Journal of Experimental Psychology:
Learning, Memory, and Cognition
2003, Vol. 29, No. 4, 626-640

0278-7393/03/$12.00 DOI: 10.1037/0278-7393.29.4.626

Learning Myopia: An Adaptive Recency Effect in Category Learning

Matt Jones
University of Michigan

Winston R. Sieck
Ohio State University

Recency effects (REs) have been well established in memory and probability learning paradigms but have received little attention in category learning research. Extant categorization models predict REs to be unaffected by learning, whereas a functional interpretation of REs, suggested by results in other domains, predicts that people are able to learn sequential dependencies and incorporate this information into their responses. These contrasting predictions were tested in 2 experiments involving a classification task in which outcome sequences were autocorrelated. Experiment 1 showed that reliance on recent outcomes adapts to the structure of the task, in contrast to models' predictions. Experiment 2 provided constraints on how sequential information is learned and suggested possible extensions to current models to account for this learning.

Recency effects (REs) are a robust phenomenon in cognitive psychology. REs are said to occur whenever more recent experiences are better remembered or are more influential in judgments about present situations. For example, in research on verbal working memory, REs are arguably among the most fundamental established phenomena, most commonly seen as increased performance on the final positions in free- or serial-recall tasks (e.g., Crowder, 1972; Murdock, 1962). Similar results have since been observed in visuo-spatial working memory (Broadbent & Broadbent, 1981), as well as in animals (Thompson & Herman, 1977; Wright, Santiago, Sands, Kendrick, & Cook, 1985). REs in working memory have often been attributed to spontaneous decay of stored information (Baddeley, 1986; Burgess & Hitch, 1999); however, this simple interpretation has been called into question by recent results showing that the rate of information loss can change, adaptively, in response to temporal statistics of the task (R. B. Anderson, Tweney, Rivardo, & Duncan, 1997). This flexibility is more consistent with a functional account of working memory (J. R. Anderson & Schooler, 1991, 2000; Schacter, 1999) and suggests that there is more underlying the phenomenon than simple architectural constraints.

Another area in which REs commonly arise is animal conditioning experiments. Common learning phenomena that depend on trial order, such as extinction, counterconditioning, and discrimination-reversal learning, all fall into the category of REs because they are characterized by behavior at the conclusion of learning being based primarily on the most recent (second) phase of training, rather than an average of both phases. However, the existence and magnitude of such trial order effects depend crucially on the relationship of physical and temporal contexts among the phases of training and testing, in a manner that is easily argued to be rational and adaptive (see Bouton, 1993, for a review). Similar REs are commonly observed in human experimentation with probability learning, for instance as a tendency to respond with whatever option was given as correct on the previous trial, independent of the reinforcement pattern over the prior history of the task (e.g., Jarvik, 1951; Nicks, 1959). This effect is easily predicted both by associative-learning models and by models that assume decay of memory for past events. However, such theories have trouble with further results indicating that REs in these tasks are subject to learning effects in response to sequential dependencies in the target sequence. In particular, when outcomes of successive trials are made to depend on each other, rather than being independently sampled, observed REs adapt accordingly (N. H. Anderson, 1960).

Some evidence for order effects has also been found in multiple-cue category learning (MCCL). For instance, Busemeyer and Myung (1988) found that humans' subjective prototypes for groups of random dot patterns showed a temporal bias involving both primacy and REs. Still, a detailed investigation of sequential effects in MCCL has yet to be performed. This situation is reflected in current models of the task, which either do not address REs or treat them as static by-products of other processes. Such a stance is at variance with the results from other domains, which suggest instead that REs may be adaptive and sensitive to statistical characteristics of the task. In particular, when viewed as an extension of probability learning, MCCL might be expected to produce effects similar to those mentioned above wherein REs adapt to sequential dependencies among outcomes. On the other hand, the presence of variable cues gives two reasons that the pattern might be different. First, there is evidence that REs in MCCL are more complex in that they depend on the similarity between present and previous stimuli (Sieck, 2000). Second, whereas uncued learners have little to attend to but the target sequence, subjects in MCCL tasks have more tangible stimuli about which they are explicitly instructed to learn. This attention shift, combined with the memory demands associated with using

Matt Jones, Department of Psychology, University of Michigan; Winston R. Sieck, Department of Psychology, Ohio State University.

This work was supported in part by National Science Foundation Grant SES-9911301.

Correspondence concerning this article should be addressed to Matt Jones, who is now at the Department of Psychology, University of Texas at Austin, 1 University Station A8000, Austin, Texas 78712. Email: mattj@psy.utexas.edu.

multiple cues, could conceivably eliminate or greatly attenuate any learning of sequential dependencies.

In the remainder of this article we describe a focused exploration of REs in category learning intended to address these issues. Our primary aims are to investigate (a) the presence and nature of REs in cued category learning, (b) the role played by cues and by stimulus similarity, and (c) potential learning effects in response to sequential dependencies in the target sequence, along with the mechanisms underlying such effects. We begin with a more detailed account of results from the probability learning literature, in which REs have been well documented and extensively investigated. Next we discuss some of the more popular current models of category learning and present a simulation study to evaluate their predictions concerning REs in situations with and without sequential dependency. We then present two experiments in a cued categorization task involving sequential dependencies (one- and two-step autocorrelations) among the correct category outcomes, designed to contrast the predictions made by these models with the predictions implied by a functional and adaptive interpretation of REs (as extrapolated from results in short-term memory and probability learning). Our main result was that, when outcomes were positively dependent on the outcomes of the previous trial (positive autocorrelation), observed REs increased, with an opposite effect for negative autocorrelation. Thus REs in MCCL are more flexible and adaptive than acknowledged by current theories.

REs in Probability Learning

In an (uncued) probability-learning paradigm, subjects are asked over repeated trials to predict which of a fixed set of events will occur. In the simplest version of the task the event probabilities are constant across trials, and subjects' ostensible task is merely to discern the relative likelihood of each response being correct. However, even in purely random situations subjects seem to try to make use of recent sequential information (see Myers, 1976, for a review). One commonly observed effect is a one-step positive RE, whereby subjects are more likely to give the response that was reinforced on the previous trial (Edwards, 1961; Jarvik, 1951; Nicks, 1959; Suppes & Atkinson, 1960, p. 196). It has been noted (Atkinson, Bower, & Crothers, 1965, p. 363; Estes, 1957) that conditioning-based theories of learning (e.g., stimulus sampling theory; Atkinson & Estes, 1963) naturally predict positive REs of this sort because cumulative iteration of their learning rules leads the most recent events to have the greatest effect on current response probabilities (Atkinson et al. 1965, p. 363; Estes, 1957).

The iterative conditioning explanation of REs predicts them to be constant, with magnitude dependent only on the learning rate. This prediction has been falsified by experiments using sequential dependencies, in which the target event for each trial depends (probabilistically) on the outcomes of the previous trial (N. H. Anderson, 1960; Engler, 1958; Hake & Hyman, 1953; Witte, 1964). For instance, N. H. Anderson (1960) varied the probability that the outcome of each trial in a two-choice task would match the outcome of the previous trial. In all cases, subjects' proportion of repetition responses (choosing the option that was previously correct) came to approximate the true repetition probability, with a small but consistent positive bias. Stated in terms of REs the one-step positive RE described above showed adaptation to the one-step autocorrelation actually present in the target sequence.

One approach to modeling of learning in situations like the one just described is to treat the task as one of discrimination learning with information from previous trials serving as cues. For instance, Burke and Estes (1957) applied a generalization of stimulus sampling theory to discrimination learning based on the immediately preceding outcome. Their model predicts conditional response probabilities to approximate conditional event probabilities, as has been found empirically (N. H. Anderson, 1960). Extensions of this idea to longer memory spans ("*k*-span models") can be found in Restle (1961, pp. 109-111) and Feldman and Hanna (1966). The major weakness of this approach is that it assumes information is encoded in terms of individual past outcomes, whereas there is good evidence that subjects attend more to patterns, such as runs of identical or alternating outcomes (Glanzer & Clark, 1962; Royer, 1967).

A contrasting theory, which we term the *aggregation hypothesis*, assumes that subjects retain sequential information from more than one trial back but merely encode some aggregate information combining the outcomes of recent trials. This assumption is embodied in run-based models, which encode outcome histories entirely in terms of homogeneous runs, that is, which outcome was seen most recently and how many consecutive times it just appeared (Gambino & Myers, 1967; Restle, 1961). Gambino and Myers's (1967) generalization model also incorporates the assumption of a confusion gradient in learning of run continuation probabilities: Reinforcement after a run of length n also affects the subject's assessment of the probability of continuation of runs of length m, to a degree that decays exponentially with $|n - m|$. Although there is clear evidence of subjects' ability to learn sequential structure beyond that contained in run length distributions (e.g., Feldman & Hanna, 1966; Rose & Vitz, 1966), run-based models have enjoyed the most success in accounting for the majority of empirical data, at least for simpler sequential manipulations (Myers, 1970; Rose & Vitz, 1966).

REs in Multiple-Cue Category Learning

As mentioned above, present models of repeated cued classification have considerably less to say when it comes to sequential effects, predicting either that such phenomena do not occur in this task or else that they occur only as by-products of other processes (e.g., iterated conditioning or decay of memory, as described below). Such predictions are plausible even in the face of contrasting results from probability learning because of the great difference in complexity between the two tasks. Whereas uncued learners have nothing to attend to but the past target sequence, subjects in cued categorization tasks have more reliable and tangible stimuli and thus may ignore sequential information.

