, and under high involvement, the thoughts revealed the same pattern as the simulated attitudes.

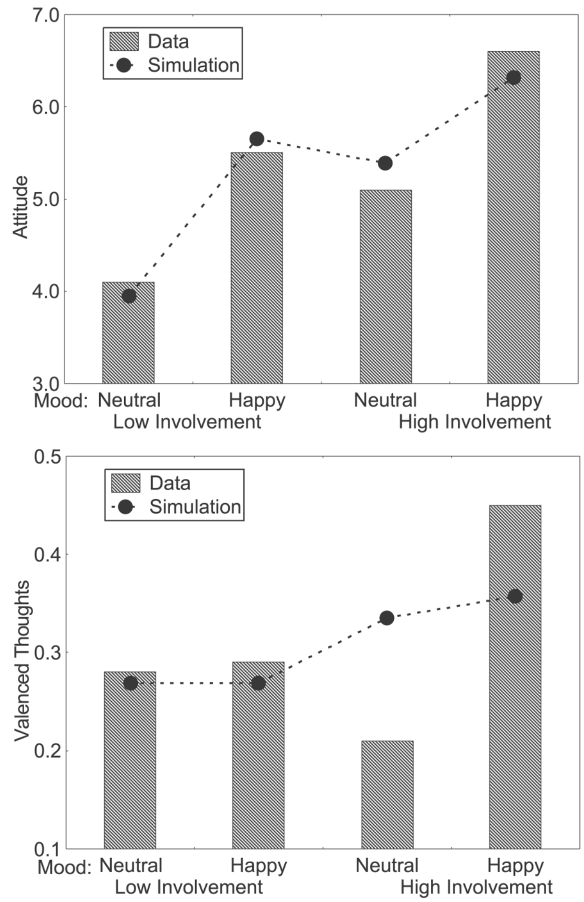

Figure 10. Mood heuristic: Observed data from Petty et al. (1993, experiment 2) and simulation results of attitudes (top panel) and valenced thoughts (bottom panel). Human data are denoted by bars, simulated values by broken lines. The human data are from Figure 3 in "Positive mood and persuasion: Different roles for affect under high- and low-elaboration conditions" by R. E. Petty, D. W. Schumann, S. A. Richman, & A. J. Strathman, 1993, *Journal of Personality and Social Psychology, 64*, p. 16. Copyright 1993 by the American Psychological Association. Adapted with permission.

These observations were verified with an ANOVA with two between-subjects factors, Involvement (low and high) and Mood (neutral and happy). The analysis on the simulated attitudes revealed a main effect of Mood, $F(1,196) = 2666.88$, $p < .0001$, indicating that a happy mood produced significantly more agreement. There was also an expected main effect of Involvement, $F(1,196) = 1702.67$, $p < .0001$, indicating that high involvement (in the processing of strong arguments) led to more agreement with the message than low involvement.

The analysis on the valenced thoughts revealed the predicted interaction between Mood and Involvement, $F(1,196) = 213.17$, $p < .0001$. A happy mood generated significantly more positive thoughts than did a neutral

[3]For their first study, Petty et al. (1993) reported a similar problem in that "it was possible … to interpret the weak arguments in a positive light" (p. 10).

Table 6. *Learning Experiences and the Mood Heuristic (Simulation 5)*

	Object & Cue		Arguments			Valence	
	Pen	Mood	Str1	Str2	Str3		
			#10 Prior Heuristic Learning: Positive (*Neutral*) Mood				
#4	0	+	0	0	0	1	0(*1*)
#4	0	+	0	0	0	1	0(*1*)
#4	0	+	0	0	0	1	0(*1*)
			Strong Message				
#1	1	1	1	0	0	*i*	*i*
#1	1	1	0	1	0	*i*	*i*
#1	1	1	0	0	1	*i*	*i*
			Test				
Attitude Toward Pen	1	0	0	0	0	?	–?
Valenced Thoughts	1	0	?	?	?	3 ?	3 ?

Note. Simplified version of the experimental design by Petty, Schumann, Richman, and Strathman (1993, exp. 2). Str = strong, Wk = weak, J = favorable; L = unfavorable; # = frequency of trial or condition, + = external activation of 0.5, *i* = internal activation (generated mainly by the arguments) is taken as external activation. Each experimental condition was run separately, and always preceded by a Prior Valence Learning phase (not shown) and Prior Heuristic Learning phase, followed by the Test phase. Trial order was randomized in each phase and condition. During Prior Valence Learning (not shown), all strong and weak argument nodes were paired with the favorable or unfavorable valence nodes respectively for 15 trials (see also Simulation 1). During Prior Heuristic Learning, each condition was repeated 10 times with 10% of the default learning rate. During heuristic processing of the experimental phase, activation was reduced to 10% for the cue and to 1% for the arguments during acquisition of novel information and testing of attribute-relevant thoughts.

mood under high involvement, $t(196) = 20.96$, $p < .0001$, but not under low involvement, $t < 1$, ns.

Implications and Extensions

All simulations in the preceding sections successfully reproduced the observed attitude and thought data from the empirical studies that tested a dual-process approach to attitude formation and change (Chaiken, 1987; Chen & Chaiken, 1999; Petty & Cacioppo, 1981, 1986; Petty & Wegener, 1999). Providing a formal account of the most important psychological processes in attitude formation by a unitary connectionist framework is an important achievement in its own right, because it organizes existing research and also because earlier attempts (e.g., Fishbein & Ajzen, 1975) articulated only fragments of these processes (e.g., central processes) at a mere input–output or computational level (cf. Marr, 1982). Perhaps more crucially, it allows making some tentative hypotheses about the nature of these underlying processes. To the extent that other comprehensive formalizations are lacking, it gives more weight to the present hypotheses than to alternative hypotheses that are not supported by a connectionist approach. In addition, it points to similarities with other connectionist models of social cognition (Van Overwalle & Labiouse, 2004; Van Rooy et al., 2003), which may suggest how attitude research can be extended to similar phenomena uncovered in these areas. Although we touched on some of these issues already, we first recapitulate some implications of the present model and then discuss empirical and theoretical extensions.

Implications for the Underlying Psychological Mechanisms

Quantitative and qualitative processing differences. Perhaps, the most central idea of this article was that the delta learning algorithm may provide a common underlying psychological mechanism responsible for different routes or modes of processing at a surface level of perceivers' intuition and awareness. We assumed that heuristic and central processes are based on different information bases (prior knowledge vs. novel information respectively) that are developed and applied somewhat differently (generalized cue knowledge vs. second-order object→valence connections) rather than involving radically different processing systems or brain structures. In this manner, we were able to account for qualitative differences in persuasion that attitude researchers have uncovered (Petty & Wegener, 1999). However, the network was endowed with sufficient flexibility through a supervisory attentional system that funneled activation to one of these information bases, so that it also could account not only for these qualitative differences, but also for quantitative differences in elaboration likelihood (Petty & Wegener, 1999). One implication of this flexibility is that it allows the network to consider heuristic cues also as information bases for deliberative scrutiny if attention is sufficiently large (see also Chen & Chaiken, 1999). For example, several studies have documented that under moderate elaboration, expertise can determine the extent of thinking (e.g., Heesacker, Petty, & Caciopppo, 1983) as can mood (e.g., Mackie & Worth, 1989; Schwarz et al., 1991). In general, the supervisory attention module in our model can account for switching of strategies although the executive processes that implement such changes in attention are not yet part of the model.

Nevertheless, there is another sense of qualitative difference that is not captured in our model. Some authors argued that central processing involves the effortful analysis of logical links via the use of propositional reasoning (e.g., Smith & DeCoster, 2000; Strack & Deutsch, 2004). Needless to say that propositional or symbolic processing is not part of our model. We only assume that once the essence of the arguments (e.g., the attributes) are symbolically understood (which is sometimes very easy when simple persuasive appeals are used such as "better" etc.), then our model proposes that their associated valences are automatically retrieved from memory and combined into a novel attitude. Thus, our model cannot fully accommodate all aspects of central processing, as it leaves propositional understanding and reasoning on the coherence and relevance of the arguments to a higher-level symbolic subsystem of the brain.

Valenced thoughts mediate attitude formation. We simulated the typical finding in dual-process research, borrowed from the cognitive response approach (Greenwald, 1968), that the quality and valence of the object's attributes as expressed in thought-listing measures determine attitude change under effortful or central processing. Although some authors claimed that attribute-relevant thoughts may just represent an alternate dependent measure of persuasion (Miller & Colman, 1981), in our connectionist model, the attitudes depended entirely on the favorable or unfavorable evaluations generated by the object's attributes, and without these, no object→valence connections would be established. Thus, consistent with the dual-process approach, our formalization points out that the evaluations generated by attribute-relevant thoughts (as later revealed in a thought-listing task) greatly affect attitude formation under central processing.

Implicit integration of valences. Our simulations also suggest that after the object's attributes have been symbolically analyzed during central processing (see aforementioned), the subsequent integration of evaluations into a single attitude estimate can occur largely outside awareness. This is because the delta learning algorithm that implements this integration does not need a central supervisory unit to control this process, as all changes involve low-level modifications in object→valence connection strength. As discussed earlier, this is consistent with recent theorizing in attitude models (e.g., Ajzen, 2002; Chen & Chaiken, 1999) and findings documenting that the processes underlying attitude formation and change are largely nonsymbolic and nonconscious (Betsch et al., 2001; Betsch, Plessner, & Schallies, 2004; Lieberman et al., 2001; Olson & Fazio, 2001). It also puts the present approach closer to lower-level processes like (subliminal) conditioning (e.g., Dijksterhuis, 2002; Riketta & Dauenheimer, 2002) and mere exposure effects, which do not require conscious attention either.

