Short Research Note

Expectancy confirmation in attitude learning: A connectionist account

J. RICHARD EISER[1]*, TOM STAFFORD[1] AND RUSSELL H. FAZIO[2]
[1]*University of Sheffield, UK*
[2]*Ohio State University, USA*

Abstract

Connectionist computer simulation was employed to model how learners acquire attitudes towards novel objects under conditions where (a) they are given prior expectancies that the objects as a whole are mostly good or mostly bad; and (b) they can only discover the true valence of the objects by approaching them. Expectancy confirmation was operationalized through modifying connection weights more after experiencing good than bad objects (positive bias), or more after experiencing bad than good objects (negative bias). Negative bias led the network to misclassify more good objects as bad, such negative attitudes resisting change because of the lack of corrective feedback relating to avoided objects. Conversely, positive bias encouraged approach and hence feedback leading to more accurate discrimination of good and bad objects, as well as to higher estimates of the valence of objects not presented during training. These findings suggest that expectancy confirmation may emerge "automatically" from basic learning processes. Copyright © 2008 John Wiley & Sons, Ltd.

Compared with the large literature on the consequences of attitudes, especially for behavior and interpersonal relations, the antecedents of attitudes have received surprisingly little attention. In particular, rather few studies have considered how attitudes are shaped by learning experiences derived from direct interaction with objects in the environment. In simple terms, we assume that, as individuals explore their environment, they will learn, among other things, which objects are good and which are bad, and this feedback will lead them to form positive and negative attitudes toward these objects. At the same time, however, the attitudes that individuals have formed will guide their exploration. In other words, they will be more likely to choose to approach and engage with objects (or persons or activities) to which they have formed more positive attitudes. Thus, we may have a positive attitude toward holidaying in a given country as a consequence of previous good holidays there. Because of this positive attitude, we will also be keener to go there, and hence gain even more experience on which to base our attitudes. Hence, the relationship between experience and attitude formation is essentially dynamic: attitudes are formed on the basis of feedback experiences, while at the same time such feedback experiences depend on behaviors that are guided by attitudes.

One consequence of this dynamic relationship is an asymmetry in the learning of good and bad objects. If learners believe that a given object is good and attractive, they will be more like to seek out experiences with it, and hence obtain yet more information about it. If they believe it to be bad and unattractive, they will be more likely to avoid it, but doing so will limit the information they gain about the object's true valence, and hence fail to learn whether their initial negative evaluations were correct or incorrect. The implication is that learners are more likely to misclassify good objects as bad than bad objects as good.

*Correspondence to: J. Richard Eiser, Department of Psychology, University of Sheffield, Sheffield S10 2TP, UK. E-mail: j.r.eiser@shef.ac.uk

Evidence of such ''learning asymmetry'' comes from experimental (Fazio, Eiser, & Shook, 2004; Shook, Fazio, & Eiser, 2007) and simulation (Eiser, Fazio, Stafford, & Prescott, 2003) studies, in which participants (or neural networks) have to learn to distinguish good from bad novel objects. The experimental task involves participants playing a computer game (''BeanFest'') in which their success depends on finding ''good beans'' that provide energy while avoiding ''bad beans'' that cost energy. The beans differ in terms of two visual dimensions (shape and speckledness), such that there are three separate clusters of good beans and three of bad beans, that together fill 36 of the possible 100 cells of the two-dimensional matrix shown in Figure 1. Each time an object is presented, learners must choose whether to approach (''eat'') or avoid it. If they approach a good object, they are rewarded. If they approach a bad object, they are punished. In the crucial (''contingent feedback'') conditions, learners only discover the valence of objects they approach, whereas avoidance provides no information. Hence, learning depends on approach, but approaching unfamiliar objects involves risk.

One strategy for reducing risk is to avoid objects of uncertain valence (like never trying a new destination for a holiday) but this also reduces the possibility of novel rewards. In real life, however, we can reduce the risk of exploration through accepting advice from others. If we are planning a holiday, a guide-book could be useful. Such socially communicated information, or *indirect* experience, provides us with evaluative expectancies to guide our behavior even before we have any *direct* experience of our own. Indirect experience has been demonstrated to influence attitude learning. In Experiment 5 of Fazio et al. (2004), participants received written ''advice,'' supposedly from previous players, that a specific subclass of objects was good or bad. The main finding was that incorrect negative beliefs or ''prejudice'' tended to be maintained through to the end of the game. This was because incorrect negative advice misled participants to avoid, and so never discover the true valence of, beans that were actually good.