One class of category learning models that do make specific a priori predictions regarding REs are those incorporating the δ-rule, an error-driven learning mechanism (Gluck & Bower, 1988; Kruschke, 1992). Simple REs arise naturally in such models for the same reason that they appear in conditioning-based theories of probability learning, namely that the most recent updating always has the strongest influence on the present strength of associations. Another major class of models of category learning, those based on explicit representations of past exemplars, have somewhat less to say about REs. Classic exemplar models (Medin & Schaffer, 1978; Nosofsky, 1986) predict no temporal effects whatsoever, as all past

instances participate equally in classification of new stimuli regardless of their time of presentation. Exemplar models can, however, be made to mimic δ-rule predictions by explicitly incorporating exponential trace decay (Estes, 1994; Estes & Maddox, 1995; Nosofsky, Kruschke, & McKinley, 1992; Nosofsky & Palmeri, 1997). These models and their predictions are detailed in the following section. In all cases, it will be shown that the REs are predicted to be static and unresponsive to characteristics of the task environment.

Simulation 1: Predictions of Categorization Models in Autocorrelated Ecologies

A simulation study was conducted to elucidate the RE predictions of three prominent models of category learning, with an emphasis on their predictions regarding statistical dependencies in the target sequence. The categorization task simulated was the same as the one used in Experiment 1 below. The cover story used in that experiment was simulated medical diagnosis: Subjects were presented with a series of hypothetical patients and asked to diagnose them as having one of two fictitious diseases *(trebitis* and *philiosis),* based on the presence or absence of three possible symptoms. Feedback was given after each trial as to what disease the patient actually had. The task involves two categories (the diseases) and three binary cues (the symptoms); both pairs of terms are used interchangeably in subsequent discussion. The design contrasts three conditions varying in terms of the one-step autocorrelation present in the target sequence: positive, negative, and control (independent sampling). Three models were simulated in this task: the adaptive network (AN) of Gluck and Bower (1988), the exemplar model with trace decay (ETD), and the attentional learning covering map (ALCOVE), a hybrid model that incorporates δ-rule learning in an exemplar-based representation (Kruschke, 1992). As all three models have been presented in detail elsewhere, we describe here only their main features, referring the reader to the original publications for further details.

Models

Adaptive network (AN). The AN model consists of a two-layer network, with input nodes representing cues and a single output node determining responses.[1] Activation u of the output node is calculated as a sum of the input node activities, each multiplied by its corresponding weight. Response probabilities are then a nonlinear (sigmoid)[2] function of u. Learning after each trial takes place based on the discrepancy between u and the feedback signal encoding the correct response. This discrepancy of prediction is termed the *d-signal.* The adjustment of each weight is proportional to this error and to the product of the weight and the corresponding input activation. Further details can be found in Gluck and Bower (1988) and the Appendix.

One relevant detail of the AN concerns the representative scheme of the input layer, for which there are at least three options. The first, which we refer to here as the *single-node network,* has one input node per cue dimension (three total for the present application), with each node being activated only when the corresponding cue is present. This network essentially assumes that only present cues are actively represented, and only they provide a basis for learning. A second alternative is the *double-node network,* which has for each cue dimension one node that encodes cue present and another encoding cue absent. Each node is active only when its condition is met. The third alternative, the *dual-node network,* also assumes that both present and absent cues are actively encoded but uses a single node per cue dimension to represent these possibilities. The node is activated positively when the cue is present and negatively when it is absent. Gluck and Bower (1988, Experiment 3) were able to reject the dual-node model due to its prediction of perfectly complementary response probabilities for complementary cue profiles (a consequence of the zero-sum nature of its input representation). However, this prediction can be relaxed by assuming that the context of the task acts as an additional, constant stimulus. Therefore we consider a version of the dual-node model that includes an additional context node, which is always active.

All three versions of the AN make qualitatively the same predictions regarding REs, with one important distinction: Because of its negative input activations, the dual-node network can predict a negative RE when the present and previous cue profiles are sufficiently complementary. The other two networks are incapable of predicting negative REs of any sort. Because of results reported later in this article, and because other than this particular prediction the models' performance is qualitatively identical, only the results of the dual-node network are presented for this simulation.

Exemplar with trace decay (ETD). The ETD model is an extension of the context model of Medin and Schaffer (1978; see also Nosofsky, 1986) that incorporates decay of memory for past cases. Specifically, the influence of past exemplars on current classifications decays exponentially with time since presentation, operationalized in terms of the number of intervening trials. This decay of trace strength can be interpreted either as an attenuated contribution to the choice process, or as a declining probability of being retrieved from memory (with all retrieved exemplars then participating equally in the choice process). Knowledge representation in the ETD model is given by direct storage of all past cases, including their values on all cue dimensions and their correct category membership. On presentation of a new instance, the odds of responding with Category A are given as the ratio of the number of stored exemplars from Category A to the number of stored exemplars from Category B, with each case being weighted by its trace strength and its similarity to the present stimulus. Similarity is defined as a product over cue dimensions, implying that the similarity between two cases will be an exponential function of the number of mismatching cue dimensions. Further details can be found in Nosofsky et al. (1992) and in the Appendix.

ALCOVE. ALCOVE is a hybrid model that uses an exemplar-based representation within a neural network updated using generalized δ-rule learning. The design for ALCOVE was taken

[1] For the remainder of this article we restrict attention to situations involving two categories. For this situation we assume a network with one output node, representing the probability of Response A versus Response B, which is formally equivalent to a network with a separate output node for each category (Gluck & Bower, 1988, p. 234, Footnote 2).

[2] For some applications, such as modeling of subjects' estimated probabilities of specific outcomes, this transformation is taken to be linear (Gluck & Bower, 1988). We have simulated the model under this alternative assumption and found it to give the same qualitative predictions.

directly from Kruschke (1992; see also Kruschke, 1990; Nosofsky et al., 1992); similar ideas can be found in Estes (1988, 1994) and in Estes, Campbell, Hatsopoulos, and Hurwitz (1989). The model architecture is a network consisting of a set of input nodes coding individual stimulus features, a layer of hidden units corresponding (in the simplest implementation) to each possible exemplar, and a pair of output nodes for the two response categories. Each hidden unit is connected to the input units so that its activation is equal to the similarity between the stimulus it represents and the one currently presented. Similarity is defined as in the ETD model except that each cue dimension contributes a variable amount determined by a learnable attention weight. Activation of each output node is determined by summing the activations of the hidden units with each multiplied by the weight from that hidden node to the output node. Response probabilities are then given by a "softmax" function (cf. Luce, 1963) of the output node activations. Learning of hidden-to-output weights is accomplished in the same manner as in the AN, using separate δ-signals for each output node. Attention weights are learned similarly by means of backpropagation (see Rumelhart, Hinton, & Williams, 1986). Further details can be found in Kruschke (1992) and in the Appendix.

Method

The task was simulated in three separate conditions, differing in the autocorrelational pattern of their target sequences. The target sequences of each condition were generated using a fixed probability that the correct outcome on each trial would match that of the previous trial: positive condition: 70%; negative condition: 30%; control condition: 50%. Once each target had been determined, the cue profile for that trial was randomly selected according to the conditional probabilities shown in Table 1. Each model was simulated on each condition for 20,000 trials.

To expand on the basic design, our method for introducing autocorrelation was to generate the disease sequence as a Markov chain. Because events in a Markov chain are independent of earlier history once conditioned on the previous time step, this approach allowed us to isolate the relevant temporal information on any given trial to the immediately preceding outcome. Once a patient's disease had been selected, symptoms were generated depending on his or her disease according to fixed conditional distributions, independent of the symptoms or diseases of past patients (see Figure 1). Thus, under the Bayesian framework, a patient's symptoms could be thought of as current evidence for inferring the patient's disease, while the outcome of the previous trial provided the prior odds ratio, according to the Markov transition probabilities. According to

Table 1
Symptom Probabilities (in Percentages) Given Diseases in Simulation 1 and Experiment 1

	Disease	
Symptom pattern	Trebitis	Philiosis
S_1	31.1	2.2
S_1, S_3	24.4	8.9
S_1, S_2	17.8	15.6
S_3	15.6	17.8
S_2	8.9	24.4
S_3, S_2	2.2	31.1

Note. S_1 = Symptom 1 (runny nose); S_2 = Symptom 2 (swollen hands); S_3 = Symptom 3 (sore throat).

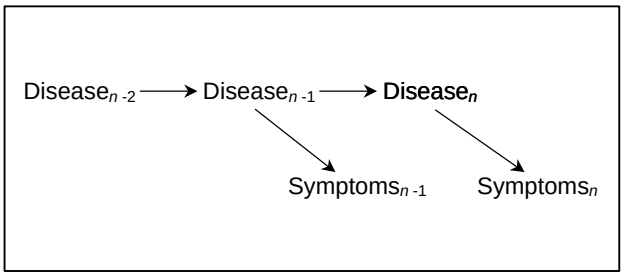

Figure 1. Method for generation of ecologies, Simulation 1 and Experiment 1. On a given trial, the disease outcome is determined solely by the previous disease outcome; symptom profiles are determined solely by present diseases.

Bayes's theorem, the correct posterior odds ratio for the disease on the n^{th} trial, once the symptoms have been revealed, is given by

$$\log\left(\frac{\mathrm{P(T)}}{1-\mathrm{P(T)}}\right) = F(S) + G(D_{n-1}). \quad (1)$$

Here T is the target trebitis, F is the likelihood-ratio function taking different values for each possible stimulus configuration S; D_{n-1} is the disease on the n-1st trial, and G is prior log-odds of trebitis, as determined by the Markov transition matrix. In the positive condition, G(T) = log[.70/(1 – .70)] = .85 and G(P) = log[.30/(1 – .30)] = –.85 (where P indicates philiosis), whereas in the negative condition these values are reversed. In the control condition G = 0.

