Through the looking glass: A dynamic lens model approach to multiple cue probability learning.

Maarten Speekenbrink and David R. Shanks
Department of Psychology
University College London

Despite what the somewhat technical name might suggest, multiple cue probability learning (MCPL) problems are commonly encountered in daily life. For instance, we may have to judge whether it will rain from cues such as temperature, humidity, and the time of year. Or, we may have to judge whether someone is telling the truth from cues such as pitch of voice, level of eye contact, and rate of eye blinks. While informative, these cues are not perfect predictors. How do we learn to solve such problems? How do we learn which cues are relevant, and to what extent? How do we integrate the available information into a judgement?

Applying a rational analysis (Anderson, 1990; Oaksford & Chater, 1998), we would answer these questions by specifying a rational model, and then compare individuals' judgements to the model-predicted judgements. Insofar as observed behaviour matches the predicted behaviour, we would conclude that people learn these tasks as rational agents. Here, we take a slightly different approach. We still use rational models, but rather than comparing predicted behaviour to actual behaviour (a comparison in what we might call observation space), we make the comparison in parameter space.

To be a little less obscure, let's take an agent who must repeatedly predict share price from past share price. A rational agent would make predictions which are optimal given an observed pattern of past share price. The question is whether a real (human) agent makes predictions as the rational agent would. Assume current share price Y_t is related to past share price Y_{t-1} as $Y_t = \beta Y_{t-1}$, where β is an unknown constant. In order to make accurate predictions, the agent must infer the value of β from repeated observations of share price. A rational agent, with optimal estimates w_t, will make predictions $\hat{y}_t = w_t Y_{t-1}$. Since the relation is rather simple, and the agent is rational, the inferences w_t (and hence predictions $\hat{y}_t$) will be accurate quite quickly. Enter the real (rational?) agent, making predictions R_t. Rather than assuming these predictions follow from rational estimates w_t, we assume they are based on $R_t = u_t Y_{t-1}$, where u_t is a coefficient not necessarily equal to w_t. Thus, we assume the same *structural* model for rational and actual predictions, but allow for different parameters u_t and w_t. By comparing u_t to w_t, we can see how the real agents' learning compares to the rational agents' learning.

Moreover, we can compare both u_t and w_t to the true value β, to see how well the rational and real agent adapt to the environment.

The values of u_t can be estimated in different ways. A common approach is to assume that the real agent learns in a similar way to the rational model; for instance, by changing u_t according to the error of predictions. To match the values of u_t to the real agents' predictions, the changes are allowed to be of a sub-optimal magnitude. The fitted learning model is then used to derive the values of u_t. This procedure is reasonable if the agent learns according to the model. However, as the estimates of u_t are constrained by the learning model, they will be biased if the agent learns in a different way. An alternative approach is to make minimal assumptions regarding the process by which u_t changes (i.e., not assuming the change is related to the error of prediction), and estimate u_t directly from the predictions and cues (e.g. Kelley & Friedman, 2002; Lagnado, Newell, Kahan, & Shanks, 2006). The resulting "unconstrained" estimates will not be biased towards the learning model, and hence provide more information regarding its validity.

The lens model

Our analysis owes much to Brunswik's (1955) lens model. The lens model has its origins in perception, and was proposed to describe how organisms perceive a distal (unobservable) criterion, through proximal (observable) cues. Hammond, Wilkins, and Todd (1966) formulated the lens model in statistical terms, and proposed it as a general framework for judgement analysis. Since then, the lens model has been a standard tool in the study of multiple cue probability learning (Cooksey, 1996). Figure 1 depicts a simple version of the lens model. The left hand side of the lens represents the environment, consisting of the cues and the criterion. How well an individual can perceive, judge, or predict the criterion (we will use the generic term "respond" from now on), depends on the relation between the cues and the criterion. Adopting the standard terminology, we will refer to these relations as *cue validity*. The right hand side of the lens represents the response system, consisting of the cues and the responses. The relations between the cues and the responses reflect how an individual actually uses the cues in forming a response. These relations are referred to as *cue utilization*. How well an individual performs in a multiple cue task depends on the relation between the criterion and the responses; central to the lens model is that this relation can be ascertained by comparing cue validity and utilization.