Adopting a similar procedure, Eiser, Shook, and Fazio (2007) had participants play the BeanFest game after advice either that ''most beans are good'' (positive advice) or that ''most beans are bad'' (negative advice). (In fact, half the beans were good and half, bad.) Participants who avoided more beans following negative advice were more likely to misclassify good beans as bad (''learning asymmetry''). At the end of the game, when asked to estimate the valence of beans not previously presented, these same participants were particularly likely to assume that these new beans were bad rather than good (''generalization asymmetry''). These findings imply that the advice influenced participants' prior *expectancies*. However, since such advice related to the objects as a whole (rather than a specific subset, as in Fazio et al., 2004, Experiment 5), it did not help participants choose which objects to approach and which to avoid. Hence, even when the advice was negative, participants still had to approach sufficient objects to survive and so receive some positive experiences that were inconsistent with their negative expectancies. Likewise, participants who followed positive advice and approached more beans also received expectancy-inconsistent feedback whenever they encountered a bad bean. This
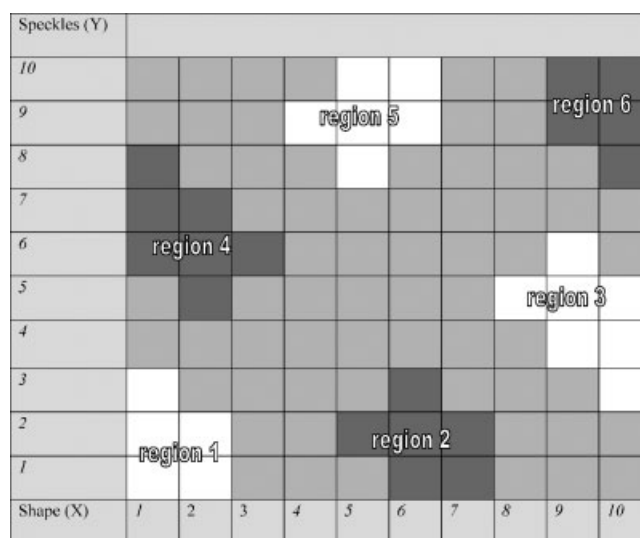

Figure 1.   The BeanFest matrix. Clear squares (regions 1, 3, and 5) represent ''good beans'' and dark gray squares (regions 2, 4, and 6) represent ''bad beans''

raises the question of whether feedback influences attitude learning differently depending on its consistency with participants' expectancies.

There is general evidence that our interpretation of ambiguous information can be influenced by expectancy-confirmation biases (e.g., Darley & Fazio, 1980; Darley & Gross, 1983). In other words, information is interpreted in the light of pre-existing hypotheses or expectancies and is more likely to be accepted as true if it confirms prior assumptions. This raises questions about how the impact of expectancies should be conceptualized within the learning process. The effects of direct feedback may arguably be viewed in terms of familiar principles of reinforcement (Sutton & Barto, 1998). However, the use of phrases such as "interpreted in the light of" and "accepted as true" to explain the effects of expectancy confirmation, this implies that a form of more deliberative, presumably conscious, reasoning is involved. Is this implication correct? Evidently, human beings are *capable* of deliberative, expectancy-based reasoning (and of expressing prejudice symbolically), but do expectancy-confirmation biases actually *require* the involvement of such higher-order mental processes, or can they "fall out," more or less automatically, from the processing of learning experiences at a more basic level?