Recent neural imaging research suggests that there are distinct subsystems responsible for the controlled integration of valenced information versus the automatic conditioning and activation of valences (located in the medial prefrontal cortex vs. amygdala respectively, cf. Cunningham, Johnson, Gatenby, Gore, & Banaji, 2003). It would be interesting to see how the results of these imaging studies would extend to a typical persuasive paradigm although neural techniques may not allow revealing in a clear-cut manner the distinction made here between controlled information access and automatic evaluative integration.

Heuristics are not abstracted rules. As noted earlier, there are at least two possible interpretations of heuristics that were not clearly distinguished in dual-process attitude theories (e.g., Chaiken et al., 1989). One interpretation is that heuristics are knowledge structures consisting of abstracted inferential rules that are activated from memory and applied as tools for cognitive work. Another interpretation, consistent with our connectionist approach, sees heuristics as exemplar-based summarized experiences that reflect people's implicit knowledge about the statistical relation between situational cues and agreement with messages. These summarized exemplars reside in cue→valence connections, which are automatically integrated into novel information upon mere perceiving or thinking about the cue. This is in line with an increasing number of connectionist simulations in other domains of psychology illustrating that many rule-like behaviors are not necessarily driven by abstract inferential rules and can be more parsimoniously explained by subsymbolic properties of connectionist models (e.g., McLeod et al., 1998; Pacton et al., 2001; Rumelhart & McClelland, 1986; Smith & DeCoster, 2000).

Initial Evidence of Heuristics as Exemplar Based

If it is indeed true that persuasion heuristics in people's minds reflect summarized exemplar knowledge rather than explicit inferential rules, this may have testable implications. For instance, our approach predicts that under peripheral processing, one can induce the application of heuristics by priming relevant exemplars more so than by priming explicit heuristic rules.

Take, for instance, the consensus heuristic. According to an inferential rule approach, priming an abstract consensus rule like "I agree often with the majority" should lead to stronger attitude agreement under heuristic processing. There is some research that investigated the effect of priming heuristic rules. Chaiken (1987) reported on two unpublished studies in which the consensus and length rule were made more accessible by priming. During an ostensibly unrelated experiment, eight sentences were provided that conveyed the gist of the rule (pp. 27–29). However, overall, there were no significant attitude effects and only participants who were low in Need For Cognition (i.e., who tend to avoid extensive thinking and elaboration; Cacioppo & Petty, 1982) were somewhat influenced by the primed rule. Chaiken (1987) admitted that none of these results "yielded statistically robust effects favoring our [rule] priming hypotheses" (p. 29).

Our connectionist or exemplar-based approach makes a different prediction. Instead of rules, priming many versus few exemplars of relevant people should lead to stronger attitude agreement. Recent studies from our lab confirmed this prediction. In one study (Van Duynslaeger & Van Overwalle, 2004) that manipulated the consensus heuristic, 292 freshmen read five weak or strong arguments about a topic ostensible given by "some student associations" and were then asked to provide their opinion about it. To induce heuristic processing, the topic was of little concern to them because it involved research in nonuniversity higher education. Before that, they were primed with either exemplars or a rule indicating low or high consensus. Specifically, in the exemplar priming condition, they were primed with one or eight exemplar sources (i.e., different students from different student associations featured in an article on the renovation of the university restaurant). In the rule priming condition, they were primed one or five times with the consensus rule (i.e., "I always agree with the opinion of the majority"), using the repeated expression procedure of Powell and Fazio (1984) that is typically applied to manipulate the activation of attitudes.

The results (see Figure 11, top panel) were largely consistent with our predictions. An ANOVA revealed the predicted interaction between type and degree of priming, $F(1, 284) = 4.95$, $p < .05$. After priming more exemplars, the participants changed their attitudes more, $F(1, 141) = 3.42$, $p < .05$ (one-tailed), whereas after priming the consensus rule, there was no change, $F < 2$. However, one aspect of the results was unexpected. The exemplar priming was effective for strong arguments, $t(141) = 2.76$, $p < .01$, but not for weak arguments, $t < 1$. Although this raises the possibility that attitude change was due to more systematic processing of the arguments or of the heuristic cues, a correlational analysis with the thought data ruled out this explanation. Perhaps, the stronger quality of the arguments induced a minimal amount of cognitive attention that was necessary for the exemplars to have an impact.

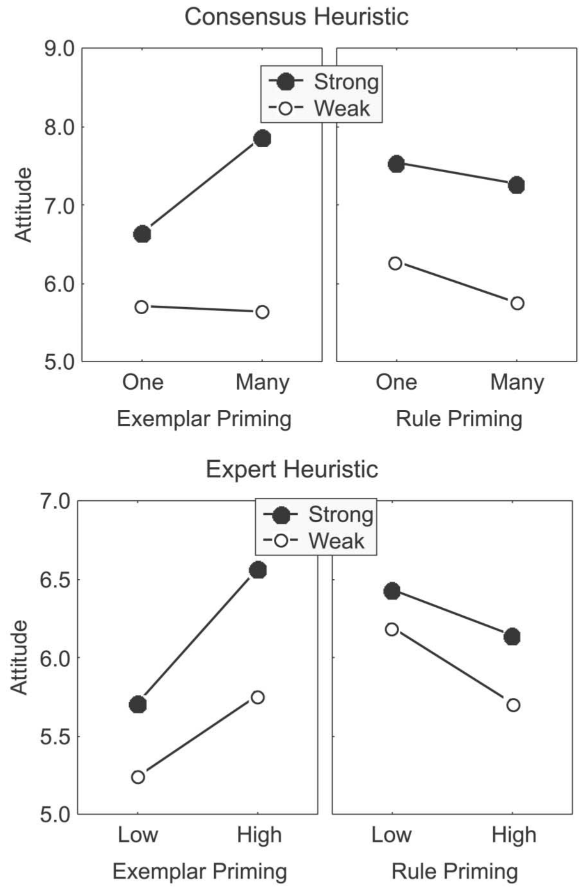

Figure 11. Heuristic use as a function of exemplar or rule priming: Observed data on the consensus heuristic from Van Overwalle and Van Duynslaeger (2004) and on the expertise heuristic from Van Duynslaeger (2004).

In another study (Van Duynslaeger, 2004) that manipulated the expert heuristic, 160 students from higher education school read the same five weak or strong arguments ostensible given by an unspecified "famous Flemish person." Before that, they were primed with either well-known exemplars varying in expertise on scientific issues (i.e., an astronaut vs. a reality show celebrity), or they were primed with the expert rule (i.e., "I often agree with the opinion of a trustworthy expert") using the repeated expression procedure. The results (see Figure 11, bottom panel) were again in agreement with our predictions. An ANOVA revealed the predicted interaction between type and degree of priming, $F(1, 152) = 4.54$, $p < .05$. After priming the expert exemplar, the participants changed their attitudes more than after priming the nonexpert exemplar, $F(1, 79) = 3.85$, $p < .05$ (one-tailed), whereas priming the expert rule did not have any effect, $F < 2$. Again, exemplar priming was more effective for strong arguments, $t(76) = 1.72$, $p < .05$ (one-tailed), than for weak arguments, t

< 1.1. Taken together, consistent with our prediction, priming heuristics had an effect only if it involved exemplars and not symbolic rules.

Extensions From Simulations in Other Domains

One of the major goals of this article was to demonstrate that attitude formation processes, like many other processes in social cognition, can be interpreted in a connectionist framework. Besides its theoretical interest as a step in the construction of a unifying theory of social thinking, it may also introduce novel cross-domain predictions that have been rarely tested in attitude research. We would like to suggest a number of such cross-domain findings that may perhaps lay the groundwork for more hypotheses and research in the future.

Attitude ambivalence. Sometimes people experience a great deal of conflict and ambivalence about attitude (e.g., Kaplan, 1972; for a review, see Priester & Petty, 1996). Ambivalence in one's attitudes may have important consequences. It may result in decreased attitude accessibility (Bargh et al., 1992) and less attitudinal confidence and persistence (Jonas, Diehl, & Brömer, 1997; MacDonald & Zanna, 1998). According to Kaplan (1972), ambivalence is determined in part by the sum of positive and negative attitude components. Thus, as more opposing beliefs are considered, the person would experience more ambivalence. Hence, in connectionist terms, the most simple and direct manner to measure ambivalence is by taking—instead of the differential activation of the favorable and unfavorable valences for measuring attitude—their summed activation. This measure provides an indication of the spread or range between the two opposing valences and is akin to a connectionist measure of people's estimates of the heterogeneity of group attributes recently proposed by Van Rooy et al. (2003). Using this measure, it is possible to "postdict" the finding of Priester and Petty (1996, Experiments 2 and 3) that ambivalence is a function of the number of deviant or conflicting pieces of information that are negatively accelerating. That is, as more conflicting information contributes to ambivalence, its contribution becomes increasingly smaller. This is precisely what the emergent acquisition property of the delta algorithm would predict.

Increased memory for inconsistent arguments. Simulations with a recurrent network using the delta learning algorithm in the domain of person perception (Van Overwalle & Labiouse, 2004) replicated the intriguing finding that inconsistent or unexpected behavioral information about an actor is often better recalled than information that is consistent with the dominant trait expectation (for a review, see Stangor & McMillan, 1992). Earlier theorizing explained this finding in terms of deeper processing of inconsistent information which results in more dense associations with the inconsistent behavior (Hastie, 1980). However, Van Overwalle and Labiouse (2004) proposed a novel emergent connectionist property of diffusion to explain this finding in terms of weakened memory for consistent behavioral information. This same emergent property may operate for attitudes and may likewise result in weaker memory for majority and consistent arguments as opposed to minority and inconsistent arguments. In addition, this emergent property predicts that the recall advantage should (a) increase for arguments at the end of a list, (b) decrease when the number of inconsistent arguments increases, but (c) remain high even when the number of consistent and inconsistent arguments is equal and inconsistency is manipulated by inducing a prior attitude.