The properties of this normative model are illustrated in Figure 2, which shows the predictive value of D_{n-1} as a function of the number of matching symptom values between the n^{th} and n-1st patients. Two features of Figure 2 are especially noteworthy. First, for the positive and negative conditions, recent information does have a unique and informative predictive role, as contrasted with the standard independent-sampling paradigm embodied in the control condition. Therefore, optimal behavior in the positive and negative conditions necessarily involves positive and negative REs (respectively), as dictated by the rightmost term in Equation 2. However, only the disease of the immediately preceding patient is relevant; information from earlier patients is redundant. Second, the predictive effects of present cues and past outcomes are purely additive and nonin-

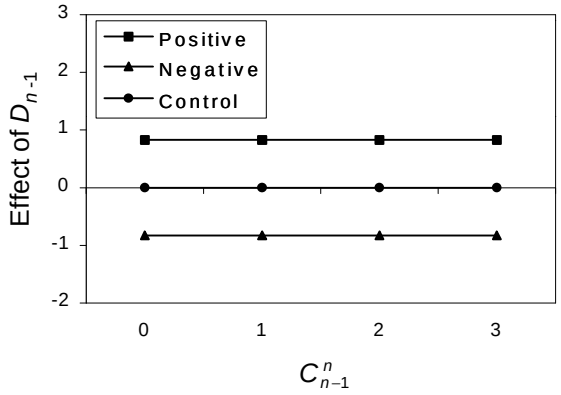

Figure 2. Predictive validity of previous outcome in ecologies of Simulation 1 and Experiment 1. The horizontal axis measures the number of matching cue values between the present and previous patients. The vertical axis gives the coefficient for the previous patient's disease in predicting the current outcome (see Equation 1). Note that these values also represent the magnitudes of one-step recency effects for normative responding.

teracting and, more specifically, the information contained in the previous patient's disease is independent of the similarity of that patient's symptoms to those of the present one. As we show next, this is in contrast to the behavior predicted by the models.

Results

Analyses of the models' performance proceeded by linear regression of response probabilities. The predictors of primary interest were the previous disease $D_{n\text{-}1}$ and its interaction with the cue commonality C^{n}_{n-1} between present and previous patients, measured as the number of symptoms that both had or did not have. In order to obtain unbiased estimates of these predictors, we also included in the regression all other predictors that were correlated with the ones of interest and that could potentially affect responses. For example, because the models all predict two-step REs, and because $D_{n\text{-}2}$ is correlated with $D_{n\text{-}1}$ in two of the three conditions, $D_{n\text{-}2}$ must be included in order to give a fair estimate of the effect of $D_{n\text{-}1}$.[3] The full set of predictors was therefore S^{1}_{n}, S^{2}_{n}, S^{3}_{n} (the three present symptom values), S^{1}_{n-1}, S^{2}_{n-1}, S^{3}_{n-1} (the symptoms of the previous patient), $D_{n\text{-}1}$, $D_{n\text{-}2}$, and the interactions $D_{n\text{-}1} \times C^{n}_{n-1}$ and $D_{n\text{-}2} \times C^{n}_{n-1}$. All variables except C^{n}_{n-1} were coded as ±1/2 so that the regression coefficient for each predictor would correspond directly to the effect of that variable on average responses. $\bar{C}^{n}_{n-1}$ was treated as a categorical (as opposed to numeric) variable for the $D_{n\text{-}1} \times \bar{C}^{n}_{n-1}$ term to more fully determine the nature of the interaction, and was coded as (0, 1, 2, 3) for the $D_{n\text{-}2} \times C^{n}_{n-1}$ term (the assumption of linearity for the latter interaction allowed for more power, especially in the empirical investigations to follow, while not affecting the qualitative pattern of the results). Although all of these terms needed to be included, as discussed above, our primary interest was in the influence of the previous disease outcome, as well as its dependence on the cue commonality. Figure 3 illustrates these primary features of interest (comprehensive tables of results from these and all subsequent regression analyses can be obtained by request from Matt Jones).

As can be seen in Figure 3, the models all predict a positive one-step RE that interacts positively with cue commonality. Closer inspection of this interaction reveals clear qualitative differences in the predictions of the three models: The AN predicts a linear interaction, with the effect of $D_{n\text{-}1}$ becoming negative for the lowest value of C^{n}_{n-1}, whereas ETD and ALCOVE predict a one-step RE that is always positive and is positively accelerated as a function of C^{n}_{n-1}. Most crucially, the level of one-step RE exhibited by each of the models is independent of the autocorrelation present in the target sequence, as shown by the close overlap of the lines in Figure 3A-C. This confirms the earlier claim that current categorization models predict REs to be invariant with respect to autocorrelation in the target sequence.

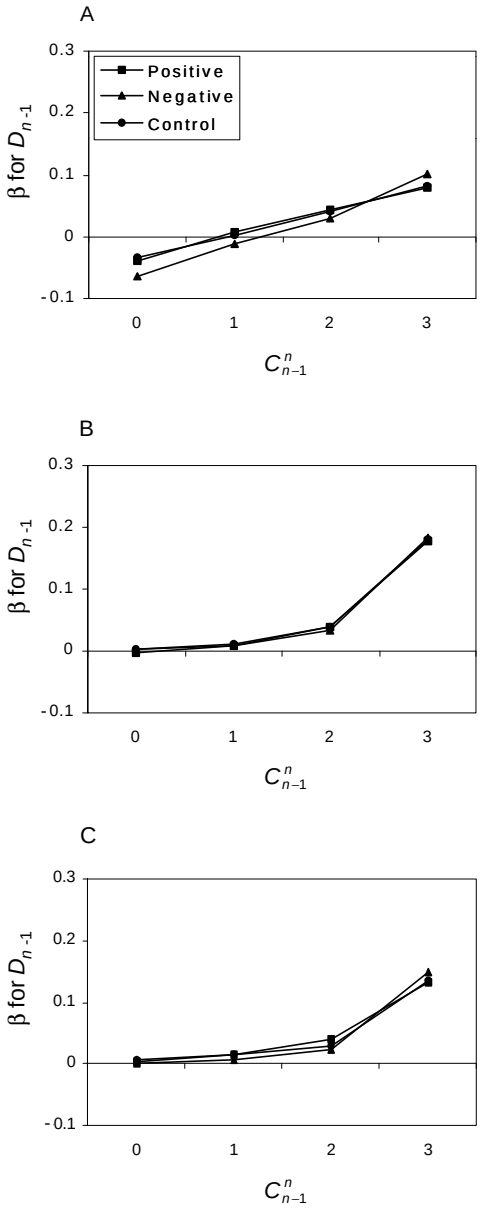

Figure 3. Influence of previous outcome on response probabilities as a function of number of common cues between present and previous trials (C^{n}_{n-1}), for adaptive network model (A), exemplar model with trace decay (B), and the attentional learning covering map model (C).

Discussion

The results of Simulation 1 are summarized as follows: All three models predict positive one-step REs that increase with cue commonality. The prediction of an overall positive RE is a result of δ-rule learning in AN and in ALCOVE, and comes directly from the trace decay assumption in the ETD model. The interaction of RE with cue commonality in all three models differs from the normative model, and occurs because similar past cases weigh more heavily on the present decision. This is an explicit assumption in the ETD model, whereas in AN and ALCOVE the effect is due to specificity of the learning mechanisms: Weights are updated in proportion to their relevance on each trial, implying that the learning on Trial $n - 1$ will influence the response on Trial n to the degree that the two stimuli activate the same input nodes. In the

[3] The same argument would seem to apply for the inclusion of $D_{n\text{-}3}$; however, the correlation between $D_{n\text{-}3}$ and $D_{n\text{-}1}$ is eliminated once $D_{n\text{-}2}$ is controlled for, due to the Markov nature of the disease sequences.

AN this same mechanism produces a negative RE when cue commonality is minimal, because negative input activations allow negative transfer of learning between stimuli with opposite cue values. Note that among the three models, the AN, and specifically the dual-node version, is unique in predicting this negative RE. The AN also differs from the other two models in that it predicts RE to be a linear function of cue commonality; by contrast, ETD and ALCOVE both predict RE to be markedly greater for a past case that highly matches the present one, as a consequence of these models' multiplicative similarity-activation functions.

Most important, the REs predicted by all three models were unchanged by the autocorrelation manipulation. This is a robust prediction that holds under broad changes in all parameter values (see the Appendix). The invariance of RE with respect to autocorrelation results from the fact that only the learning rate (in AN and in ALCOVE) or the rate of forgetting (in ETD) determine the preferential reliance on recent outcomes. In particular, the models include no mechanism for recognizing sequential dependencies. We turn now to an empirical test of this assumption.

Experiment 1: Recency and the Effects of Autocorrelation

The purpose of Experiment 1 was to verify the presence of REs in category learning, and to test their dependence on cue commonality and autocorrelation among outcomes. In order to determine whether REs in category learning are due to immutable or adaptable mechanisms, we divided subjects into three groups differing in the manner of one-step autocorrelation present in the sequence of diseases they encountered: positive, negative, or none (control). If REs were immutable, as predicted by the models simulated above, then β weights for the previous outcomes in predicting present responses would be equivalent across experimental conditions. If, however, REs were adaptive, then βs would be expected to shift in the direction supported by each ecology.