One way of addressing this question (without debating broader issues surrounding the notion of automaticity, e.g., Bargh & Ferguson, 2000, Fazio, 2001) is to investigate whether expectancy confirmation can be produced "automatically" within an artificial learning system by definition incapable of conscious reasoning, if we make some minimal assumptions about the relative weighing of particular kinds of information. In other words, if an "automaton" can show conceptually similar biases, this supports the argument that expectancy confirmation may become automaticized in humans. The present study therefore employs the methodology of connectionist computer simulation to examine how expectancies and experience might combine within a single learning system designed to model the acquisition of attitudes or evaluative judgments. Since this technique has hitherto been mainly used to consider the effect of direct experience only (but see Van Overwalle & Heylighen, 2006), additional assumptions are needed to incorporate the influence of expectancies or indirect experience. By definition, connectionist learning systems cannot be "aware" of another's mental state, but this does not entail that minds (whether real or artificial) are *tabulae rasae* that process all inputs equally, rather than being selectively attentive to, and affected by, certain kinds of stimuli. Hence, remembering connectionism's aspirations to "biological plausibility," and broader arguments that we may inherit predispositions to derive particular beliefs or "memes" from our experience (Dawkins, 1976; Dennett, 1991), we see no principled objection to particular learning outcomes having different prior probabilities. Moreover, granted the evolutionary advantages of imitation, even primitive learning systems may exhibit biases towards reproducing the action tendencies or "preferences" of conspecifics.

A straightforward way of simulating biased processing of this kind is through controlling the rate at which a learning system updates its representations depending on the particular kinds of information presented to it. In human terms, stated informally, the assumption is that the learner places greater reliance or trust in information or feedback that is consistent with socially communicated expectancies than with information that is inconsistent. Stated more formally, feedback that is consistent with expectancies should influence learning more strongly than feedback inconsistent with expectancies.[1]

Connectionist Simulation: Basic Principles

Our simulations employ a form of connectionist neural network. A neural network, or net, consists of interconnected nodes or units. Each unit has a level of activation that depends on external stimuli and/or information received from other units. Input units receive activations from presented stimuli. These activations are then fed forward through a layer of "hidden units" to an output layer, reflecting the decisions or predictions made by the net. How much activation spreads from one unit to another depends on the strength, or weight, of the connection between them. Crucially, such connection weights are modified through training. For example, when the net receives feedback on the correctness of any decision it makes on the basis of some input, a learning algorithm is applied in order to modify the connection weights so as to improve the net's performance on subsequent trials. Commonly used learning algorithms involve calculating the error or discrepancy (Δ)

[1]Note that this view of expectancy confirmation is entirely compatible with the well-known finding that expectancy-incongruent information attracts more attention and tends to be better remembered than expectancy-congruent information (Stangor & McMillan, 1992). The fact that incongruent information is better remembered does not entail that it will have more impact on learning, rather that more 'cognitive work' will be needed to reconcile it with prior beliefs. In this sense, expectancy-confirmation biases are one of a number of devices that might be used to protect prior beliefs by discounting the force of troublesome evidence (see e.g., Abelson, 1959; Janis & Mann, 1977).

between the net's prediction and the "correct" target value for a given input, and adjusting the connection weight so that Δ is reduced on subsequent presentation of that input. To control the extent of modification to the weights made on any given trial, Δ is typically multiplied by a constant (<1) termed the "learning rate" parameter. This parameter is usually set so that learning improves in very small increments over hundreds or thousands of trials or "epochs." However, the choice of any given learning rate is typically a pragmatic rather than theoretically inspired one. A novel feature of the simulations to be reported here is that we effectively re-scaled the learning rate so that learning (connection weight adjustment) was accelerated or decelerated, conditional on the type of feedback received by the net and its consistency with pre-defined expectancies.