Subtyping of deviant sources. Another finding in group processes that was simulated in a recurrent network by Van Rooy et al. (2003) is subtyping. Members of a group with extreme positions on an issue are typically subtyped into subcategories and separated from the rest of the group, more so than members with moderate deviating positions. This insulates the group from dissenting members, so that the content of the existing group stereotype is preserved. Van Rooy et al. (2003) explained this phenomenon by the delta algorithm's emergent property of competition, which predicts that the more information is concentrated in a few members, the more it must compete against the group stereotype and is discounted. This emergent property may also apply in attitude formation. Hence, we predict that extreme deviant positions on issues that are defended by a few sources are more easily discounted than mildly deviant positions supported by many sources. Consequently, the best tactic to change attitudes is to distribute disconfirming information among as many sources as possible, so as to avoid subtyping of extreme deviant sources.

Contrast effect in ease of retrieval. This approach can be extended to other heuristic effects, such as the ease of retrieval effect under central processing conditions (Tormala, Petty, & Briñol, 2002). The ease of retrieval effect refers to the phenomenon that when people are asked to come up with arguments for a given attitude position, they are more in favor of communication if they have to generate only a few arguments, and less in favor if they have to generate a high number of arguments (see also Wänke & Bless, 2000; Wänke, Bless, & Biller, 1996; Wänke, Bohner, & Jurkowitsch, 1997). People thus reveal a contrast away from the requested number of arguments (e.g., less in favor if more arguments are requested). This effect can be simulated based on the idea that the requested number of arguments serves as a standard of comparison. Recently, Van Overwalle and Labiouse (2004) proposed a connectionist account for contrast effects in person perception through the presence or priming of exemplary others who serve as a standard of comparison. They documented that this contrast effect may be due to the emergent property of competition against a standard, and this idea might be extended to the attitude domain. To understand how this might work, we will describe a simulation of the ease of retrieval effect in somewhat more detail.

Simulation 6: Ease of Retrieval Effect

Key experiment. Tormala et al. (2002, Experiment 2) asked their participants to read a persuasive communication concerning a new exam policy and requested them to generate either 2 or 10 favorable arguments in response. They found that under central processing, participants were more in favor of the communication if they had to generate only 2 arguments, and they were less in favor if they had to generate 10 arguments. However, under peripheral processing, the opposite pattern was found. Presumably participants did not consider the material thoroughly but were rather influenced by the sheer number of the arguments required and generated. Tormala et al. (2002) explained these results by arguing that subjective confidence influences judgments. The easier it is to generate a list of arguments (because a low number is required), the more confidence an individual has in them. The more difficult it is to generate a list of arguments (because a high number is required), the less confidence an individual has in them. Confidence in one's thoughts is especially important under central processing when people's motivation and ability to process the information is relatively high, and less so under peripheral processing.

The previous explanations of the ease of retrieval effect rely on metacognitive processes, that is, the subjective sense of ease or difficulty of generating arguments, or confidence. These processes are not part of our model. Therefore, we suggest an alternative connectionist explanation of Tormala et al.'s (2002) findings that does not involve metacognitive processes and where these subjective feelings are merely an epiphenomenon of an underlying connectionist mechanism.

To understand our approach, it is important to realize that participants in this type of research typically retrieve fewer arguments than the requested high number and are thus forced to generate novel arguments (e.g., Wänke et al., 1996). We argue that these novel arguments are less convincing because the difficulty in generating them "will be attributed some qualitative aspect of the information" (Wänke & Bless, 2000, p. 158) or, alternatively, because they are mostly redundant with respect to the already retrieved arguments. For instance, after retrieving "good health" as a reason for engaging in sports, people might construe "on doctor's advice" as a novel argument that actually only rephrases the original one. Although observers read novel arguments as equally convincing in isolation (Wänke et al., 1996), participants themselves find them less "compelling" (Wänke & Bless, 2000) and less "strong" (Haddock, 2000; but see Haddock, Rothman, & Schwarz, 1996) and have less "confidence" in them (Tormala et al., 2002). There is also research demonstrating that the ease of retrieval effect is found regardless of whether the requested arguments are actually listed or not (Wänke et al., 1997), suggesting that participants have the intuition that they only rephrase or use less compelling arguments. Because the exact reason for the reduced convincingness is not known, future research may explore in more depth the source of it and whether argument overlap plays a significant role (e.g., by assessing the perceived redundancy of newly construed arguments). In the simulation, we implemented our interpretation in terms of reduced perceived quality by ignoring additional constructed (but less convincing) arguments, thus keeping the same amount of spontaneously retrieved arguments in all conditions.

We suggest that the required number of arguments may act like a situational length cue. Under peripheral processing, this promotes the operation of the length heuristic and so dominates attitude formation in much the same way as in Simulation 2. However, under central processing, the reverse effect of ease of retrieval is due to a contrast effect of the heuristic length cue against one's own spontaneously retrieved number of arguments. In line with Van Overwalle and Labiouse (2004, Simulation 5) analysis of contrast effect in person impression formation, this latter effect relies on the emergent connectionist property of competition. This property arises when multiple factors compete in predicting or causing an outcome and produces a lowering of the connection weights, similar to discounting in causal attribution (Kelley, 1971) and blocking in the conditioning literature (Rescorla & Wagner, 1972). The connectionist mechanism behind competition is that the internal activation in the valence nodes is determined by the sum of all activations received from the attitude object and all other external cues present. Discounting of an attitude occurs when the connection weights of external cues are already strong so that any additional growth of the object→valence connection is blocked.

In the simulations, the competition effect ensues most strongly under central processing, because the novel information receives full activation so that it may compete more against prior knowledge. Specifically, when a high number of arguments is requested, competition will ensue between the strong cue→valence connection and the object→valence

Table 7. *Learning Experiences and Ease of Retrieval Effect (Simulation 6)*

	Object & Cue		Arguments[a]			Valence	
	Exam	Length	Str1	Str2	Str3		
	#10 Prior Heuristic Learning: Short *(Long)* Strong Message						
#0 (2)	0	+	+	0	0	*i*	*i*
#1 (2)	0	+	0	+	0	*i*	*i*
#0 (2)	0	+	0	0	+	*i*	*i*
	Strong Message						
#1	1	1	1	0	0	*i*	*i*
#1	1	1	0	0	1	*i*	*i*
	Test						
Attitude Toward Exam	1	0	0	0	0	?	–?

Note. Simplified version of the experimental design by Tormala, Petty, and Briñol (2002, exp. 2). Exam = Exam policy, Str = strong, J = favorable; L = unfavorable; # = frequency of trial or condition, + = external activation of 0.5, *i* = internal activation (generated mainly by the arguments) is taken as external activation. Each experimental condition was run separately, and always preceded by a Prior Valence Learning phase (not shown) and Prior Heuristic Learning phase, followed by the Test phase. Trial order was randomized in each phase and condition. During Prior Valence Learning (not shown), all strong argument nodes were paired with the favorable valence nodes respectively for 15 trials (see also Simulation 1). During Prior Heuristic Learning, each condition was repeated 10 times with 10% of the default learning rate. During heuristic processing of the experimental phase, activation was reduced to 10% for the cue and to 1% for the arguments during acquisition of novel information and testing of attribute-relevant thoughts.

[a]The arguments during prior learning are completely different from those in the experimental and test conditions, but are shown in the same columns to conserve space. The arguments during prior heuristic learning serve to drive the cue's valence into a positive or negative direction, but are of no further importance.

connection, resulting in discounting of the latter connection. This mechanism produces a contrastive effect away from the advocated position in the communication (i.e., ease of retrieval effect). In contrast, when a low number of arguments is requested, little competition will ensue between the weak cue→valence connection and the object→valence connection, and so the attitude will be relatively favorable.

Simulation. Table 7 represents a simplified simulated learning history of Tormala et al. (2002; Experiment 2). The Prior Valence and Prior Heuristic Learning phases were identical to Simulation 2 of the length heuristic. A request of a low versus high number of arguments was simulated in a Prior Heuristic Learning phase by simulating one or six arguments, which is roughly equivalent to the number of arguments used by Tormala et al. (2002). Next, the Experimental phase replicated the generation of arguments. In this line of research, the two requested numbers are selected by the experimenter such that they are smaller and larger, respectively, than the number of arguments that people would retrieve spontaneously (see, e.g., Wänke et al., 1996). Accordingly, we choose two arguments.

Results. The statements of each condition listed in Table 7 were processed by the network for 50 participants in each condition with different random orders. Figure 12 depicts the mean test activation for all simulated attitude measures, projected on top of the empirical data from Tormala et al. (2002). As can be seen, the simulation closely matched the attitude data. An ANOVA with Involvement (low and high) and Number of Requested Arguments (low and high) as between-subjects factors, revealed that the expected interaction was significant, $F(1, 396) = 308.13$, $p < .0001$. Since Tormala et al. (2002) did not report subjective ease of retrieval, we were not able to assess their role in this simulation.

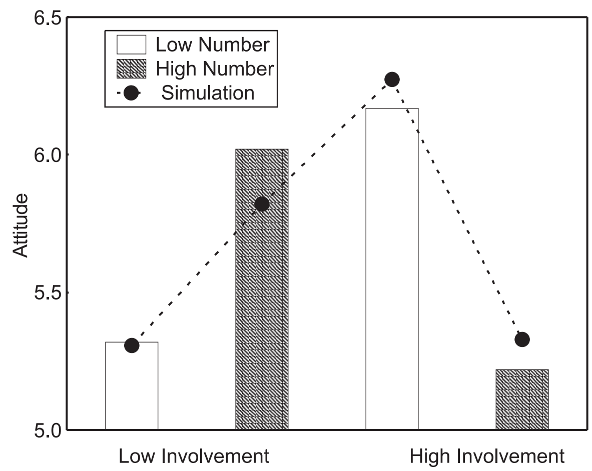

Figure 12. Ease of retrieval effect: Observed data from Tormala et al. (2002, experiment 2) and simulation results. Human data are denoted by bars, simulated values by broken lines. The human data are from Figure 1 in "Ease of retrieval effect in persuasion: A self-validation analysis" by Z. L. Tormala et al., 2002, *Personality and Social Psychology Bulletin, 28*, pp. 1705—1707. Copyright 2002 by the Society for Personality and Social Psychology. Adapted with permission.