Method

Subjects. Subjects were 100 undergraduates from the University of Michigan who received partial credit in an introductory psychology course for participation. The experiment was conducted in groups of 10 to 20, with subjects randomly divided among three conditions: positive (n = 34), negative (n = 33), and control (n = 33).

Design. There were three between-subjects conditions with positive, negative, and null (control) autocorrelational patterns, exactly as described in Simulation 1.

Procedure. The stimulus associated with each successive hypothetical patient was presented on a Dell PC (Austin, TX) as a pair of vertical lists of symptoms, arranged side by side, titled *The patient presents with* and *The patient does NOT exhibit.* Each of three symptoms—runny nose (S1), swollen hands (S_2), and sore throat (S_3)—was contained in one of the two lists. Below the symptom display were buttons labeled with two fictitious diseases—trebitis and philiosis—from which subjects were instructed to choose their diagnosis, along with a window in which subjects typed their percentage of confidence in their diagnosis (subjects entered their responses using a keyboard and a mouse). The confidence response was constrained to be between 50% and 100%. Detailed instructions on the meaning and use of the probability scale were provided, in accordance with earlier probability judgment research (Sieck & Yates, 2001; Yates, Lee, Shinotsuka, Patalano, & Sieck, 1998). After both responses were made, the correct disease was displayed, with the symptoms and the subject's responses still visible. The subject then clicked a button to clear the screen and begin the next patient.

Subjects were each presented with 150 hypothetical patients in two blocks of 75, with a message displayed on the screen after the first block telling them to rest, stretch, et cetera, and to hit *RETURN* when they were ready to continue. The entire experimental session was completed within 50 min.

Results

To measure contributions of present and past information to subjects' responses, we analyzed the data from each subject using regression. Subjective probability estimates for the outcome trebitis were obtained from subjects' confidence judgments in their diagnoses (by subtracting the given response from 100% on all trials where the diagnosis was philiosis) and fit with the same linear regression model used in Simulation 1. Diagnoses were also fit to a logistic regression model using the same predictors. Note that this logistic model is a direct generalization of the normative model (Equation 2) with additional terms included. Because the central results of the two analyses were qualitatively identical, only the probability data will be presented here.

Coefficients for the previous outcome $D_{n\text{-}1}$ at all four values of cue commonality C^n_{n-1} are shown graphically in Figure 4. These coefficients represent the magnitude of subjects' one-step RE conditioned on each possible degree of similarity between the present and previous patients. For example, in the control condition on trials where the present and previous patients matched on all symptom dimensions (C^n_{n-1} = 3), reported probabilities of trebitis were 15% higher when the previous patient had trebitis than when that patient had the alternative disease philiosis. The pattern of effects in Figure 4 reveals a heightened one-step RE for the positive condition as compared with control, with an approximately symmetric and opposite effect in the negative condition. All three conditions show an increased RE for larger values of C^n_{n-1}.

Formal tests of these observations were carried out by means of a repeated measures analysis of variance (ANOVA), with four observations per subject corresponding to the β weights for $D_{n\text{-}1}$ at all four values of C^n_{n-1}.[4] Condition was included as a between-subjects variable. The test of the main effect of C^n_{n-1} was significant, $F(2.2, 213.7) = 51.9$, $p \approx 0$, using Geisser-Greenhouse (G-G) correction for violation of sphericity with ε = .734, indicating that the influence of $D_{n\text{-}1}$ does depend on C^n_{n-1}. Furthermore the quadratic contrast for C^n_{n-1} was significant, $F(1, 97) = 16.4$, $p < .001$, implying that the dependence of RE on cue commonality is nonlinear. The main effect of condition was also significant, $F(2, 97) = 3.479$, $p < .05$, confirming that the level of RE differed across conditions. Finally the interaction between condition and C^n_{n-1} was not statistically significant, F(4.4,

[4] The two-stage approach used here, in which β weights are obtained separately for each subject and then compared across subjects, is nearly equivalent to a mixed-effects model that analyzes all of the data concurrently, using subject as a random effect (interacting with all 10 fixed effects) nested within condition. The present analysis gives qualitatively identical results with approximately the same statistical power, but avoids the difficulty of dependence among trials for a given subject (because, e.g., $D_{n\text{-}1}$ for Trial k is the same as $D_{n\text{-}2}$ for Trial k – 1). The only independence assumption required under our approach is that the β weights for different subjects be independent.

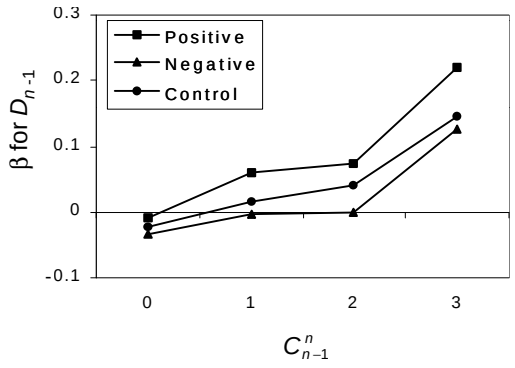

Figure 4. Influence of previous outcome on responses in Experiment 1 as a function of C^n_{n-1} (probability judgment data).

213.7) = 0.757, using G-G correction. This corresponds to the approximate parallelism among the lines in Figure 4 and is consistent with the interpretation that while the overall level of RE differed among conditions, the dependence of RE on cue commonality was unchanged.

Discussion

The results of Experiment 1 confirm the presence of REs in category learning. In addition, they demonstrate the positive dependence of REs on cue commonality, confirming this particular prediction of the categorization models discussed above. The form of this interaction is nonlinear and positively accelerated, consistent with the exemplar-based models (ETD and ALCOVE) but inconsistent with the version of the AN model considered. In contrast to the predictions of all three categorization models, the data clearly show that people can adapt their use of temporal information to the statistical structure of the task, in that the level of one-step RE changes in the direction of the one-step autocorrelation present in the target sequence. There is no need to attempt quantitative fitting of the models in order to reject them in their current forms, as their architectural assumptions (specifically the lack of sequential-learning mechanisms) imply a stronger qualitative prediction, independent of parameter values, that is contradicted by the present results.

Experiment 2: Manipulation of Autocorrelation at Multiple Lags

The primary purpose of Experiment 2 was to determine better the specific manner in which sequential dependencies affect categorization behavior through simultaneous manipulation of target autocorrelation at multiple lags. Since Simulation 1 and Experiment 1 showed that extant categorization models yielded inadequate theoretical guidance regarding sequential dependencies, Experiment 2 tested the previously described competing explanations offered by earlier models from the probability learning literature.

It has been shown in probability learning (Jarvik, 1951) that subjects can exhibit a positive RE over single trials concurrent with a negative RE (gambler's fallacy) over longer spans. Similarly, Experiment 2 included a new condition in which the current disease depended positively on the immediately previous outcome and negatively on that of the patient two trials back. This dual condition allowed testing of interference effects between temporal information at different lags. The disease sequences for this condition were generated as a two-step Markov chain, that is with each outcome dependent on the two before it. The transition probabilities (shown in Table 2 along with the other three conditions) were specified so that the effects of the previous two patients on the log odds for the present disease were additive, with D_{n-1} having a positive effect and D_{n-2} having a negative effect of equal magnitude. The information carried by the n-1st patient was identical in the dual and positive conditions. The major questions regarding this new condition were whether the additional two-back information would influence diagnoses (as compared with control), and whether its presence would affect subjects' learning of the one-back dependency (by comparison to the positive condition).

The hypotheses concerning learning of sequential dependencies that arose in the probability learning literature make differing predictions for these comparisons. Recall that the first hypothesis was that the previous k trial outcomes would be incorporated as cues (k-span model). An important special case is the single-span model where only the previous outcome is incorporated. The second hypothesis, which we dubbed the aggregation hypothesis, is that a summary of the previous trials is incorporated, for example, the last trial outcome along with the number of consecutive occurrences. Recall further that the aggregation hypothesis enjoyed the greatest empirical support in simple probability learning. But note that generalization to MCCL cannot be assumed presumptively for the same reasons complicated sequential processing could not be assumed, that is, there are substantially greater demands on attention and memory due to the need to process the nominal cues.

The k-span model (for $k > 1$) predicts that subjects will be able to fully learn the dependencies in the dual condition; thus they should exhibit a one-step RE equal to that seen in the positive condition, and a two-step RE lower than subjects in the control condition. The single-span model predicts no learning of two-back information (two-step REs in all conditions should be equal) and a one-step RE in line with the one-step conditional probabilities in the target sequence.[5] Finally, the aggregation hypothesis predicts interference between the two types of autocorrelation in the dual condition, because they make opposite contributions to the correlation between aggregated recent history and current outcomes. The presence of the negative two-step autocorrelation in the dual condition will therefore lead to a one-step RE that is less than that in the positive condition. This model differs from the single-span model in that it also predicts differences among the conditions in the level of two-step RE. The prediction for the comparison between dual and control conditions on this measure is undetermined, because the interference from the one-step autocorrelation

[5] The one-back model predicts a one-step RE in the dual condition that is less than that in the positive condition although still greater than in control. This is because of the positive correlation between D_{n-1} and D_{n-2} in the dual condition (due to the positive one-step autocorrelation), which makes the contribution of D_{n-1} to the odds ratio for D_n less when D_{n-2} is not controlled for than when it is. Therefore, even though the predictive contribution of D_{n-1} is equal in the positive and dual conditions when D_{n-2} is controlled for, this contribution is weaker in dual when D_{n-2} is ignored ($P[D_n = T \mid D_{n-1} = T] = 3/4$ in positive condition, = 9/14 in dual).