METHOD

Network Architecture

We employed the neural network architecture as shown in Figure 2, and described in more detail in Eiser et al. (2003; Study 2). This comprises a separate learning system and action selection mechanism. The *learning system* is a fully connected, three-layer, feed-forward network with an input layer to encode the stimulus attributes, a hidden layer that produces a condensed representation of this input information, and a single output unit that represents an "evaluation" of each input pattern (bean) on a continuum from "bad" (0) to "good" (1). The input layer has 22 units, of which 11 are used to encode one dimension (representing shape) and the remaining 11 the other dimension (representing speckles). The encoding is arranged (see Eiser et al., 2003) so that ("distributed") patterns of activation for stimuli that are more similar to each other in terms of their "visual" attributes (i.e., closer to each other within the matrix) overlap. This allows the network to form more generalized representations of good and bad "areas," in other words, to assign similar valence to objects that are "visually" similar.
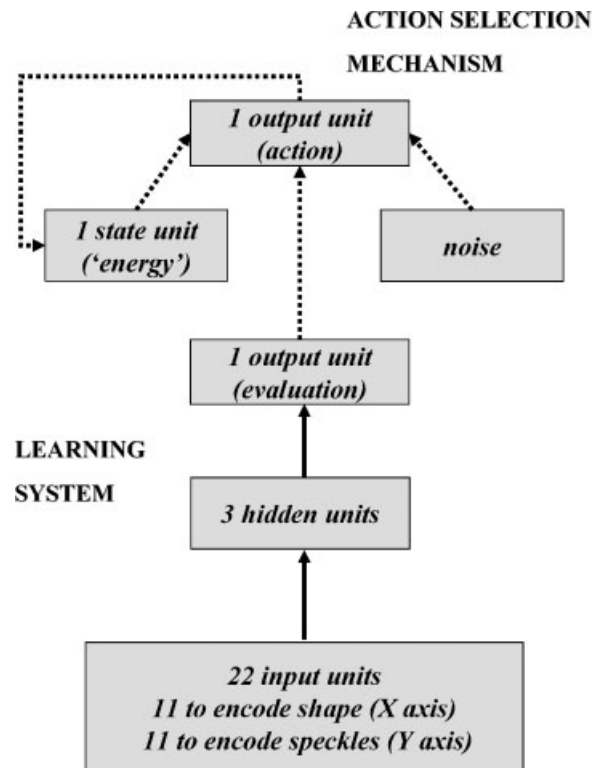

Figure 2.   Network architecture. Solid lines indicate connections modified by training

The purpose of the *action selection mechanism* is to produce a "choice." A logistic function is first applied to the evaluation generated by the learning system, meaning that, as an evaluation moves away from neutral (0.5), it is rapidly converted into a quite polarized value (i.e., close to 0 or 1). This converted evaluation, or "judgment" is then fed into the action selection mechanism along with additional activation from a state unit ("energy") that essentially provides a running count of the impact of previous gains or losses in energy as a result of "eating" good or bad beans, and losses due to time. The initial value of this "energy unit" was set at 1 and varied between 0 and 1. There is a steady decay in the activation of the energy unit, at the rate of 0.0001 per bean presentation. Eating a good bean increases the activation of the energy unit by 0.001 and *eating* a bad bean decreases it by 0.001 (whereas *avoiding* a bad bean produces only the time-related decay of 0.0001). The input to the action selection mechanism from this energy unit is determined by a "hunger function" (i.e., the "neutral hunger" function in Eiser et al., 2003, Study 2, p. 1230), so that hunger increases to an asymptote of 1 as energy drops to 0, but takes a minimum value of 0 when energy is at its maximum 1.

Finally, a stochastic probability function ("noise") is introduced. This effectively produces a random variation in the threshold value (above or below 0.5) against which the final outcome activation is compared, so as to determine whether the network will "eat" (activation above threshold) or not. The effect of this stochastic element is that the network will occasionally "eat" beans it has categorized as "bad" (i.e., evaluated as <0.5), and avoid beans categorized as "good." The purpose of this is to introduce an element of chance into the net's decision-making and hence feedback experience, rather than making its approach and avoidance behaviors entirely determined by prior attitudes.

Training Procedure

This was based on the standard back propagation of error ("backprop") algorithm (Rumelhart, Hinton, & Williams, 1986) for modifying the connection weights. Within each epoch, the network is presented, in a random order, with the 36 input patterns corresponding to the 18 good and 18 bad "beans" shown in Figure 1. The learning system of the network then generates a judgment (between 0 and 1) representing its assessment of how bad or good each bean is. This judgment is then compared with a "training signal" or "target value," defined as 1 for a "good" bean and 0 for a "bad" bean and the error or Δ (target minus judgment) is then calculated. This is used to modify the connections within the network so as to reduce the discrepancy between judgments and target values on subsequent trials. In all simulations to be described, training proceeded for 5000 epochs, although most modification of the weights occurred during the first few hundred epochs. The following variations of this procedure were compared.

Full versus Contingent Feedback

Under "full feedback" (essentially a control condition), Δ was calculated and connection weights were modified both when the net "chose to eat" (i.e., when the action selection mechanism produced an output equal to or greater than threshold) *and* when it did not. This represents a situation where someone makes a prediction about a bean's value and then is always told whether the bean is good or bad. Under "contingent feedback," connection weights were only modified if the selected action was equal to, or greater than, threshold–i.e., if the net had "chosen to eat." No learning (modification of weights) took place on any trial where the output was below threshold. This represents a situation where someone has to eat a bean to discover whether it is good or bad. If a bean is avoided, no information is provided about its true value and so no learning takes place on that trial.