Model Comparisons

What do other recent approaches in the literature besides dual-process models have to say about attitude formation, and how do they compare with the present approach? In addition, how robust are the present simulations with respect to other possible connectionist implementations? We begin with the last issue and then turn to a comparison with alternative models.

Alternative Implementations of the Simulations

The simulations that we have reported replicate the empirical data or theoretical predictions reasonably well. However, it is possible that this fit is due to some procedural choices of the simulations rather than conceptual validity. To explore whether our simulations are robust to changes in implementational choices, we applied a number of alternative encodings and processing parameters. First, we compared the present localist coding with a distributed coding to add more realism to the simulations. In a distributed coding, each concept or object is represented by a set of nodes, instead of one node as in localist encoding. The advantage is that these nodes reflect subsymbolic features that we are not always aware of but that nevertheless may influence our attitudes and that may also include the context in which the attitude is typically applied. As a first attempt toward such a more realistic implementation, each concept was represented by a distributed pattern of activation across four nodes, drawn from a normal distribution with mean as indicated in the learning histories and *SD* .10. In addition, in all prior learning phases, random noise drawn from a normal distribution with mean 1 and *SD* .10 was added to these activations to reflect the variation in past experiences. Second, we compared the present dual encoding of favorable versus unfavorable evaluations with a unitary valence encoding in which unfavorable evaluations are represented by negative activation levels (instead of positive activation levels of an unfavorable valence). Third, we also used, instead of the present linear updating activation algorithm with two internal cycles, a nonlinear activation updating algorithm as used by other social researchers (Read & Montoya, 1999; Smith & DeCoster, 1998). Note that apart from these alterations, all alternative implementations used the same simulation specifications, unless noted otherwise in Table 8.

Table 8 lists the correlations between the simulated values and the observed or theoretical data from the original as well as from each of the alternative implementations. As can been seen from the mean correlations (see last two bottom rows), although some of the alternative implementations are adequate, the present simulations are often superior. First, the distributed coding is adequate for all simulations, except that the biasing effect of the expert heuristic on ambiguous information (Simulation 4) was not replicated. This finding is puzzling, and we have no clear answer for it. Second, the unitary valence coding generally leads to a weaker fit with the data (especially for thoughts), presumably because a single valence node does not allow for a neutral or ambivalent evaluation that dual valence nodes can support (by coding the valences as both low or both high respectively). Finally, a nonlinear recurrent model (with parameter values very close to the linear model) provides the weakest fit with the data. Overall, the results suggest that the present specifications are preferable to a unitary valence coding scheme and a nonlinear activation update algorithm, and the results of the distributed coding were equivalent, with one exception. Of course, other implementations of the

Table 8. *Fit and Robustness of the Simulations, Including Alternative Encoding and Models*

Number and Topic	Dependent Measure	Original Simulation	Distributed Coding	Unitary Valence	Nonlinear Recurrent
1. Central	Attitude	1.00*	.99*	.88	.99*[a]
2. Length	Attitude	.96*	.94*	.97*	.90*[a]
	Thoughts	.80*	.76*	.54	.75*[a]
3. Consensus	Attitude	.88*	.85*	.91*	.71*[a]
	Thoughts	.84*	.62*	.21	.56*[a]
4. Expertise	Attitude	.94*	.73*[b]	.89*	.70*
	Thoughts	.88*	.84*[b]	.74*	.86*
5. Mood	Attitude	.97*	.92*	.78	< 0
	Thoughts	.45	.69	.49	.79
6. Retrieval	Attitude	.95*	.92*	.92*	.82*
Means	Attitude	.95	.89	.89	.65
	Thoughts	.74	.73	.50	.74

Notes: Cell entries are correlations between mean simulated values (averaged across randomizations) and empirical data or theoretical predictions. The learning rate parameter for the distributed coding was the best fitting value between .01 and .05. The parameters for the nonlinear recurrent model were similar to the linear model, except that Decay = .77 (McClelland & Rumelhart, 1988).
[a]Decay = .70. [b]The biasing effect of the heuristic on ambiguous information was not reproduced. *$p < .05$ (one-tailed).

learning history and the network architecture that we did not explore are also conceivable.

Siebler's (2002) Parallel Constraint Satisfaction Network

Recently, Siebler (2002) proposed a connectionist parallel-constraint-satisfaction model (McClelland & Rumelhart, 1988; McLeod et al., 1998) with a single connectionist mechanism to account for dual-processing routes in attitude formation and change (see also Kunda & Thagard, 1996; Read & Marcus-Newhall, 1993; Shultz & Lepper, 1996; Spellman & Holyoak, 1992; Thagard, 1989). Siebler's constraint satisfaction model involves the simultaneous satisfaction of multiple, sometimes conflicting constraints on an individual's cognitions, including the attitude itself, positive and negative heuristic cues, weak and strong arguments and favorable and unfavorable cognitive responses. The model architecture assumes that cues are associated with attitudes in a relatively direct manner, whereas arguments are associated with attitudes more indirectly, via cognitive responses. These connections impose constraints that are soft rather than hard, so that they are desirable, but not essential to satisfy.

Although Siebler (2002) reports excellent fits with the empirical data of two experiments manipulating source expertise (Chaiken & Maheswaran, 1994; Petty et al., 1981), the constraint satisfaction network has a number of shortcomings. First, the constraint satisfaction network has no learning mechanism. The process of developing the connections in the network is not modeled. As a result, the model is nonadaptive, as the connections have to be hand set by the experimenter and do not develop automatically from prior or current learning. Second, the constraint satisfaction network limits attitude formation and change to temporary changes of activation in the network, driven by satisfying all constraints present. Hence, the network reflects only a short-lived mental state of attitude that occurs only when all relevant prior beliefs, heuristic cues, and persuasive arguments are activated (consciously or subconsciously) in the individual's mind. However, this is contradicted by a variety of empirical research showing that there is no substantial correlation between argument recall and attitude. Instead, it is now well established that most attitudes are formed online, and novel information is encoded and processed. To allow such online adjustment, a learning algorithm is essential.

Van Overwalle and Jordens's (2002) Feedforward Network of Cognitive Dissonance

Our attitudes are not only driven by immediate evaluations of attitude objects, but sometimes also by reactions to our own behaviors, especially when these behaviors go against our initial preferences. This dilemma has been investigated under the heading of cognitive dissonance (Festinger, 1957). For instance, when induced to write an essay that runs counter to one's initial attitude (e.g., a student defending stricter exam criteria), an individual will tend to reduce dissonance by changing his or her attitude in the direction of the position taken in the essay. This tendency is stronger when alternative explanatory factors or justifications, such as high payment or social pressure, are absent. In contrast, when such external demands provide sufficient justification for engaging in the dissonant behavior, dissonance reduction does not occur (e.g., Linder, Cooper, & Jones, 1967; Cooper & Fazio, 1984).

Van Overwalle and Jordens (2002) provided a feedforward model of this cognitive dissonance process. Their network involves the same type of concepts and connections as in the present recurrent model, with the important addition of a connection between the attitude object and behaviors performed by the person. The rationale was that individuals attempt not only to understand their evaluations, but also to justify their discrepant behavior. Both outcomes influence their attitudes. When alternative causal explanations for the discrepant behavior are absent, only the attitude object is sufficiently connected to these novel outcomes, resulting in the psychological realization that the object is liked more than initially thought. This results in attitude change. Conversely, when sufficient external explanations are available, their connections may sufficiently explain the outcomes, resulting in discounting and little attitude change. The mechanism responsible for this latter process in the network model of Van Overwalle and Jordens (2002) is the emergent property of competition.

This connectionist implementation of cognitive dissonance is largely consistent with the present network. First, in persuasive communication, little effect derives from one's behaviors, so that this factor could be safely ignored here. Second, because feedforward networks are more limited than recurrent models in the type of connections and the flow of activation, Van Overwalle and Jordens's (2002) feedforward network can be subsumed in the present more general recurrent model. They reported that their feedforward simulations of cognitive dissonance could easily be "upgraded" with very similar results to a recurrent architecture. Third, Van Overwalle and Jordens's (2002) hand coded all valences as +1, whereas in this model they were indirectly "coded" by the recurrent activation accumulated through the object's attributes. In this respect, again the recurrent model is more general than theirs. This leads to the conclusion that the present recurrent model encompasses perhaps a large range of earlier findings and models in the attitude literature, including attitude change due to cognitive dissonance

(Festinger, 1957; see also Van Overwalle & Jordens, 2002).

Eiser et al.'s (2003) Back-Propagation Network

Recently, Eiser et al. (2003) developed a connectionist model of attitude acquisition that differed from the previous connectionist models in that it assumes a more active role of the perceiver. Thus, not only passive exposure was modeled, but also active exploration of the environment. It was assumed that through such active learning, people will choose to engage in activities that they find enjoyable and avoid unpleasant activities as much as possible. Consequently, perceivers are much less accurate at identifying enjoyable versus unpleasant objects, resulting in an asymmetry in the appreciation of objects. In particular, unfamiliar objects are often seen as more negative than positive. Eiser et al. (2003) reported empirical support for this prediction, and also replicated these results with a multilayer feedforward model with the generalized delta (or back propagation) learning algorithm. The main advantages of their model are (a) a hidden layer that facilitates generalization from one situation to another and (b) a behavioral output component that makes active exploration in a virtual environment possible. However, a limitation is that the Eiser et al. (2003) model does not include nodes to represent the objects' attributes. Even after incorporating these, it remains to be seen to what extent this model is capable of replicating the empirical findings that were covered in this article. Our recurrent model does not incorporate a behavioral exploration process and hence does not expect differential learning of positive and negative attitudes, something that in any case was not reported in the persuasive literature discussed so far.