Table 2
Conditional Probabilities of Target Event T in Markov Model for Diseases in Experiment 2

	(D_{n-1}, D_{n-2})			
Condition	(T, T)	(T, P)	(P, T)	(P, P)
Positive	.75	.75	.25	.25
Negative	.25	.25	.75	.75
Dual	.50	.90	.10	.50
Control	.50	.50	.50	.50

Note. D_{n-1} = disease of patient in Trial $n-1$; D_{n-2} = disease of patient in Trial $n-2$; T = trebitis; P = philiosis.

could either over- or underbalance learning of the two-step contingency. However, because the dual and positive conditions differ only in terms of the two-step autocorrelation present in the disease sequence, the aggregation hypothesis must predict that subjects in the former condition will exhibit a lower value of two-step RE.

Method

Subjects. Subjects were 237 undergraduates from the University of Michigan who received partial credit in an introductory psychology course for participation. The experiment was run with groups of 15 to 25, with subjects randomly divided among four conditions: positive (n = 60), negative (n = 59), dual (n = 60), and control (n = 58).

Design. There were four between-subjects conditions with distinct autocorrelational patterns in their respective disease sequences. Diseases were generated in sequence according to the conditional probabilities displayed in Table 2. Outcome probabilities were set at 50% for the first trial for all participants. For subjects in the dual condition, the second disease was selected so as to have a 75% probability of matching the first. Once the disease sequence for a subject had been determined, each symptom value for each patient was randomly selected, dependent only on that patient's disease, according to the conditional probabilities shown in Table 3.

Procedure. The procedure was the same as in Experiment 1, except that subjects were presented with 300 hypothetical patients in three blocks of 100, with short breaks between blocks. The entire experimental session was completed within 70 min.

Results

Our general approach was once again to extract β weights from the models for each subject, and then to compare these betas across conditions to test for effects of the sequential manipulations. Here, the data from all but the first two trials for each subject were subjected to an extended version of the regression analysis described earlier. The results of primary interest were the influence of each of the prior two outcomes as a function of their respective cue commonalities, that is D_{n-1} for each value of C^n_{n-1} and D_{n-2} for each value of C^n_{n-2}.

Subjects' reported confidence in their diagnoses was again converted into probability judgments for the disease trebitis and fit with linear regression. Diagnoses (choice data) were fit using the corresponding logistic model. The logistic model was unidentified for 5 subjects (2 each in negative and dual, 1 in control) who were therefore excluded from analyses on the choice data. The coefficients for previous disease D_{n-1} are graphed as a function of cue commonality C^n_{n-1} in Figures 5A and 5B. As in Experiment 1, all four conditions showed an interaction between one-step RE and cue commonality, with the effect of D_{n-1} being greater for larger values of C^n_{n-1}; furthermore, this relationship was nonlinear. In addition, the overall level of one-step RE was greater in the positive condition and lesser in the negative condition, both as compared with control. The dual condition was nearly indistinguishable from control. These observations were confirmed through repeated measures ANOVAs on each data set, using the coefficients for D_{n-1} at the four levels of C^n_{n-1} as observations for each subject (see Table 4). Post hoc comparisons by means of Tukey's test, using a global alpha level of .05, revealed significant differences between all pairs of conditions except dual versus control.

Figures 5C and 5D show the corresponding data for Trial $n-2$. Again, the lines for all four conditions slope upward, reflecting an interaction between D_{n-2} and C^n_{n-2}, or two-step RE and two-step cue commonality. The separation of the lines suggests greater overall two-step RE in the positive condition, and possibly less in the dual condition, as compared with the other two conditions. Repeated measures ANOVAs on the four coefficients for D_{n-2} for each subject confirmed the main effects of condition and cue commonality, with no significant interaction (see Table 4). Pairwise comparisons among conditions by means of Tukey's test, with a global alpha of .05, showed differences between positive and dual (both data sets) and between positive and control (choice data only).

To gain a more direct measure of the effect of cue commonality on recency, we performed an additional analysis using only the responses from the second trial for each subject. This restriction allowed estimates of the effects of previous outcomes (i.e., the disease from Trial 1) that were not confounded by prior learning of sequential dependencies. Table 5 shows the proportion of subjects (collapsed across conditions) who chose on Trial 2 the disease that was correct on Trial 1, as a function of the cue commonality C^2_1 between the first two patients. This proportion is denoted R_2. Also shown is the mean probability assigned on the second trial to the correct disease of the first trial, $P_2(D_1)$, obtained from confidence judgments on Trial 2 by subtracting from 100% when appropriate. The positive relationship between C^2_1 and both R_2 and $P_2(D_1)$ is another manifestation of the RE by cue commonality interaction. A further observation is that responses appear biased away from D_1,

Table 3
Conditional Probabilities (in Percentages) of Symptom Profiles Given Disease in Experiment 2

	Disease	
Profile	Trebitis	Philiosis
S_1	9.1	19.0
S_1, S_3	19.0	9.1
S_1, S_2	4.4	39.4
S_3	39.4	4.4
S_2	9.1	19.0
S_3, S_2	19.0	9.1

Note. S_1 = Symptom 1 (runny nose); S_2 = Symptom 2 (swollen hands); S_3 = Symptom 3 (sore throat).

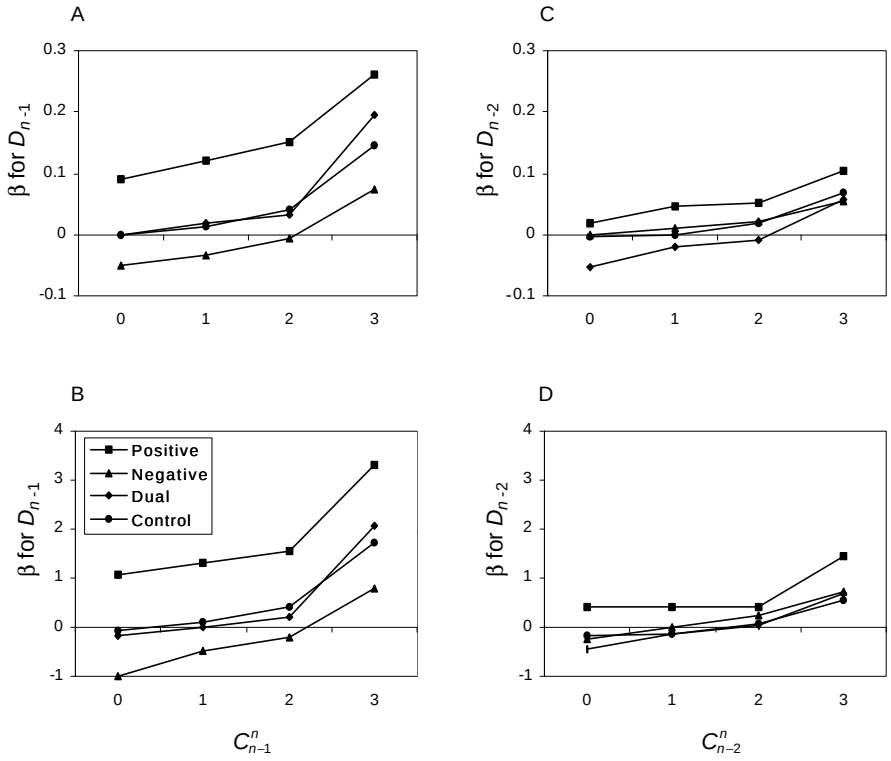

Figure 5. Influence of past outcomes on responses in Experiment 2. A: Average β weights for D_{n-1} as a function of C^n_{n-1} in analysis of probability judgments. B: Average β weights for D_{n-1} as a function of C^n_{n-1} in analysis of diagnoses. C: Average β weights for D_{n-2} as a function of C^n_{n-2} in analysis of probability judgments. D: Average β weights for D_{n-2} as a function of C^n_{n-2} in analysis of diagnoses.

that is both R_2 and $P_2(D_1)$ are less than 50%, for low values of C_1^2. Exact binomial tests on R_2 showed this bias to be significant for $C_1^2 = 0$ ($p < .0001$) and $C_1^2 = 1$ ($p < .001$). A one-sample t test comparing $P_2(D_1)$ to 50% for $C_1^2 = 0$ was also significant, $t(46) = -3.13$, $p < .01$. Thus, for sufficiently dissimilar stimuli the default RE, prior to any sequential learning effects, was negative. Neither exemplar model predicts this effect.

Discussion

The results of Experiment 2 replicated those of Experiment 1, providing further support that REs in cued category learning are (nonlinearly) dependent on cue commonality and can adapt to autocorrelations among outcomes. The primary manipulation of Experiment 2, the dual condition, showed no evidence of learning of one-back information even though this information was the same as in the positive condition where learning effects were large. The results were similar for two-back information: The dual condition was not significantly different from control, whereas the positive condition (in which the second previous outcome carried no unique information) showed heightened RE at this lag. These two results—mutual cancellation of the sequential dependencies in the dual condition and generalization of one-step autocorrelation to a two-step RE in the positive condition—contradict both the single-span and *k*-span models of sequence learning, but provide good support for the proposal that the aggregation hypothesis generalizes to MCCL.