Simulating Expectancy Confirmation

This was achieved by incorporating a "bias" parameter (G). Along with the learning parameter (set at 0.02), this had the effect of re-scaling Δ before adjustment of the connection weights. Our approach involved multiplying Δ by G (where $G \geq 1$), thereby accelerating learning, on all "consistent" trials (i.e., where feedback was consistent with prior expectancies) and dividing Δ by G, thereby decelerating learning, on all "inconsistent" trials. Using this approach, we produced a *negative* bias condition, in which Δ was multiplied by G following experience with a bad bean and by $1/G$

following experience with a good bean. Under *positive* bias, Δ was multiplied by *G* following experience with a good bean and by 1/*G* following experience with a bad bean. A *neutral* condition is included in all simulations to be described, in which Δ is not multiplied by *G*.

We set *G* either to a "strong" level of 1.5 or a "mild" level of 1.2 for both positive and negative bias. Together with a neutral condition, this yielded a 2 × 5 × 2 (Feedback × Bias × Valence) factorial with repeated measures on the last factor. Valence refers simply to whether the presented input patterns (beans) were good or bad. Feedback was either full (connection weights modified by back propagation of error on every trial regardless of selected action) or contingent (connection weights modified only if the selected action corresponded to approach, i.e., "eating" a bean).

Analyses were based on 10 completed replications (i.e., runs with different randomly initialized weights) in each cell, with training continuing for 5000 epochs. The term "completed" signifies the fact that some networks might "die"—i.e., run out of energy—before learning how to discriminate the good and bad beans adequately. There were three such "deaths" (all under contingent feedback, mild negative bias) for which replacement simulations were run with new random settings. The data to be reported describe network performance both early in training, i.e., at 200 epochs, and at the conclusion of training at 5000 epochs. *Discrimination* is measured by considering the outputs of the learning system, i.e., the raw "evaluations" of the 36 beans presented during training (recall that these could range between 0 and 1) before their transformation by the action selection mechanism. These are treated as follows. A judgment of 0.5 or above is defined as representing a positive evaluation or attitude toward a specific bean, and a judgment below 0.5 is defined as a negative evaluation. We then counted the number of bad beans out of 18 that were evaluated negatively and converted this to a proportion score between 0 and 1, termed *P bad correct*. We then calculated a similar proportion score for the number of good beans evaluated positively, termed *P good correct*. In addition, to see how the network *generalized* from its learning of the presented input patterns to the combination of attributes in the remaining 64 cells of the matrix, the network was presented with these 64 untrained patterns as inputs and the outputs of the learning system were recorded (without any subsequent modification of the weights). The mean outputs of the learning system were then averaged over all these 64 novel stimuli and recoded from −1 (bad) to +1 (good) so as to be directly comparable with the human data reported by Fazio et al. (2004).

With respect to *discrimination* we anticipated that the bias would affect the learning asymmetry through influencing the rate at which the net's evaluations of the objects are modified by feedback. Positive bias leads to a faster acquisition of positive evaluations following positive feedback, but slower acquisition of negative evaluations following negative feedback. This increases the likelihood of objects being sampled and so the likelihood of the net receiving feedback to rectify false-negative evaluations. Hence, the learning asymmetry effect should be reduced. Negative bias leads to a faster acquisition of negative evaluations following negative feedback, but slower acquisition of positive evaluations following positive feedback. This should lead to fewer objects being sampled, and hence a maintenance or strengthening of the learning asymmetry effect. With respect to *generalization*, we anticipated that (except in the full feedback conditions, where more complete learning could eliminate the effects of the bias manipulation), the estimates made by the net concerning the 64 untrained patterns would be lowest in the negative bias conditions, and highest in the positive bias conditions. Furthermore, we anticipated that, under contingent but not full feedback, generalization scores would tend, overall, to be negative. This "generalization asymmetry" has been found in previous simulations (Eiser et al., 2003), reflecting the fact that contingent feedback typically leads to lower exposure to positive test stimuli. It has also been found with human participants (Fazio et al., 2004), where the data indicate that negative information may be more highly weighted than positive.