Kruglanski and Thompson's (1999) Unimodel

Kruglanski and Thompson (1999) questioned the assumption of dual-process theories that attitude change is attainable via two qualitatively distinct routes, and instead argued that these routes are functionally equivalent and differ only to the extent that they involve cognitive effort in decoding simple versus complex persuasive information. They proposed a unimodel that adopts a more abstract level of analysis in which the two persuasion modes are viewed as special cases of the same underlying process. Specifically, heuristic rules that are derived from prior beliefs and schemata stored in memory as well as explicit thoughtful elaboration of persuasive arguments rest on the same type of propositional if–then reasoning leading from evidence to a conclusion. For instance, heuristics are represented by an if–then propositional logic such as "if an opinion is offered by an expert, then it is valid" (p. 90), and central processing is also represented by if–then reasoning such as "if something contributes to the thinning of the ozone layer, then it should be prohibited" (p. 90). Thus, Kruglanski and Thompson (1999) concluded, "rule-based reasoning is common to both persuasion modes" (p. 104). In a series of experiments, Kruglanski and Thompson (1999) demonstrated that if the heuristic information (e.g., on the expertise of the source) is sufficiently complex, then such "heuristic" information might also require central processing before it has any impact. Conversely, if the "central" arguments are sufficiently brief and simple, they can have an impact under peripheral processing.

Our connectionist network is in agreement with the unimodel in its claim that there is a single core mechanism underlying both central and peripheral routes of persuasion, and both perspectives see the degree of elaboration as the main quantitative difference between the two processing routes. Although both models were developed independently from each other, they are remarkably similar in these respects. However, there are some noteworthy differences. First, there is a radically different view on the underlying mental processes responsible for attitude formation. Instead of the unimodel's explicit, symbolic, and sequential if–then reasoning logic of evidence, we proposed a lower-level connectionist mechanism, with a parallel, subsymbolic, and implicit processing of this information, ultimately leading to an explicit attitude belief. As noted earlier, we believe that this perspective is more in agreement with neurological evidence on the working of the brain and with recent findings showing that attitudes can be formed with little awareness of the integration process (Betsch et al., 2001; Lieberman et al., 2001). Second, our model proposes qualitative differences between processing modes, in that heuristic and central processes are based on different information bases (prior knowledge vs. novel information, respectively) that are developed and applied differently (generalized cue knowledge embedded in cue→valence connections vs. second-order object→valence knowledge), whereas the unimodel treats these as similar and built from the same if–then propositional logic.

Betsch et al.'s (2004) Value–Account Model

Recently, Betsch et al. (2004) put forward a Value-Account model that assumes that aggregation of preferences into a summary evaluation or value-account is by default implicit and automatic. Only when provided with sufficient motivation and capacity, perceivers will develop an aggregated attitude through explicit deliberation and weighting of specific information or episodes. This is consistent with most recent views on attitude processes, including ours. More importantly, based on a large series of experiments (e.g., Betsch et al., 2001), Betsch and colleagues suggested that implicit aggregation is guided by a summation principle, whereas explicit aggregation follows an averaging principle. Similarly, in evaluative conditioning research, De Houwer et al. (2001) argued that a summation pattern (based on a simple Hebbian learning algorithm) might be more typical of implicit preference learning, whereas an averaging pattern (by the delta learning algorithm) is more typical of explicit signal learning. How can these findings be reconciled with our model?

Recall that the delta algorithm predicts a negatively accelerating learning curve. In the beginning of learning, the learning error is still large so that each novel input results in substantial weight adjustments that are added to each other, reflecting a summation of the favorable and unfavorable valences. However, toward asymptote, the error is much reduced (i.e., people reached an overall evaluative estimate based on the evidence given), so that novel evidence results in less weight adjustments, reflecting an averaging of prior and novel valences (see Appendix B). Hence, the different patterns of integration can be accommodated in the present model by making the same assumption as we did for heuristic reasoning, that is, by taking the assumption that implicit learning is slower or shallower than explicit learning. Applied in the present context, this suggests that during implicit learning, the delta algorithm is still in its early "summation" phase, whereas explicit learning is faster so that delta learning enters its later "averaging" phase much more quickly. An interesting implication of this assumption is that implicit learning should attain an averaging phase after an extended time in which more information is presented.

Wilson et al.'s (2000) Model of Dual Attitudes

An intriguing challenge to the present approach was recently posed by the dual-attitude model of Wilson, Lindsey, and Schooler (2000). According to these authors, people may hold in memory different explicit and implicit attitudes toward the same attitude object. When such dual attitudes exist, the implicit attitude is activated automatically, whereas the explicit attitude requires more capacity and motivation to retrieve. The implicit attitude changes more slowly like old habits, whereas the explicit attitude changes relatively easily. Most attitude researchers agree that this distinction exists, but there is disagreement as to what may cause it.

In some cases, different outcomes from explicit and implicit measures may be due to the fact that each measure focuses on different aspects of the same attitude object in memory. In terms of this model, it would reflect testing the network by priming the subfeatures of the same attitude object with a different distributed activation pattern.[4] In other cases, it seems evident that the explicit attitude reflects some sort of suppression of illegitimate or unwanted thoughts, such as when people remove racial attitudes in explicit measures but disclose their implicit racial stereotypes when measured implicitly (e.g., under time constraints). Another example is when people realize that the information received earlier was incorrect, but still hold in memory the (incorrect) association between the object and their negative evaluations, a phenomenon known as evaluative perseverance (Wilson et al., 2000). This model cannot account for this dissociation because we did not model episodic (recent) and semantic (old) memory as separate memory structures but simply as different learning phases in time.

As noted earlier, several authors have made proposals for a dual memory system of the brain that may explain the dissociation between old and recent memories (French, 1997; McClelland et al., 1995; Smith & DeCoster, 2000). One subsystem would be dedicated to the rapid learning of unexpected and novel information and the building of episodic memory traces (e.g., the main experimental phase in our simulations). However, not only the learning, but also the decay of episodic traces is relatively fast in this subsystem. Hence, episodic memory lasts only a few days. In contrast, the other subsystem would be responsible for slow incremental learning of statistical regularities of the environment and gradual consolidation of information learned in the first subsystem, resulting in stable and more lasting semantic memory traces (e.g., the prior learning phases of our simulations). Because this latter subsystem has more permanent memory traces, novel information has relatively little effect so that the older attitudes often persist over time. The process of consolidation of recent memory into lasting memory could start a few minutes after receiving the novel information and last for several days. Consistent with this idea, Schooler (1990, cited in Wilson et al., 2000) reported that explicit attitude change resulted in a substantial dissociation between implicit and explicit attitude measures immediately afterwards, but that 48 hr later this difference gradually began to wear off.

[4]In addition to differences between implicit and explicit attitudes, people sometimes report different explicit attitudes depending on what information was activated before, what subset of data they attended to or retrieved from memory, what standard of comparison was salient, and so on (e.g., Wilson & Hodges, 1992). Our approach can accommodate many (but not all) of these results as reflecting the impact of a person's recent, pre-message learning history, through priming or attentional focus.

General Conclusions

This article introduced a novel connectionist framework of attitudes that provides an integrative account of many earlier perspectives of attitude formation, change, and use. The proposed model rests on the shoulders of pioneering work that was incorporated in its architecture and processing mechanisms. The model's architecture adopted the three-component view on attitudes as consisting of beliefs, evaluations, and behavioral tendencies (Katz & Stotland, 1959; Rosenberg & Hovland, 1960) and also implemented the basic idea from spreading activation networks that attitudes consist of object–evaluation associations in memory (Fazio, 1990). The model's learning algorithm was based on older work on associative learning processes (Rescorla & Wagner, 1972) and classical conditioning of attitudes (Olson & Fazio, 2001; Staats & Staats, 1958) and was shown to incorporate algebraic approaches to attitude formation (Fishbein & Ajzen, 1975). Of most importance was that this model could simulate heuristic and central processing as proposed in earlier dual-process models (Chaiken, 1987; Petty & Cacioppo, 1981, 1986; Petty & Wegener, 1999).

The proposed connectionist perspective offers a novel view on how information may be encoded in the brain, how it may be structured and activated, and how it may be retrieved and used for attitude judgments. One major advantage of a connectionist perspective is that it incorporates a learning algorithm that allows the model to associate patterns that reflect social concepts and evaluations by means of very elementary learning processes. Hence, complex social reasoning and learning can be accomplished by putting together an array of simple interconnected elements, which greatly enhance the network's computational power without the need for a central executive or awareness of its processing mechanisms. In addition, connectionist models have other capacities that we did not address such as its content-addressable memory, its ability to do pattern completion, and its noise tolerance (for more on these issues, see McClelland & Rumelhart, 1988; McLeod et al., 1998; Smith, 1996).

Given the extensive breadth of attitude research, we inevitably were not able to include many other interesting findings and phenomena that have now been simulated by similar connectionist models, such as cognitive dissonance (Van Overwalle & Jordens, 2002) and impression formation about persons or groups (Kashima et al., 2000; Queller & Smith, 2002; Smith & DeCoster, 1998; Van Overwalle & Labiouse, 2004; Van Rooy et al., 2003). There are other obvious limitations of the present model such as the lack of hidden or exemplar nodes (e.g., Kruschke & Johansen, 1999; McClelland & Rumelhart, 1988; McLeod et al., 1998; O'Reilly & Rudy, 2001) which limit its computational power, and the lack of distinct episodic and semantic memory structures to overcome the problem of "catastrophic interference" (McCloskey & Cohen, 1989; Ratcliff, 1990) and the parallel existence of different implicit and explicit attitudes (Wilson et al., 2000).