Usually, the parameters of the lens model are estimated by multiple regression analysis. For instance, estimates of cue validity are obtained as the regression coefficients of a model which relates the criterion to the cues. Similarly, by regressing the responses onto the cues, we can obtain estimates of cue utilization. In its original form, the lens model is static, and provides constant cue validity and utilization weights for a period in time (e.g., the duration of the task). As such, these estimates are valid under the assumption that the environment and the response system are stationary. However, in a typical MCPL task, the environment is unknown to the agent, who must infer its structure as the task progresses. As such, even if the environment is stationary, the response system (which relies on these inferences) will be non-stationary. Of course, we could apply the lens model after an extensive period of training, such that learning will have converged to a particular representation of the environment. However, we are interested in the learning process itself, in how an agent learns to adapt his/her predictions to the structure of the environment. To study this process, we must deal with non-stationarity.

Previously, "rolling regression" has been proposed as a method to estimate the changes
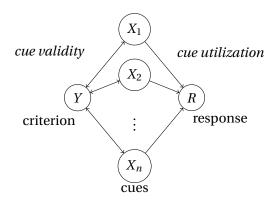
Figure 1. The lens model.

in cue utilization (Kelley & Friedman, 2002; Lagnado et al., 2006). In rolling regression, the regression model is repeatedly fitted to a moving window of trials. For a window of size W, cue validity and utilization at trial t are estimated from regression models applied to observations from trials $t - W + 1$ to t. While this method is a straightforward extension of usual lens model analysis, it has certain drawbacks. For one, utilization and validity can only be estimated from trial W onwards. Since the window size may have to be large to get reliable estimates, this will be problematic if we are interested in the early stages of learning. And although the window size allows manipulation of something akin to memory limits, the type of memory represented is not very plausible. More particularly, each observation within the window affects estimates to a similar extent, but observations outside the window have no effect. A more gradual form of memory loss seems more appropriate. Finally, as a method of tracking slowly changing weights, rolling regression is not optimal in a statistical sense. In this chapter, we will propose an alternative method, which has clear advantages over rolling regression.

Overview

In the remainder of this chapter, we will first describe a generalized lens model, which is still static, but incorporates situations in which linear regression is unsuitable. We then extend this framework into a dynamic lens model, which is suitable for non-stationary environments and response systems. Of particular interest in the context of rational analysis is that we can replace part of the model with a rational learning model. This allows a comparison between the structure of the environment, optimal inference, and actual inference as evident from the responses made. Finally, we will illustrate the methods by re-analysing data from an earlier MCPL study (Lagnado et al., 2006).

A framework for MCPL

In multiple cue probability learning tasks, the objective is to make a prediction R of the value of a criterion variable Y, on the basis of a number of cues X_j. The relation between criterion and cues $\mathbf{X} = (X_1, \ldots, X_n)^\mathsf{T}$ is probabilistic, and each cue pattern is associated with a probability

distribution over the possible values of Y,

$$P(Y|\mathbf{x}) = P(Y, \mathbf{X} = \mathbf{x})/P(\mathbf{X} = \mathbf{x}). \tag{1}$$

This conditional probability distribution is initially unknown, and has to be inferred from repeated observations of training pairs $(y, \mathbf{x})$ drawn from the environmental distribution $P(Y, \mathbf{X})$.

Generalized linear model

We consider environments which fit in the framework of generalized linear models (McCullagh & Nelder, 1983, see also Dobson, 2002). That is, we assume the conditional distribution $P(Y|\mathbf{X})$ to be a member of the exponential family, and the existence of a function g and parameters $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_m)^\mathsf{T}$ such that

$$g(\mu_x) = \mathbf{z}^\mathsf{T} \boldsymbol{\beta}, \tag{2}$$

where T denotes the (matrix) transpose. The link function g linearises the relation between the criterion and the *effective cue pattern* $\mathbf{z}$, which is a vector of dimension m. If Y is Normally distributed, a good choice is the identity function $g(\mu_x) = \mu_x$. When the distribution of Y is binomial, a logit function $g(\mu_x) = \log(\mu_x/(1 - \mu_x))$ is often used. The effective cue pattern $\mathbf{z}$ is related to the actual cue pattern $\mathbf{x}$ (a vector of dimension n) as