RESULTS

Discrimination

Table 1 shows the proportion of correct evaluations of good and bad stimuli within each condition early in, and after, training. Once again, full feedback leads to perfect or near-perfect identification of both bad and good stimuli, though with two interesting exceptions. First, under strong negative bias, identification of good stimuli is close to chance at 200 epochs (though perfect by 5000 epochs). Second, under strong positive bias, the identification of bad stimuli is only 76% correct at

Table 1. Mean proportions of correct responses to bad and good stimuli and generalization index at 200 and 5000 epochs, as a function of feedback and bias

Feedback	Bias	200 epochs			5000 epochs		
		P bad correct	P good correct	Generalization	P bad correct	P good correct	Generalization
Full	Strong Negative	.98	.56	−.23	1	1	.02
	Mild Negative	.97	.97	−.02	1	1	.02
	Neutral	.98	.97	.02	1	1	.05
	Mild Positive	1	1	.03	1	1	−.02
	Strong Positive	.76	1	.26	.90	1	.19
Contingent	Strong Negative	.99	.13	−.40	1	.70	−.39
	Mild Negative	.96	.61	−.33	1	.67	−.45
	Neutral	.98	.67	−.31	1	.67	−.44
	Mild Positive	.91	.97	.21	1	1	.15
	Strong Positive	.61	.98	.35	1	1	.33

200 epochs (90% at 5000 epochs). This suggests that the bias manipulation (at least when "strong") influences early training even under full feedback. Eventually, however, the full feedback enables the network to overcome this initial bias.

However, when feedback is contingent on the network selecting an approach response, the pattern of results is unambiguous. Already by 200 epochs, there is a strong main effect for Valence, $F(1,45) = 100.51$, $p < .001$, confirming an overall learning asymmetry effect; a Bias × Valence interaction, $F(4,45) = 91.76$, $p < .001$, reflecting near-perfect responses to bad beans under negative bias and to good beans under positive bias; and a Bias main effect, $F(1,45) = 28.02$, $p < .001$. (Once again, note the very low mean for good beans under strong negative bias, and the relatively low mean for bad beans under strong positive bias). By the end of training, identification of bad beans is perfect, whereas errors occur relatively frequently to good beans (Valence: $F(1,45) = 841.00$, $p < .001$) unless the network is trained under a positive bias (Bias × Valence: $F(4,45) = 141.00$, $p < .001$). Hence, learning asymmetry is reduced by a positive bias, that is, by the network modifying the connection weights at a higher rate following positive than negative feedback.

Generalization

Next, we considered the generalization scores derived for the network's responses to the 64 untrained input patterns (see Table 1). Considering just the contingent feedback conditions, the effect of bias is highly significant both at 200 epochs, $F(4,45) = 274.23$, $p < .001$, and at 5000 epochs, $F(4,45) = 133.33$, $p < .001$. However, the main differences are between the two positive bias and the three remaining conditions. (The means for the strong negative and neutral conditions differ significantly, $p = .003$, at 200 but not 5000 epochs). The overall generalization asymmetry is highly significant both at 200 epochs, $F(1,45) = 107.22$, $p < .001$, and 5000 epochs, $F(1,45) = 123.15$, $p < .001$. In order to see whether this generalization asymmetry was a by-product of the learning asymmetry, we performed a hierarchical regression analysis to predict the generalization score at 5000 epochs from Bias at step 1 and learning asymmetry at step 2 (contingent feedback conditions only). This showed that, the effect of Bias on generalization was greatly reduced (from $t(48) = 10.15$, $p < .001$ at step 1 to $t(47) = 2.88$, $p < .01$ at step 2) when controlling for learning asymmetry, indicating a partial but highly significant mediation effect as assessed by a Sobel test ($z = 7.51$, $p < .001$).