Given the importance of attention and motivation in attitude formation and change, it will ultimately be necessary to incorporate these factors into an improved model. For the time being, we manipulated the overall attention by a supervisory activation module. However, other aspects of attention are not part of the dynamics of our network. For instance, salient situational factors such as heuristic cues can sometimes motivate people to scrutinize persuasive information more carefully. Credible and likeable sources, or majority positions may motivate people to consider the message arguments more attentively, because these sources are more likely to provide correct or valuable information (Erb et al., 1998; Heesacker et al., 1983; Mackie, 1987; Roskos-Ewoldsen, Bichsel, & Hoffman, 2002) whereas negative mood may signal that the message content is problematic (Mackie & Worth, 1989; Sinclair et al., 1994; Wegener & Petty, 1996; Worth & Mackie, 1987; but see Bohner & Weinerth, 2001). It strikes us that the next step in connectionist modeling of attitudes and social cognition in general will involve exploring connectionist architectures built from separate but complementary systems with more consideration for the interaction between different subsystems of the brain.

References

Adolphs, R., & Damasio, A. (2001). The interaction of affect and cognition: A neurobiological perspective. In J. P. Forgas (Ed.). *Handbook of affect and social cognition* (pp. 27–49). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Ajzen, I. (1988). *Attitudes, personality and behavior*. Homewood, IL: Dorsey Press.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179–211.

Ajzen, I. (2002). Residual effects of past on later behavior: Habituation and reasoned action perspectives. *Personality and Social Psychology Review, 6,* 107–122.

Ajzen, I., & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology, 22*, 453–474.

Anderson, N. H. (1967). Averaging model analysis of set size effect in impression formation. *Journal of Experimental Psychology, 75*, 158–165.

Anderson, N. H. (1971). Integration theory and attitude change. *Psychological Review*, *78*, 171–206.

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

Anderson, N. H. (1981b). Integration theory applied to cognitive responses and attitudes. In R. E. Petty, T. M. Ostrom, & T. C. Brock (Eds.), *Cognitive responses in persuasion* (pp. 361–397). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ans, B., & Rousset, S. (1997). Avoiding catastrophic forgetting by coupling two reverberating neural networks. *Académie des Sciences de la Vie, 320,* 989–997.

Ans, B., & Rousset, S. (2000). Neural networks with a self-refreshing memory: Knowledge transfer in sequential learning tasks without catastrophic forgetting. *Connection Science, 12,* 119.

Axsom, D., Yates, S., & Chaiken, S. (1987). Audience response as a heuristic cue in persuasion. *Journal of Personality and Social Psychology, 53*, 30–40.

Baker, A. G., Berbier, M. W., & ValléeTourangeau, F. (1989). Judgments of a 2 × 2 contingency table: Sequential processing and the learning curve. *The Quarterly Journal of Experimental Psychology, 41B*, 65–97.

Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology, 87,* 5–22.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology, 62*, 893–912.

Berkowitz, L., & Knurek, D. A. (1969). Label-mediated hostility generalization. *Journal of Personality and Social Psychology, 13*, 200–206.

Betsch, T., Plessner, H., & Schallies, E. (in press). The value-account model of attitude formation. In G. Haddock & G. R. Maio (Eds.), *Contemporary perspectives on the psychology of attitudes* (pp. 251–274). New York: Psychology Press.

Betsch, T., Plessner, H., Schwieren, C., & Gütig, R. (2001). I like it but I don't know why: A value-account approach to implicit attitude formation. *Personality and Social Psychology Bulletin, 27*, 242–253.

Bohner, G., Ruder, M., & Erb, H. P. (2002). When expertise backfires: Contrast and assimilation effects in persuasion. *British Journal of Social Psychology, 41,* 495–519.

Bohner, G., & Weinerth, T. (2001). Negative affect can increase or decrease message scrutiny: The affect interpretation hypothesis. *Personality and Social Psychology Bulletin, 27*, 1417–1428.

Bower, G. H. (1981). Emotional mood and memory. *American Psychologist, 36*, 129–148.

Cacioppo, J. T., & Petty, R. E. (1982) The need for cognition. *Journal of Personality and Social Psychology, 42,* 116–131.

Canli, T., Desmond, J. E., Zhao, Z., Glover, G., & Gabrieli, J. D. E. (1998). Hemispheric asymmetry for emotional stimuli detected with fMRI. *NeuroReport, 9*, 3233–3239.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology, 39*, 752–766.

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario Symposium* (Vol. 5, pp. 3–39). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Chaiken, S., Duckworth, K. L., & Darke, P. (1999). When parsimony fails…. *Psychological Inquiry, 10*, 118–123.

Chaiken, S., & Eagly, A. H. (1983). Communication modality as a determinant of persuasion: The role of communicator salience. *Journal of Personality and Social Psychology, 45*, 241–256.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford Press.

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology, 66*, 460–473.

Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory and Cognition, 18*, 537–545.

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 73–96). New York: Guilford Press.

Cooper, J., & Fazio, R. H. (1984). A new look at dissonance theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17, pp. 229–266). New York: Academic Press.

Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2003). Neural components of social evaluation. *Journal of Personality and Social Psychology, 85*, 639–649.

Darke, P. R., Chaiken, S., Bohner, G., Einwiller, S., Erb, H. P., & Hazlewood, J. D. (1998). Accuracy motivation, consensus information, and the law of large numbers: Effects on attitude judgment in the absence of argumentation. *Personality and Social Psychology Bulletin, 24*, 1205–1215.

De Houwer, K., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127*, 853–869.

Dijksterhuis, A. (2002, September). *A subliminal road to happiness: Enhancing implicit self-esteem by subliminal evaluative conditioning*. Paper presented at the 4th European Social Cognition Network Meeting, Paris, France.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes.* San Diego, CA: Harcourt Brace.

Ebbesen, E. B., & Bowers, R. J. (1974). Proportion of risky to conservative arguments in a group discussion and choice shifts. *Journal of Personality and Social Psychology, 29*, 316–327.

Eiser, J. R., Fazio, R. H., Stafford, T., & Prescott, T. J. (2003). Connectionist simulation of attitude learning: Asymmetries in the acquisition of positive and negative evaluations. *Personality and Social Psychology Bulletin, 29*, 1221–1235.

Erb, H. P., Bohner, G., Schmälzle, K., & Rank, S. (1998). Beyond conflict and discrepancy: Cognitive bias in minority and majority influence. *Personality and Social Psychology Bulletin, 24*, 620–633.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 13, pp. 75–109). San Diego, CA: Academic Press.

Fazio, R. H., Powell, M. C. (1997). On the value of knowing one's likes and dislikes: Attitude accessibility, stress, and health in college. *Psychological Science, 8,* 430–436.

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.

Festinger, L. (1957). *A theory of cognitive dissonance.* Evanston, IL: Row, Peterson.

Fiedler, K. (1996). Explaining and simulating judgment biases as an aggregation phenomenon in a probabilistic, multiple-cue environment. *Journal of Personality and Social Psychology, 103*, 193–214.

Fiedler, K., Walther, E., & Nickel, S. (1999). The auto-verification of social hypotheses: Stereotyping and the power of sample size. *Journal of Personality and Social Psychology, 77*, 518.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. London: Addison Wesley.

Forgas, J. P. (2001). *Handbook of affect and social cognition*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Försterling, F. (1989). Models of covariation and attribution: How do they relate to the analogy of analysis of variance? *Journal of Personality and Social Psychology, 57*, 615–625.

French, R. (1997). Pseudo-recurrent connectionist networks: An approach to the "sensitivity–stability" dilemma. *Connection Science, 9*, 353–379.

French, R. (1999). Catastrophic forgetting in connectionist networks: Causes, consequences and solutions. *Trends in Cognitive Sciences, 3*, 128–135.

Gollwitzer, P. M. (1990). Action phases and mind-sets. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 53–92). New York: Guilford Press.

Greenwald, A. G. (1968). Cognitive learning, cognitive response to persuasion, and attitude change. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.), *Psychological foundations of attitudes* (pp. 147–170). San Diego, CA: Academic Press.

Haddock, G. (2000). Subjective ease of retrieval and attitude-relevant judgments. In H. Bless & J. P. Forgas (Eds.), *The message within: The role of subjective experience in social cognition and behavior* (pp. 125–142). Philadelphia: Taylor & Francis.

Haddock, G., Rothman, A. J., & Schwarz, N. (1996). Are (some) reports of attitude strength context dependent? *Canadian Journal of Behavioral Science, 28*, 313–316.

Hastie, R. (1980). Memory for behavioral information that confirms or contradicts a personality impression. In R. Hastie, T. M. Ostrom, E. B. Ebbesen, R. S. Wyer, D. L. Hamilton, & D. E. Carlston (Eds.), *Person memory: The cognitive basis of social perception* (pp. 155–177). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Haugtvedt, C. P., Schumann, D. W., Schneier, W. L., Warren, W. L. (1994). Advertising repetition and variation strategies: Implications for understanding attitude strength. *Journal of Consumer Research, 21*, 176–189.

Heesacker, M., Petty, R. E., Cacioppo, J. T. (1983). Field dependence and attitude change: Source credibility can alter persuasion by affecting message-relevant thinking. *Journal of Personality, 51*, 653–666.