The additional analyses of Trial 2 responses shed more light on the RE by cue commonality interaction, showing that RE is by default negative for sufficiently dissimilar cases ($C_1^2 = 0$ or 1). The restriction to Trial 2 allowed us to avoid sequence-learning effects and to obtain a direct estimate of the generalization from one instance to another. Only the dual-node version of the AN, with its negative input activations, can explain the observed negative generalization effect. Other versions of the AN, along with ETD and ALCOVE, have no way of producing negative REs of any sort.

Simulation 2: Sequential Learning Mechanisms

A second simulation was performed in order to explicitly test whether inclusion of mechanisms for sequential processing can explain our central results of adaptation of REs to autocorrelation. Because of the support for the aggregation hypothesis from the results of Experiment 2, we implemented that hypothesis here within the architecture of ALCOVE. Specifically, ALCOVE was augmented so as to have access to aggregated information about recent outcomes, for use as an additional cue in learning. For simplicity, this cue was taken to be the sum of the previous two outcomes, that is, the number of times trebitis appeared on the last

Table 4
Results of Repeated Measures ANOVAs on Effects of Previous Outcomes in Experiment 2

Effect	*F*	*df*	*p*
One-back effects			
Probability data			
Condition	23.7	3, 233	$<10^{-12}$
C^n_{n-1}	120.4	1.69, 394.5	$<10^{-39}$
Quadratic	92.1	1, 233	$<10^{-18}$
Condition × C^n_{n-1}	1.6	5.08, 394.5	.11
Choice data			
Condition	34.4	3, 228	$<10^{-17}$
C^n_{n-1}	118.7	1.88, 429.4	$<10^{-43}$
Quadratic	84.1	1, 228	$<10^{-16}$
Condition × C^n_{n-1}	1.3	5.65, 429.4	.24
Two-back effects			
Probability data			
Condition	5.8	3, 233	<.001
C^n_{n-2}	63.1	2.36, 548.6	$<10^{-29}$
Quadratic	18.5	1, 233	<.0001
Condition × C^n_{n-2}	1.4	7.06, 548.6	.19
Choice data			
Condition	3.6	3, 228	<.05
C^n_{n-2}	31.4	1.99, 452.8	$<10^{-12}$
Quadratic	11.1	1, 228	<.001
Condition × C^n_{n-2}	0.8	5.96, 452.8	.59

Note. Dependent variable is β weight for previous disease for one-back effects, and for disease of two trials back for two-back effects. Main effects of C^n_{n-1} and C^n_{n-2}, and their interactions with condition, were all tested using Geisser-Greenhouse correction with the following epsilon values: one-back probability: .564; one-back choice: .628; two-back probability: .785; two-back choice: .662. C^n_{n-1} and C^n_{n-2} = cue commonality measures; Quadratic = the quadratic contrast for C^n_{n-1} or C^n_{n-2}.

two trials. Beyond this modification the implementation of ALCOVE was exactly as in Simulation 1.

Method

The modified version of ALCOVE was simulated separately on all four conditions of Experiment 2. Each simulation consisted of 20,000 trials.

Results

The model's response probabilities were fit using the same regression model as was used in Experiment 2. Again, the focal predictors of this model are D_{n-1} for each value of C^n_{n-1} and D_{n-2} for each value of C^n_{n-2}; that is, the effects of the two previous outcomes and their interactions with cue commonality. The coefficients for these predictors are shown in Figure 6. The graph of the one-back effects (Figure 6A) shows that modified ALCOVE reproduces the increased RE found in the positive condition, along with the opposite effect in the negative condition. As with human subjects, the dual condition shows little or no learning of one-back information (as compared with control). The situation with two-back effects is nearly identical: An increased two-step RE appears in the positive condition while again the dual and control conditions are indistinguishable (as was found empirically). One contrast with the results of Experiment 2 is the negative two-step RE in the negative condition.

Discussion

The results of this second simulation demonstrate that augmentation of extant category-learning models with the capacity for learning of sequential dependencies can provide an adequate fit to the central results presented here. In particular, the simulation shows that the aggregation hypothesis is able to explain the interference and generalization effects across lags that were observed in Experiment 2. These effects occur in the model because the learned association from recent history to present responses comes to reflect an average of the autocorrelation coefficients present in the ecology, and furthermore acts equally at different lags. Therefore in the dual condition the positive and negative autocorrelations average out to produce a near-null association from history to responses. Likewise, in the positive condition there is an effect of events two trials back because they influence the averaged recent history as much as do events one trial back. Thus, the association learned in response to the one-step autocorrelation leads to a two-step RE.

Closer comparison of Figures 5 and 6 does reveal some discrepancies between the empirical and simulated data. For instance, the model produces REs at lags of one and two trials that are approximately equal in magnitude, whereas human subjects showed significantly stronger one-step REs. Also, the model predicts a negative two-step RE in the negative condition for which there was no evidence with humans. Further investigation is required to determine whether these represent true inconsistencies and, if so, how they can be eliminated. Nevertheless, the present simulation shows that perhaps the simplest possible implementation of the aggregation hypothesis provides a reasonable fit to the pattern of adaptive REs found here.

General Discussion

The present experiments have shown clear evidence of REs in category learning analogous to those found in the memory and probability learning literatures. The presence of nominal cues adds complexity to the phenomenon beyond that in uncued learning, with REs stronger when past cases are more similar to the present case. This interaction of RE with stimulus similarity was seen to be nonlinear and, in particular, positively accelerated. If we assume that cue commonality is negatively (and linearly) related to the psychological distance between stimuli, then this observation is

Table 5
Average Responses on Trial 2

	C_1^2			
Statistic	0	1	*2*	3
R_2 (%)	19.1	30.3	54.7	87.2
$P_2(D_1)$ (%)	41.4	48.5	51.9	65.9
SE (%)	2.8	1.7	1.5	3.2
N	47	76	75	39

Note. C_1^2 = the number of matching cue dimensions between the first and second patients; R_2 = proportion of subjects choosing on Trial 2 the correct disease from Trial 1; $P_2(D_1)$ = average probability assigned on Trial 2 to the correct disease from Trial 1; *SE* = standard error of $P_2(D_1)$; *N* = number of subjects at each level of C_1^2.

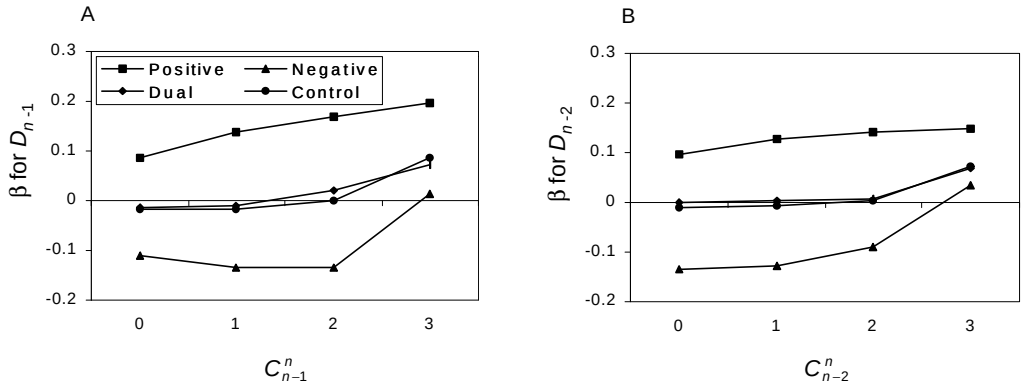

Figure 6. Influence of past outcomes on responses in the modified version of the attentional learning covering map model (Simulation 2). A: Average β weights for D_{n-1} as a function of C^n_{n-1}. B: Average β weights for D_{n-2} as a function of C^n_{n-2}.

consistent with Shepard's (1958, 1987) proposal of exponential decay of generalization, which states that prior learning will affect present behavior to a degree that falls off exponentially with the psychological distance between present and previous stimuli.

Finally, and most significantly, we found that REs can change adaptively in response to levels of autocorrelation in the target sequence, showing that they are more than passive effects. This result is in stark contrast to the predictions of all three categorization models considered, which produce identical patterns of REs regardless of the sequential structure of the task. Although there are many other models of human category learning that we did not explicitly test (e.g., J. R. Anderson, 1991; Ashby & Maddox, 1993; Kruschke, 2001; Nosofsky, Palmeri, & McKinley, 1994; Smith & Minda, 1998), none of these anticipates our results either. Contrary to all existing models of which we are aware, the findings of the present experiments lead us to conclude that subjects are able to adaptively modify their use of recent information in response to temporal statistics of the task. This conclusion is consistent with parallel observations in probability learning (e.g., N. H. Anderson, 1960) and related results from verbal working memory (R. B. Anderson et al., 1997). Further results from causal learning in humans (Matute, Vegas, & De Marez, 2002) and conditioning in animals (Bouton, 1993) show that the context and nature of testing can also affect relative use of recent information. The present results imply that complicated sequence processing is routine even in situations that demand attention to cues that have greater temporal proximity to target events. Hence, mechanisms that support such sequence processing need to be incorporated into models of cued category learning.

The interference effects found for the multiple autocorrelations in the dual condition, as well as the learned two-step RE in the positive condition (which is not supported by the ecology), shed light on the question of how people learn and use sequential information in this sort of task. Specifically, these results support the hypothesis that recent past outcomes are aggregated into a single additional cue that can be used in determining present responses. Subjects thereby learn whether the present case is more or less likely to be in the category that has been recently most prevalent, without distinguishing among sequential information at different lags. This is consistent with the run-based sequence-learning models of Restle (1961) and Gambino and Myers (1967), and contrasts with alternative hypotheses that assume encoding of past information and learning of sequential dependencies is done in terms of separate individual events (Burke & Estes, 1957; Feldman & Hanna, 1966).