Correspondence to Human Data

The closest human parallel to the situation simulated here is the Eiser et al. (2007) experiment, where participants were advised either that most beans were bad, or that most were good. This manipulation was set deliberately weak, so that many participants in fact disregarded this advice. Nonetheless, if the data are re-analyzed just for those participants who most closely followed the advice given to them (i.e., those with "adherence" scores at least 0.5 SDs above the mean, as
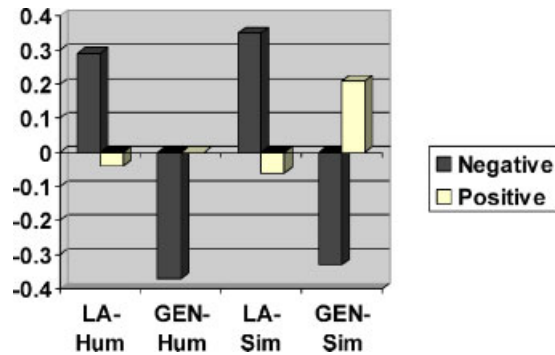
Figure 3.   Comparison of mean learning asymmetry (LA) and generalization (GEN) scores from high-adherence participants in the Eiser et al. (2007) experiment (-Hum) and the present simulations (-Sim) under mild bias at 200 epochs

defined in the original paper), the relative differences between conditions broadly resemble those observed in our simulations. Figure 3 compares the mean learning generalization asymmetry scores from these human data with those from the mild negative and positive conditions at 200 epochs.

DISCUSSION

Our purpose in these simulations was not to produce an *exact* replication of a specific human experiment, but to offer a "proof of concept" that expectancy confirmation can occur in a learning system incapable of conscious reasoning. Our findings are congruent with other evidence that apparent biases can emerge from the operation of "basic" processes such as reinforcement learning. For example, risk aversion (preferring high-probability moderate rewards over low-probability high rewards) has been demonstrated in simulations of organizational decision-making (March, 1996) and bees' foraging (Ziv, Joel, Meilijson, & Ruppin, 2002).

Expectancy confirmation was operationalized by adjusting the learning rate differentially for good or bad outcomes, depending on the condition. Thus, a positive bias meant that the network updated its representation of the input space (i.e., the beans) more emphatically (or "confidently") whenever it experienced good beans rather than bad outcomes, whereas the opposite was true under a negative bias. The familiar learning asymmetry effect under contingent feedback, already discernible at 200 epochs, was weaker or eliminated by positive bias but stronger under negative bias. Continuing training to 5000 epochs improved recognition of bad beans even further, but errors still persisted in the recognition of good beans under neutral and negative bias conditions.

Turning to the generalization data, by the end of training novel beans were rated on average more positively under positive than neutral bias. However, the neutral and negative bias conditions did not differ. This is readily explained by reference to the learning asymmetry effect, which persisted under neutral bias to the same extent as under negative bias, but was eliminated under positive bias. In other words, the network's more negative judgment of novel beans except under positive bias reflects its underestimation of the proportion of good beans among the 36 presented during the training phase, This does not apply under full feedback, since the whole point of full feedback (if allowed to continue throughout training) is that it eliminates the selectivity of experience of good and bad beans on which the learning asymmetry effect depends.

In interpreting these data, it is important to remember that connectionist networks are merely formal systems for identifying abstract relationships between symbols or patterns of data. There is nothing in the formal systems themselves that entails a correspondence between such abstract symbols and relationships and any concrete objects or phenomena. Within our simulations, we chose target values of 0 to represent "bad" and 1 to represent "good," but *by themselves* these numbers, and the direction of the difference between them, are completely arbitrary. It is only under contingent feedback that this distinction becomes meaningful. This is because, firstly, "good" is effectively defined operationally as "more likely to be approached" and, secondly, feedback depends on approach. Valence is thus defined within our paradigm in terms of its relationship to approach-avoidance, and in terms of the asymmetrical relationship between approach-avoidance and feedback. The theoretical contribution of simulations therefore depends less on *whether*

(human) experimental findings can be reproduced than on *how*. The single most important message is that supposedly complex phenomena may emerge from the iteration of extremely simple underlying processes.

Whereas connectionist methods are being used to simulate an increasingly wide range of social psychological phenomena (e.g., Read & Miller, 1998; Smith, 1996; Van Overwalle & Siebler, 2005), our own approach is distinctive in terms of its focus on the dependence of attitude formation on actions taken by the learner. Introducing a bias on the learning system illustrates one way in which we may protect our existing attitudes from the effects of contradictory experience. In more standard applications of the backprop algorithm, as noted earlier, no particular theoretical meaning is attached to the learning rate parameter used to control how quickly the network learns. A novel feature of our simulation was effectively to treat the learning rate as reflective of a ''readiness to learn'' that controlled the speed of learning, conditional on the kind of reward received, effectively accelerating learning following feedback that confirmed expectancies, and decelerating it following feedback that did not.[2]