Isen, A. M. (1984). Towards understanding the role of affect in cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 179–236). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ito, T. A., & Cacioppo, J. T. (2001). Affect and attitudes: A social neuroscience approach. In J. P. Forgas (Ed.), *Handbook of affect and social cognition* (pp. 50–74). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Jonas, K., Diehl, M., & Brömer, P. (1997). Effects of attitudinal ambivalence on information processing and attitude-intention consistency. *Journal of Experimental Social Psychology, 33*, 190–210.

Kaplan, K. J. (1972). On the ambivalence-indifference problem in attitude theory and measurement: A suggested modification of the semantic differential technique. *Psychological Review, 77*, 361–372.

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.

Kashima, Y., Woolcock, J., & Kashima, E. S. (2000). Group impression as dynamic configurations: The tensor product model of group impression formation and change. *Psychological Review, 107*, 914–942.

Katz, D., & Stotland, E. (1959). A preliminary statement to a theory of attitude structure and change. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 3, pp. 423–475). New York: McGraw-Hill.

Kelley, H. H. (1971). Attribution in social interaction. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 1–26). Morristown, NJ: General Learning Press.

Kruglanski, A. W., & Thompson, E. P. (1999). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry, 10*, 83–109.

Kruschke, J. K., & Johansen, M. K. (1999). A model of probabilistic category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1083–1119.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103*, 284–308.

LaBerge, D. (1997). Attention, awareness and the triangular circuit. *Consciousness and Cognition, 6*, 149–181.

LaBerge, D. (2000). Networks of attention. In M. S. Gazzaniga (Ed.), *The new cognitive neuroscience* (pp. 711–724). Cambridge, MA: MIT Press.

Lane, R. D., Reiman, E. M., Bradley, M. M., Lang, P. J., Ahern, G. L., Davidson, R. J., et al. (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia, 35*, 1437–1444.

Lieberman, M. D., Ochsner, K. N., Gilbert, D. T., & Schacter, D. L. (2001). Attitude change in amnesia and under cognitive load. *Psychological Science, 12*, 135–140.

Linder, D. E., Cooper, J., & Jones, E. E. (1967). Decision freedom as a determinant of the role of incentive magnitude in attitude change. *Journal of Personality and Social Psychology, 6*, 245–254.

MacDonald, T. K., & Zanna, M. P. (1998). Cross-validation ambivalence toward social groups: Can ambivalence affect intentions to hire feminists? *Personality and Social Psychology Bulletin, 24*, 427–441.

Mackie, D. M. (1987). Systematic and non-systematic processing of majority and minority persuasive communications. *Journal of Personality and Social Psychology, 53*, 41–52.

Mackie, D. M., & Worth, L. T. (1989). Processing deficits and the mediation of positive affect in persuasion. *Journal of Personality and Social Psychology, 57*, 27–40.

Maheswaran, D., & Chaiken, S. (1991). Promoting systematic processing in low-motivation settings: Effect of incongruent information on processing and judgment. *Journal of Personality and Social Psychology, 61*, 13–25.

Manis, M., Dovalina, I., Avis, N. E., & Cardoze, S. (1980). Base rates can affect individual predictions. *Journal of Personality and Social Psychology, 38*, 231–248.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

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

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology, 114*, 159–188.

McClelland, J. L., & Rumelhart, D. E. (1988). *Explorations in parallel distributed processing: A handbook of models, programs and exercises*. Cambridge, MA: Bradford.

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. *The Psychology of Learning and Motivation, 24*, 109–165.

McLeod, P., Plunkett, K., & Rolls, E. T. (1998). *Introduction to connectionist modeling of cognitive processes*. Oxford, UK: Oxford University Press.

Miller, N., & Colman, D. E. (1981). Methodological issue in analyzing the cognitive mediation of persuasion. In R. E. Petty, T. M. Ostrom, & T. C. Brock (Eds.), *Cognitive responses in persuasion* (pp. 105–125). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience. *American Psychologist, 56*, 717–734.

Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science, 12*, 413–417.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience*. Cambridge, MA: MIT Press.

O'Reilly, R. C., & Rudy, J. W. (2001). Conjunctive representations in learning and memory: Principles of cortical and hippocampal function. *Psychological Review, 108*, 311–345.

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General, 130,* 401–426.

Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion: Central and peripheral routes to attitude change*. New York: Springer.

Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology, 46*, 69–81.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123–205). San Diego, CA: Academic Press.

Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology, 41*, 847–855.

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research, 10*, 135–147.

Petty, R. E., Schumann, D. W., Richman, S. A., & Strathman, A. J. (1993). Positive mood and persuasion: Different roles for affect under high- and low-elaboration conditions. *Journal of Personality and Social Psychology, 64*, 5–20.

Petty, R. E., & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 41–72). New York: Guilford Press.

Posner, M. I. (1992). Attention as a cognitive and neural system. *Current Directions in Psychological Science, 1*, 1114.

Powell, M. C., & Fazio, R. H. (1984). Attitude accessibility as a function of repeated attitudinal expression. *Personality and Social Psychology Bulletin, 10,* 139–148.

Priester, J. R., & Petty, R. E. (1996). The gradual threshold model of ambivalence: Relating the positive and negative bases of attitudes to subjective ambivalence. *Journal of Personality and Social Psychology, 71*, 431–449.

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

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

Ratneshwar, S., & Chaiken, S. (1991). Comprehension's role in persuasion: The case of its moderating effect on the persuasive impact of source cues. *Journal of Consumer Research, 18*, 52–62.

Read, S. J., & Marcus-Newhall, A. (1993). Explanatory coherence in social explanations: A parallel distributed processing account. *Journal of Personality and Social Psychology, 65*, 429–447.

Read, S. J., & Miller, L. C. (1998). *Connectionist models of social reasoning and social behavior*. New York: Lawrence Erlbaum Associates, Inc.

Read, S. J., & Montoya, J. A. (1999). An autoassociative model of causal reasoning and causal learning: Reply to Van Overwalle's critique of Read and Marcus-Newhall (1993). *Journal of Personality and Social Psychology, 76*, 728–742.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–98). New York: Appleton-Century-Crofts.

Riketta, M., & Dauenheimer, D. (2002). *Manipulating self-esteem through subliminally presented words*. Manuscript submitted for publication.

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective and behavioral components of attitudes. In C. I. Hovland & M. J. Rosenberg (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 1–14). Hew Haven, CT: Yale University Press.

Roskos-Ewoldsen, D. R., Bichsel, J., & Hoffman, K. (2002). The influence of accessibility of source likeability on persuasion. *Journal of Experimental Social Psychology, 38*, 137–143.

Rumelhart, D., & McClelland, J. (1986). On learning the past tense of English verbs: Implicit rules or parallel distributed processing? In J. M. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing* (Vol. 2, pp. 216–271). Cambridge, MA: Bradford.

Schuette, R. A., & Fazio, R. H. (1995). Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model. *Personality and Social Psychology Bulletin, 21*, 704–710.

Schumann, D. W., Petty, R., & Clemons, D. S. (1990). Predicting the effectiveness of different strategies of advertising variation. A test of the repetition-variation hypothesis. *Journal of Consumer Research, 17*, 192–201.

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 527–561). New York: Guilford Press.

Schwarz, N., Bless, H., & Bohner, G. (1991). Mood and persuasion: Affective states influence the processing of persuasive communications. *Advances in Experimental Social Psychology*, *24*, 161-199.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.

Shanks, D. R. (1985). Forward and backward blocking in human contingency judgment. *Quarterly Journal of Experimental Psychology, 37b*, 1–21.

Shanks, D. R. (1987). Acquisition functions in contingency judgment. *Learning and Motivation, 18*, 147–166.

Shanks, D. R., Lopez, F. J., Darby, R. J., & Dickinson, A. (1996). Distinguishing associative and probabilistic contrast theories of human contingency judgment. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation* (Vol. 34, pp. 265–311). New York: Academic Press.

Shultz, T., & Lepper, M. (1996). Cognitive dissonance reduction as constraint satisfaction. *Psychological Review, 2*, 219–240.

Siebler, F. (2002). *Connectionist modeling of social judgment processes*. Unpublished doctoral dissertation, University of Kent, Canterbury, United Kingdom.

Sinclair, R. C., Mark, M. M., & Clore, G. L. (1994). Mood-related persuasion depends on (mis)attributions. *Social Cognition, 12*, 309–326.

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

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

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review, 4*, 108–131.

Spellman, B. A., & Holyoak, K. J. (1992). If Saddam is Hitler who is George Bush? Analogical mapping between systems of social roles. *Journal of Personality and Social Psychology, 62*, 913–933.

Staats, A. W., & Staats, C. K. (1958). Attitudes established by classical conditioning. *Journal of Abnormal and Social Psychology, 57*, 37–40.

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin, 111*, 42–61.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*, 220–247.

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

Thorpe, S. (1994). Localized versus distributed representations. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 949–952). Cambridge, MA: MIT Press.

Tormala, Z. L., Petty, R. E., Briñol, P. (2002). Ease of retrieval effect in persuasion: A self-validation analysis. *Personality and Social Psychology Bulletin, 28*, 1700–1712.

Van Duynslaeger, M., & Van Overwalle, F. (2004) *Do persuasive heuristics need abstract rules?* Unpublished data, Free University of Brussels, Belgium.

Van Overwalle, F. (1998) Causal explanation as constraint satisfaction: A critique and a feedforward connectionist alternative. *Journal of Personality and Social Psychology, 74*, 312–328.

Van Overwalle, F., & Jordens, K. (2002). An adaptive connectionist model of cognitive dissonance. *Personality and Social Psychology Review, 6*, 204–231.

Van Overwalle, F., & Labiouse, C. (2004). A recurrent connectionist model of person impression formation. *Personality and Social Psychology Review, 8*, 28–61.