Although the three categorization models discussed (AN, ETD, ALCOVE) all failed to predict the sequential-learning effects observed here, they did fare well in terms of their other temporal predictions. The interaction of RE with cue commonality expected by all three models was confirmed, as was the positively accelerated form of the relationship as predicted by the two exemplar-based models. The AN's prediction of a linear function would seem to be a problem for this model; however, a version that assumes additional input nodes for configural cues predicts the nonlinearity (Gluck, 1991; see also Gluck, Bower, & Hee, 1989). Finally, it was seen in the control condition, where sequential learning effects are presumably minimal, that positive REs are still present. This result is consistent with the positive one-step REs found in probability learning with independent trials (Nicks, 1959) and with the overestimation of event repetition probabilities found by Anderson (1960), and is well predicted by all three categorization models. Note that all of the findings just listed clearly contrast with the strictly normative model.[6] Taken together, these results provide support for the sequential predictions that come from all three models. This, plus the fact that the variation of sequential dependencies had no effect on the RE by cue commonality interaction, implies that perhaps the REs induced by the autocorrelation manipulations are mechanistically independent of those correctly predicted by the models (whether the latter come from decay of memory or from iterated error-driven learning). The results of Simulation 2, showing how ALCOVE can be brought to reproduce our central results with the addition of a mechanism for processing sequential information, further support this conclusion. This suggests the possibility that augmenting standard categoriza-

[6] One further result from Experiment 2, not presented here but nevertheless predicted by all three categorization models, was an effect of each of the previous case's cues, in a direction opposite that of the same cue acting on the present trial.

tion models to use sequential information according to the aggregation hypothesis may provide the basis for a unified theory of category learning and sequence learning.

The biggest current challenge we see for such an integrated theory comes from the Trial 2 analyses, which showed a negative RE from Trial 1 when the first two patients had sufficiently different symptom profiles. As mentioned earlier, this effect is observed prior to any possible effects of sequence learning, and is thus a direct measure of the generalization from the stimulus on Trial 1 to that on Trial 2. Intuitively it seems that what is needed to capture this result is an implementation of the idea that a past case that is strongly dissimilar from the present one will be used as evidence against the present case's membership in that past case's category. The potential for this type of negative generalization arises naturally in an AN model that allows for negative input activations (as well as perhaps in certain prototype-based models; cf. Reed, 1972), but would seem to require significant modifications in either of the exemplar-based models due to the multiplicative nature of their similarity rules. However, the real challenge for both types of model comes from the fact that the negative RE is only present early on in the sequence; in our analyses of the full session REs were never significantly negative except in the negative condition. This suggests a Bayesian interpretation, in which the prior expectation of a balanced base rate between the two categories initially produces the effect but is eventually drowned out by subsequent learning. Further research is required to determine whether an approach based on this idea could accurately reproduce the present results.

Regardless of the mechanisms responsible for REs, it is instructive to consider their role in the relationship between normative and descriptive behavior. One possible view of REs is as a type of suboptimal behavior, perhaps resulting from limitations of the human memory system (as is implied by the trace decay assumption of the ETD model). In a stationary ecology with independently sampled outcomes (the norm in experimental research), all past cases are equally informative as to the statistical parameters of the task and the most reliable strategy will therefore weight all past cases equally. Behavior involving a positive RE in this case amounts to basing decisions on a smaller sample of the available information, which leads to higher variance estimates of outcome probabilities.

If, however, recent events happen to be more informative, then REs can in fact be advantageous (cf. Sieck & Yates, 2001). For instance, consider a doctor diagnosing patients, using both symptoms and knowledge of disease base rates. Because diseases occur at different rates at different times (e.g., flu season) and can occur in outbreaks, there is an above-chance tendency for patients who get sick at the same time to have the same disease. Thus, the doctor can benefit from using information about other recently encountered patients to inform current diagnoses. Such a strategy would manifest a positive RE and would be justified by the varying base rates in the ecology. In addition, if there exists the possibility that symptom manifestations associated with diseases are slowly changing over time (perhaps because of mutation) then a temporal bias in learning of cue-category correspondences would be adaptive as well. Such a bias would appear as an interaction between REs and cue commonality of the type reported here. This type of rational analysis of REs in categorization is analogous to similar empirical analyses of need probabilities for memory retrieval as a function of past occurrences, which have been interpreted as providing rational justification for REs observed in memory tasks (Anderson & Schooler, 1991, 2000). Both arguments rely critically on the increased relevance and reliability of recent information, suggesting a deeper functional, and perhaps mechanistic, connection between REs in the two domains.

Another example of the benefit of preferential reliance on recent outcomes comes from experimental work on flower foraging decisions of bumblebees. Real (1991) argued that the short-term memory exhibited by bees' strategies is made advantageous by the presence of spatial autocorrelation in the nectar rewards of flowers (which translates into a temporal autocorrelation in the bee's experience). Dukas and Real (1993) subsequently showed that bees' decisions are based largely on the last flower visited. A similar result was found by Cuthill, Kacelnik, Krebs, Haccou, and Iwasa (1990), concerning flight times in starlings' predatory behavior as a function of recent experience. Thus, it could be argued that REs are an adaptation to a natural environment that is predominantly autocorrelated. Our natural environment is far from a static one, and temporally distant information may often be less reliable than that which is more immediate. Positive REs observed in tasks involving independent sampling could result from subjects coming into the experiment with an expectation (implicit or explicit) that contiguous events are related in that they have a tendency toward the same outcome (but see Gilovich, Vallone, & Tversky, 1985, for an alternative interpretation). A similar argument could be made concerning the RE by cue commonality interaction: Although not normatively justified by the task environment in the present experiments, such behavior could be a manifestation of subjects' default expectation of nonstationarity of cue-category correspondences (in the manner discussed above). If so, then this effect would be expected to extinguish after sufficient training. As is often the case with hypotheses concerning functionality, a direct test is rather difficult, but results like the ones reported here, showing that REs respond adaptively to artificially induced autocorrelation, can be interpreted as evidence in favor of the functional interpretation.

A similar approach was taken by Flood (1954) in his defense of the rationality of probability-matching, another supposed suboptimality in repeated judgments whereby people tend to make responses with frequencies approximating the probabilities of each outcome, rather than simply selecting the most likely alternative every time (Grant, Hake, & Hornseth, 1951). Flood's idea was that subjects in these experiments expect probabilities to change over time, and therefore continue to occasionally explore responses even if they have not paid off well in the past. Although the results of Flood's experimental manipulations designed to influence subjects' expectations of stationarity were less than conclusive, his argument highlights the connection between REs and probability matching: Both have similar ecological justification, and (as a simple argument can show) mechanisms for the former can in turn produce the latter. This along with other work implicating REs as a cause of people's overconfidence in their probability of making correct predictions (Sieck & Yates, 2001) suggests that REs may play a fundamental role in those aspects of human decision-making behavior not captured by current normatively inspired theories.

References

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review, 98,* 409-429.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2,* 396-408.

Anderson, J. R., & Schooler, L. J. (2000). The adaptive nature of memory. In E. Tulving (Ed.), *The Oxford handbook of memory* (pp. 557-570). New York: Oxford University Press.

Anderson, N. H. (1960). Effects of first-order conditional probability in a two-choice learning situation. *Journal of Experimental Psychology, 59,* 73-93.

Anderson, R. B., Tweney, R. D., Rivardo, M., & Duncan, S. (1997). Need probability affects retention: A direct demonstration. *Memory & Cognition, 25,* 867-872.

Ashby, F. G., & Maddox, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology, 37,* 372-400.

Atkinson, R. C., Bower, G. H., & Crothers, E. J. (1965). *An introduction to mathematical learning theory.* New York: Wiley.

Atkinson, R. C., & Estes, W. K. (1963). Stimulus sampling theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology: Vol. 2* (pp. 121-268). New York: Wiley.

Baddeley, A. D. (1986). *Working memory.* New York: Oxford University Press.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114,* 80-99.

Broadbent, D. E., & Broadbent, M. H. P. (1981). Recency effects in visual memory *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 33(A),* 1-15.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106,* 551-581.

Burke, C. J., & Estes, W. K. (1957). A component model for stimulus variables in discrimination learning. *Psychometrika, 22,* 133-146.

Busemeyer, J. R., & Myung, I. J. (1988). A new method for investigating prototype learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 3-11.

Crowder, R. G. (1972). Visual and auditory memory. In J. F. Kavanaugh & I. G. Mattingly (Eds.), *Language by ear and by eye.* New York: McGraw Hill.

Cuthill, I. C., Kacelnik, A., Krebs, J. R., Haccou, P., & Iwasa Y. (1990). Starlings exploiting patches: The effect of recent experience on foraging decisions. *Animal Behaviour, 40,* 625- 640.

Dukas, R., & Real, L. A. (1993). Effects of nectar variance on learning by bumble bees. *Animal Behaviour, 45,* 37-41.

Edwards, W. (1961). Probability learning in 1000 trials. *Journal of Experimental Psychology, 62,* 385-394.

Engler, J. (1958). Marginal and conditional stimulus and response probabilities in verbal conditioning. *Journal of Experimental Psychology, 55,* 303-317.

Estes, W. K. (1957). Theory of learning with constant, variable, or contingent probabilities of reinforcement. *Psychometrika, 22,* 113-132.

Estes, W. K. (1988). Toward a framework for combining connectionist and symbol-processing models. *Journal of Memory and Language, 27,* 196-212.