Note that all nets started from similarly randomized settings of the connection weights, so that there is no implication that the net was ''reading across'' representations from another's mental states. As operationalized here, ''expectancies'' consisted merely in a differential readiness to learn from specific stimulus-outcome combinations. Likewise, expectancies do not simply over-ride direct experience. Although the positive bias slows down the *rate* at which false-positive evaluations will be corrected, so long as the net continues to approach bad beans, it will receive corrective feedback that will eventually turn such evaluations negative. Conversely, the negative bias slows down the rate of true-positive identification of good beans while speeding up the true-negative identification (and hence, subsequent avoidance) of encountered bad beans. At the same time, direct experience does not eliminate the effect of earlier advice. In other words, when the net receives feedback contrary to expectancies derived from ''advice,'' it (somewhat) discounts the feedback, but does not lose ''trust'' in the advice as such; on the next trial where feedback is received, the bias operates at full strength.[3]

This emphasis on how direct and indirect experience interacts may have more general implications for the role of social influence in attitude formation. Perhaps because of the experimental paradigms within which such processes are typically studied, advice from others or persuasive messages tend to be treated as ''one-off'' social impacts. To the extent that the source of the message is liked and/or trusted, the recipient is predicted (by several theorists from e.g., Osgood and Tannenbaum, 1955, onwards) to shift his or her position to become closer to that advocated by the source. But what happens afterwards, when individuals act upon the advice they have received or gain further information from other sources? Is the initial advice merely averaged in with later inputs (e.g., Anderson, 1981), or does it influence the *interpretation* of subsequent information and experience, as implied by the broader literature on expectancy-confirmation processes (e.g., Darley & Fazio, 1980; Darley & Gross, 1983)? The present study supports this latter possibility, through helping identify how the consistency of feedback with prior expectancies may shape how individual acquire attitudes through exploration of their environment.

ACKNOWLEDGEMENTS

REFERENCES

Abelson, R. P. (1959). Modes of resolution of belief dilemmas. *Journal of Conflict Resolution*, *3*, 343–352.
Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

[2]Strictly, we did not manipulate the learning rate itself as a function of bias. However, Δ is multiplied by the bias and learning rate parameters together, so the combined result amounts to a rescaling of the learning rate.
[3]In principle, our simulation could be modified to make the strength of reliance on advice itself dependent on feedback (for example, by making G a dynamic parameter that strengthened following expectancy-consistent feedback, and weakened following expectancy-inconsistent feedback). While conceptually plausible, this would not at all undermine the learning asymmetry effect under negative bias, since the whole point is that, by avoiding a larger proportion of objects, the net avoids much of the expectancy-inconsistent feedback that might be a signal to reduce reliance on such advice.

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*, 925–945.

Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist*, *35*, 867–881.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, *44*, 20–33.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Dennett, D. C. (1991). *Consciousness explained*. Boston, MA: Little, Brown.

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.

Eiser, J. R., Shook, N. J., & Fazio, R. H. (2007). Attitude learning through exploration: Advice and strategy appraisals. *European Journal of Social Psychology*, *37*, 1046–1056.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion, 15*, 115–141.

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

Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. New York: Free Press.

March, J. G. (1996). Learning to be risk averse. *Psychological Review*, *103*, 309–319.

Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review*, *62*, 42–55.

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

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1, pp. 318–362). Cambridge, MA: MIT Press.

Shook, N. J., Fazio, R. H., & Eiser, J. R. (2007). Attitude generalization: Similarity, valence and extremity. *Journal of Experimental Social Psychology, 43*, 641–647.

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

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*, *11*, 42–61.

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

Van Overwalle, F., & Heylighen, F. (2006). Talking nets: A multiagent connectionist approach to communication and trust between individuals. *Psychological Review*, *11*, 606–627.

Van Overwalle, F., & Siebler, F. (2005). A connectionist model of attitude formation and change. *Personality and Social Psychology Review*, *9*, 231–274.

Ziv, Y., Joel, D., Meilijson, I., & Ruppin, E. (2002). Evolution of reinforcement learning in uncertain environments: A simple explanation for complex foraging behaviors. *Adaptive Behavior*, *10*, 5–24.