Van Overwalle, F., & Van Rooy, D. (1998). A connectionist approach to causal attribution. In S. J. Read & L. C. Miller (Eds.), *Connectionist models of social reasoning and social behavior* (pp. 143–171). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Van Rooy, D., Van Overwalle, F., Vanhoomissen, T., Labiouse, C., & French, R. (2003). A recurrent connectionist model of group biases. *Psychological Review, 8*, 28–61.

Wänke, M., & Bless, H. (2000). The effects of subjective ease of retrieval on attitudinal judgments: The moderating role of processing motivation. In H. Bless & J. P. Forgas (Eds.), *The message within: The role of subjective experience in social cognition and behavior* (pp. 143–161). Philadelphia: Taylor & Francis.

Wänke, M., Bless, H., & Biller, B. (1996). Subjective experience versus content of information in the construction of attitude judgments. *Personality and Social Psychology Bulletin, 22*, 1105–1113.

Wänke, M., Bohner, G., & Jurkowitsch, A. (1997). There are many reasons to drive a BMW: Does imagined ease of argument generation influence attitudes? *Journal of Consumer Research, 24*, 170–177.

Wegener, D. T., & Petty, R. E. (1996). Effects of mood on persuasion processes: Enhancing, reducing and biasing scrutiny of attitude-relevant information. In L. L. Martin & A. Tesser (Eds.), *Striving and feeling: Interactions among goals, affect and self-regulation* (pp. 329–362). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107*, 101–126.

Wood, W., Kallgren, C. A., Preisler, R. M. (1985). Access to attitude-relevant information in memory as a determinant of persuasion: The role of message attributes. *Journal of Experimental Social Psychology, 21*, 73–85.

Worth, L. T., & Mackie, D. M. (1987). Cognitive mediation of positive affect in persuasion. *Social Cognition, 5*, 76–94.

Zanna, M. P., Kiesler, C. A., & Pilkonis, P. A. (1970). Positive and negative attitudinal affect established by classical conditioning. *Journal of Personality and Social Psychology*, *14*, 321–328.

Appendix A: The Linear Autoassociative Model

In an autoassociative network, concepts are represented in nodes that are all interconnected. Processing information in this model takes place in two phases. In the first phase, the activation of the nodes is computed, and in the second phase, the weights of the connections are updated (see also McClelland & Rumelhart, 1988).

Node Activation

During the first phase of information processing, each node in the network receives activation from external sources. Because the nodes are all interconnected, this activation is then spread throughout the network where it influences all other nodes. The activation coming from the other nodes is called the internal activation. Together with the external activation, this internal activation determines the final pattern of activation of the nodes, which reflects the short-term memory of the network.

In mathematical terms, every node i in the network receives external activation, termed ext_i. In the autoassociative model, every node i also receives internal activation int_i, which is the sum of the activation from the other nodes j (denoted by a_j) in proportion to the weight of their connection w_{ij}, or

$$int_i = \Sigma(a_j * w_{ij}), \tag{1}$$

for all $j \neq i$. Typically, activations and weights range approximately between –1 and +1. The external activation and internal activation are then summed to the net activation, or

$$net_i = E * ext_i + I * int_i, \tag{2}$$

where E and I reflect the degree to which the net activation is determined by the external and internal activation, respectively. In a recurrent network, the activation of each node i is updated during a number of cycles until it eventually converges to a stable pattern that reflects the network's short-term memory. According to the linear activation algorithm (McClelland & Rumelhart, 1988, p. 167), the updating of activation is governed by the following equation:

$$\Delta a_i = net_i - D * a_i, \tag{3}$$

where D reflects a memory decay term. In the present simulations, we used the parameter values $D = I = E = 1$. Given these simplifying parameters, the final activation of node i reduces simply to the sum of the external and internal activation, or:

$$a_\mathrm{i} = net_i = ext_i + int_i \tag{3'}$$

Weight Updating

After this first phase, the autoassociative model enters into its second learning phase, where the short-term activation is consolidated in long-term weight changes to better represent and anticipate future external activation. Basically, weight changes are driven by the discrepancy between the internal activation from the last updating cycle of the network and the external activation received from outside sources, formally expressed in the delta algorithm (McClelland & Rumelhart, 1988, p. 166):

$$\Delta w_{ij} = \varepsilon\ (ext_i - int_i)\ a_j, \tag{4}$$

where Δw_{ij} is the weight of the connection from node j to i, and ε is a learning rate that determines how fast the network learns.

Generating Evaluative Reactions

As noted in the text, to generate evaluative responses that the network recognizes as genuine, the internal activation generated at the valence nodes is taken as external activation. First, the internal activation arriving at the valence nodes is computed by Equation 1. Next, this internal activation is further "boosted" toward the extremes of +1 and –1 by running 10 internal cycles of the nonlinear activation updating algorithm. In mathematical terms,

$$\text{if } net_i > 0 \text{ then } \Delta a_i = net_i\ (1 - a_i) - D * a_i \tag{5a}$$

$$\text{if } net_i < 0 \text{ then } \Delta a_i = net_i\ (a_i + 1) - D * a_i \tag{5b}$$

where D reflects a memory decay term. After cycling 10 times, the resulting activation a_i is then taken as external activation.

Appendix B: Fishbein and Ajzen's (1975) Model and the Delta Algorithm

This appendix demonstrates that the delta algorithm converges at asymptote to the expectancy-value model of attitude formation by Fishbein and Ajzen (1975; Ajzen, 1991). According to this model, an attitude is formed by summing the multiplicative combination of (a) the strength of a salient belief that a behavior will produce a given outcome and (b) the subjective evaluation of this outcome, or (Ajzen, 1991, p. 191):

$$\text{attitude} \approx \Sigma b_i e_i, \tag{6}$$

were b_i represents the strength of the belief and e_i the evaluation. Beliefs and evaluations are typically scored on 7-point scales. Although Fishbein and Ajzen (1975) suggest that the integration (of the multiplication of beliefs and evaluations) is a summative process, they acknowledge that evidence in favor of summation versus averaging is rather inconsistent and inconclusive (pp. 234–235). Moreover, to prevent their summative function to grow out of bounds, they restrict their formula to salient beliefs about an attitude object (typically not more than 10). Because of this implicit boundary assumption and because there is "no rational a priori criterion we can use to decide how the belief and evaluation scales should be scored" (Ajzen, 1991, p. 193), the preceding formula can be normalized by dividing it by the mean belief strengths, or:

$$\text{attitude} \approx \Sigma b_i e_i / \Sigma b_i \tag{7}$$

This proof uses the same logic as Chapman and Robbins (1990) in their demonstration that the delta algorithm converges to the probabilistic expression of covariation. In line with the conventional representation of covariation information, attitude relevant information can be represented in a contingency table with two cells. Cell a represents all cases where the attitude object is followed by a given (positive) evaluation, and cell b represents all cases where the same object is followed by the opposite (negative) evaluation. For simplicity, we use only an object with a single expectation or belief although this proof can easily be extended to multiple beliefs.

In a recurrent connectionist architecture with localist encoding, the object j and the evaluation i are each represented by a node, which are connected by adjustable weights w_{ij}. We use a localist encoding to simplify the proof. When the object is present, its corresponding node receives external activation, and this activation is spread to both valence nodes. As defined in the text, we assume that the overall internal activation received at the valence nodes i after priming the object node j reflects the attitude.

According to the delta algorithm in Equation 4, the weights w_{ij} are adjusted proportional to the error between the actual evaluation (represented by its external activation *ext*) and the evaluation as predicted by the network (represented by its internal activation *int*). If we take the default activation for a_j (which is 1), then the following equations can be constructed for the two cells in the contingency table:

$$\text{For the a cell: } \Delta w_{ij} = \varepsilon(e_1 - int)_, \tag{8}$$

$$\text{For the b cell: } \Delta w_{ij} = \varepsilon(e_2 - int)_. \tag{9}$$

Note that e_1 reflects a positive evaluation and e_2 a negative evaluation. The change in overall attitude is the sum of Equations 8 and 9 weighted for the corresponding frequencies a and b, in the two cells, or:

$$\Delta w_{ij} = a[\varepsilon(e_1 - int)] + b[\varepsilon\ (e_2 - int)] \tag{10}$$

These adjustments will continue until asymptote, that is, until the error between actual and expected category is zero. This implies that at asymptote, the changes will become zero, or $\Delta w_{ij} = 0$. Consequently, Equation 8 becomes

$$0 = a[\varepsilon(e_1 - int)] + b[\varepsilon(e_2 - int)]$$
$$= a[e_1 - \mathrm{int}] + b[e_2 - int]$$
$$= [a * e_1 + b * e_2] - [a + b]\ int$$

so that

$$int = [a * e_1 + b * e_2] / [a + b],$$

As noted earlier, the internal activation *int* received at the valence nodes after priming the object node reflects the attitude. Because e_1 was expressed in positive terms and e_2 in negative terms, the equation reflects the differential internal activation of the favorable and unfavorable valence nodes. Hence, the left side of the equation can simply be interpreted as the attitude. In addition, the right side of the equation can be rewritten in Fishbein and Ajzen's (1975) terms as

$$\text{attitude} = \Sigma f_i e_i / \Sigma f, \quad (11)$$

where *f* represents the frequency that the attitude object leads to a given outcome and evaluation (which we assume determine the belief strength *b*). The equivalence between Equations 7 and 11 demonstrates that the delta algorithm predicts a (normalized) multiplicative function at asymptote for making attitude judgments, where the strength of the beliefs is determined by the frequencies by which the attitude object and evaluations co-occur.

Note that although the delta algorithm predicts an averaging multiplicative function after a large amount of input (i.e., at asymptote), in its beginning phase, the algorithm actually predicts an additive function. At the start of learning, every new piece of information results in relatively substantial weight adjustments because the error is still large. The more learning occurs, the greater the likelihood that the error decreases, so that novel information has less effect and is integrated with older information, resulting in a sort of averaging of earlier nd recent novel input.