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

Estes, W. K., Campbell, J. A., Hatsopoulos, N., & Hurwitz, J. B. (1989). Base-rate effects in category learning: A comparison of parallel network and memory storage-retrieval models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 556-571.

Estes, W. K., & Maddox, W. T. (1995). Interactions of stimulus attributes, base rates, and feedback in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1075-1095.

Feldman, J., & Hanna, J. F. (1966). The structure of responses to a sequence of binary events. *Journal of Mathematical Psychology, 3,* 371-387.

Flood, M. M. (1954). Environmental non-stationarity in a sequential decision-making experiment. In R. M. Thrall, C. H. Coombs, & R. L. Davis (Eds.), *Decision processes* (pp. 287- 299). New York: Wiley.

Gambino, B., & Myers, J. L. (1967). Role of event runs in probability learning. *Psychological Review, 74,* 410-419.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17,* 295-314.

Glanzer, M., & Clark, W. H. (1962). The verbal loop hypothesis: Binary numbers. *Journal of Verbal Learning and Verbal Behavior, 2,* 301-309.

Gluck, M. A. (1991). Stimulus generalization and representation in adaptive network models of category learning. *Psychological Science, 2,* 50-55.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117,* 227-247.

Gluck, M. A., Bower, G. H., & Hee, M. R. (1989). A configural-cue network model of animal and human associative learning. *Proceedings of the 11th annual conference of the Cognitive Science Society,* Ann Arbor, MI (pp. 323-332). Hillsdale, NJ: Erlbaum.

Grant, D. A., Hake, H. W. & Hornseth, J. P. (1951). Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *Journal of Experimental Psychology, 42,* 1-5.

Hake, H. W., & Hyman, R. (1953). Perception of the statistical structure of a random series of binary symbols. *Journal of Experimental Psychology, 45,* 64-74.

Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology, 41,* 291-297.

Kruschke, J. K. (1990). A connectionist model of category learning (Doctoral dissertation, University of California, Berkeley, 1990). *Dissertation Abstracts International, 51,* 4624.

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

Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology, 45,* 812-863.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 103-189). New York: Wiley.

Matute, H., Vegas, S., & De Marez, P. (2002). Flexible use of recent information in causal and predictive judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 714-725.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85,* 207-238.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64,* 482-488.

Myers, J. L. (1970). Sequential choice behavior. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation: Vol. 4. Advances in research and theory* (pp. 109-170). New York: Academic Press.

Myers, J. L. (1976). Probability learning. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Vol. 3. Approaches to human learning and motivation* (pp. 171-205). Hillsdale, NJ: Erlbaum.

Nicks, D. C. (1959). Prediction of 2-choice decisions from event runs. *Journal of Experimental Psychology, 57,* 105-114.

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

Nosofsky, R. M., Kruschke, J. K., & McKinley, S. C. (1992). Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 211-233.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104,* 266-300.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. K. (1994). Rule-plus-exception model of classification learning. *Psychological Review, 101,* 53-79.

Real, L. A. (1991, August 30). Animal choice behavior and the evolution of cognitive architecture. *Science, 253,* 980-986.

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology, 3,* 382-407.

Restle, F. (1961). *Psychology of judgment and choice.* New York: Wiley.

Rose, R. M., & Vitz, P. C. (1966). The role of runs in probability learning. *Journal of Experimental Psychology, 72,* 751-760.

Royer, F. L. (1967). Sequential complexity and motor response rates. *Journal of Experimental Psychology, 74,* 199-202.

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

Schacter, D. L. (1999). The seven sins of memory. Insights from psychology and cognitive neuroscience. *American Psychologist, 54,* 182-203.

Shepard, R. N. (1958). Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space. *Journal of Experimental Psychology, 55,* 509-523.

Shepard, R. N. (1987, September 11). Toward a universal law of generalization for psychological science. *Science, 237,* 1317-1323.

Sieck, W. R. (2000). *Overconfidence in judgment for repeatable events.* Unpublished doctoral dissertation, University of Michigan.

Sieck, W. R., & Yates, J. F. (2001). Overconfidence effects in category learning: A comparison of connectionist and exemplar memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1003-1021.

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1411- 1436.

Suppes, P., & Atkinson (1960). *Markov learning models for multi-person interactions.* Stanford, CA: Stanford University Press.

Thompson, R. K., & Herman, L. M. (1977, February 4). Memory for lists of sounds by the bottle-nosed dolphin: Convergence of memory processes with humans? *Science, 195,* 501-503

Witte, R. S. (1964). Long-term effects of patterned reward schedules. *Journal of Experimental Psychology, 68,* 588-594.

Wright, A. A., Santiago, H. C., Sands, S. F., Kendrick, D. F., & Cook, R. G. (1985, July 19). Memory processing of serial lists by pigeons, monkeys, and people. *Science, 229,* 287- 289.

Yates, J. F., Lee, J-W., Shinotsuka, H., Patalano, A. L., & Sieck, W. R. (1998). Cross-cultural variations in probability accuracy: Beyond general-knowledge overconfidence? *Organizational Behavior and Human Decision Processes, 74,* 89-117.

Appendix: Model Details

The implementation details of the models in both simulations are given as follows. Parameter values for the data presented (see Figures 3 and 6) as well as ranges of other values tested for Simulation 1 (all giving the same qualitative predictions) are displayed in Table A1.

Table A1
Parameter Values for Simulations

		Tested	
Parameter	Shown	Low	High
AN			
ε	.10	.01	.25
θ	1.00	.10	10.00
ETD			
s	.14	.01	.5
t	.90	.10	.99
C	1.00	0.00	10.00
ALCOVE			
c	1.00	.10	10.00
ε_{ex}	.20	.01	.50
ε_{att}	.10	0.00	.50
ϕ	1.00	.10	10.00

Note. Tested columns give ranges of other values tested in Simulation 1. Shown column indicates parameter values used for data presented. AN = adaptive network; ETD = exemplar model with trace decay; ALCOVE = attentional learning covering map; ε = learning rate; θ = scaling parameter; *s* = similarity parameter; *t* = decay parameter; *C* = background noise constant; *c* = hidden node tuning specificity; ε_{ex} = learning rate for association weights; ε_{att} = learning rate for attention weights; ϕ = choice temperature.

Adaptive Network (AN)

Each input node takes value 1 when its cue is present and –1 when absent. Activation of the context node is constant at 1. Activation *u* of the output node is given by

$$u = \sum_i a_i \cdot w_i, \tag{A1}$$

where *i* indexes input nodes and a_i and w_i represent the corresponding activities and weights, respectively. The probability of responding with Category A is given by a sigmoid transformation of *u:*

$$\mathrm{P(A)} = \frac{1}{1+e^{-\theta u}}, \tag{A2}$$

where 0 > 0 is a scaling parameter. After feedback is given, weights are updated according to

$$\Delta w_i = \varepsilon \cdot a_i \cdot (f - u), \tag{A3}$$

where ε is the learning rate. The feedback *f* is equal to 1 if Category A was correct and 0 if B was correct.

Exemplar with Trace Decay (ETD)

Similarity of a stored exemplar *k* to the present case *n* is calculated as

$$S(k,n) = s^d \text{ ,} \tag{A4}$$

where *d* gives the number of cue dimensions on which the cases disagree. The similarity parameter *s* is taken to be between 0 and 1. Response probability is given as

(Appendix continues)

$$\mathrm{P(A)} = \frac{C + \sum_{\mathrm{A}} S(k,n) \cdot t^{n-k}}{C + \sum_{\mathrm{A}} S(k,n) \cdot t^{n-k} + \sum_{\mathrm{B}} S(k,n) \cdot t^{n-k}} \,. \qquad \text{(A5)}$$

Here the sums are taken separately over those stored exemplars for Category A and those for B. The decay parameter t, strictly between 0 and 1, controls how quickly exemplars fade from memory. The background noise constant C serves to bias response probabilities towards 50%, especially early in the session.

ALCOVE

Activation a_i of the i^{th} input node, coding the i^{th} cue for the present case, is 1 when that cue is present and 0 otherwise. For the additional sequential cue used in Simulation 2, a is equal to the number of times trebitis was correct over the last two trials. Activation b_j of the jth hidden unit is then calculated as:

$$b_j = e^{-c\sum_i \alpha_i |h_{ji} - a_i|} \,. \qquad \text{(A6)}$$

Here h_{ji} is the value on the i^{th} cue dimension of the exemplar represented by hidden node j, and the α_i represent attention weights. The parameter c determines the tuning specificity of the hidden nodes.

Activation of output nodes is determined by a weighted sum of the hidden activations:

$$o_k = \sum_i w_{kj} \cdot b_j \,, \qquad \text{(A7)}$$

where w_{kj} is the weight from hidden node j to output node k. Response probability is then given by:

$$\mathrm{P(A)} = \frac{e^{\phi \cdot o_{\mathrm{A}}}}{e^{\phi \cdot o_{\mathrm{A}}} + e^{\phi \cdot o_{\mathrm{B}}}} \,, \qquad \text{(A8)}$$

where ϕ is the choice temperature parameter.

The weights w_{kj} and α_i are updated by means of back-propagation, with respective learning rates $\varepsilon_{\mathrm{ex}}$ and $\varepsilon_{\mathrm{att}}$. The feedback signal given to each output node is equal to 1 if the corresponding category was correct, and –1 otherwise; however the δ-signal is taken to be 0 if the actual activation overshoots the desired value (e.g., an activation of 2 at the correct category node).

Received September 6, 2002
Revision received January 14, 2003
Accepted January 15, 2003 ■