$$\mathbf{z} = h(\mathbf{x}). \tag{3}$$

Often, h will be the identity function, so that the effective cue pattern is the original cue pattern. This corresponds to the assumption that the cues affect the criterion independently. If this assumption does not hold, the effective cue vector can be extended to include interaction between the cues. In this case, the dimension of $\mathbf{z}$ will differ from the dimension of $\mathbf{x}$. When the cues are categorical variables, function h will map the cue patterns to corresponding patterns of dummy variables. Finally, function h can also represent the mapping of cue patterns to their representations in some psychological space.

The lens model fits neatly into our framework. We assume that $P(Y|\mathbf{X})$ and $P(R|\mathbf{X})$ are of the same form (e.g., both normal or both binomial), with expectations $E(Y|\mathbf{x}) = E(Y|\mathbf{z}) = \mu_z$ and $E(R|\mathbf{x}) = E(R|\mathbf{z}) = \rho_z$, that g and h are identical for both the environment and response system,

$$g(\mu_z) = \mathbf{z}^\mathsf{T} \mathbf{v}, \tag{4}$$

and

$$g(\rho_z) = \mathbf{z}^\mathsf{T} \mathbf{u}. \tag{5}$$

Here, $\mathbf{v}$ represents cue validity, and $\mathbf{u}$ cue utilization. For our purposes, it is crucial that h, and as a result the effective cue vector $\mathbf{z}$, is identical in both models. Without this equivalence, cue validity and utilization are not directly comparable.

Since the dimension of $\mathbf{v}$ is usually smaller than the number of possible cue patterns, learning is simplified. Rather than having to infer the conditional distribution (or expected value) of the criterion for each cue pattern $\mathbf{x}$ directly, the agent infers the value of $\mathbf{v}$, and determines the conditional distribution from these. Estimates of $\mathbf{v}$ will be more reliable than direct estimates of $P(Y|\mathbf{x})$ or $E(Y|\mathbf{x})$, and also require less storage in memory than a direct representations of these.

Optimal responses

Associated with the response and the criterion is a loss function $L(R, Y)$, and the objective for the agent is to give responses $R = r_x^*$ to cue patterns $\mathbf{x}$ which minimize the expected loss

$$r_x^* = \arg\min_r E(L(r, Y)|\mathbf{x}). \tag{6}$$

For real valued criterion and predictions, a squared loss function

$$L(r, y) = (r - y)^2, \tag{7}$$

is often used, and the optimal prediction is the expected value

$$r_x^* = E(Y|\mathbf{x}) \equiv \mu_x. \tag{8}$$

In terms of cue validity, this can be rewritten as

$$r_x^* = g^{-1}(\mathbf{z}^\mathsf{T}\mathbf{v}), \tag{9}$$

where g^{-1} denotes the inverse of the link function g. For nominal criterion and responses in the same set, a 0-1 loss function

$$L(r, y) = \begin{cases} 0 & \text{if } r = y \\ 1 & \text{if } r \neq y \end{cases}, \tag{10}$$

is usually appropriate, and the optimal prediction is

$$r_x^* = \arg\max_y P(Y = y|\mathbf{x}). \tag{11}$$

When the criterion $Y = \{0, 1\}$ is dichotomous, this can be rewritten in terms of cue validity as

$$r_x^* = \begin{cases} 0 & \text{if } \mathbf{z}^\mathsf{T}\mathbf{v} < 0 \\ 1 & \text{otherwise} \end{cases}. \tag{12}$$

Dynamic lens model

As noted earlier, standard lens model analysis is static, by which we mean that validity and utilization are assumed to be constant over trials t. In standard MCPL tasks presented in laboratory studies, the environment is often stationary, so that this assumption will hold for cue validity. However, as a result of learning, cue utilization will change with t. To study learning, we need to take a dynamic viewpoint in which we explicitly focus on trial-by-trial changes in cue utilization.

In dynamic lens model analysis, both cue validity and utilization are allowed to vary over time. A graphical representation of the structure of the dynamic lens model is given in Figure 2. To account for the changes in cue validity and utilization, we need a (stochastic) model for the change process. In most MCPL experiments, the environment is under control of the experimenter, and the dynamics of cue validity will be known. Determining the dynamic process of cue utilization will be more challenging. While there are many possibilities, the assumption of a simple random walk will suffice in most cases. That is, we assume

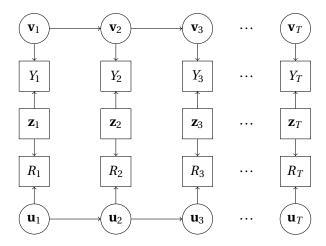$$\mathbf{u}_t = \mathbf{u}_{t-1} + \mathbf{e}_t, \tag{13}$$

Figure 2. The dynamic lens model. Observed variables are depicted in squares, while unobserved (latent) variables are depicted in circles. The criterion Y_t is dependent on the effective cue vector $\mathbf{z}_t$ and cue validity $\mathbf{v}_t$. The response R_t depends on the effective cue vector and cue utilization $\mathbf{u}_t$. Cue validity $\mathbf{v}_t$ depends on previous cue validity $\mathbf{v}_{t-1}$, and utilization $\mathbf{u}_t$ on previous utilization $\mathbf{u}_{t-1}$.

with $\mathbf{e}_t \sim N(\mathbf{0}, \Sigma_u)$, i.e., following a zero-mean multivariate normal distribution with covariance matrix Σ_u. The model for the responses resulting from the combination of Equations 5 and 13 is known as a dynamic generalized linear model (West & Harrison, 1997). When the process for environmental change is also unknown, a similar model can be used for cue validity.

Estimation of $\mathbf{u}_t$ from all observations $R_t, \mathbf{x}_t$, $t = 1, \ldots, T$, where T denotes the total number of trials, is known as smoothing. When Σ_u is known and the conditional distribution $p(R_t | \mathbf{x}_t, \mathbf{u}_t)$ is Normal with known variance, optimal estimates can be derived analytically, and the solution is known as the Kalman smoother (e.g., Durbin & Koopman, 2001; West & Harrison, 1997). When R_t is not normally distributed, analytical solution are usually unavailable, and approximate methods must be used. In the application described later, we will use a Monte Carlo Markov Chain (MCMC) procedure for this purpose.

Incorporating rational models

In the dynamic lens model, trial-by-trial variation in utilization is assumed to be independent of (variation in) cue validity. While this assumption may be false if an agent makes accurate inferences of cue validity, the assumption reflects the wish to "let the data speak for itself". If validity and utilization are indeed related, this should be evident from the (estimated) validity and utilization weights. As such, the dynamic lens model allows one to analyse learning without making explicit assumptions regarding how this learning proceeds. Of course, rational models do make such assumptions, which is part of their strength. To incorporate rational learning models, we need to extend the response part of the dynamical lens model.

A graphical representation of the extended model is given in Figure 3. The model now consists of three related sub-models, one for the environment, one for learning, and one for the responses. The model of the environment, consisting of the criterion, cues and cue validity, is identical to the one in Figure 2. The learning model consists of the criterion, effective cue pattern, and inference. Here, the agent infers validity $\mathbf{v}_t$ from training pairs $(y_t, \mathbf{z}_t)$. We will denote
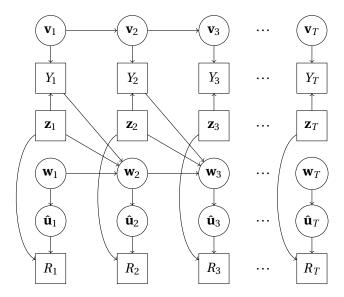
Figure 3. A general model of learning and prediction in MCPL. The criterion Y_t depends on the effective cue vector $\mathbf{z}_t$ and cue validity $\mathbf{v}_t$. Inferred cue validity $\mathbf{w}_{t+1}$ depends on the previous inference $\mathbf{w}_t$, effective cue vector $\mathbf{z}_t$, and criterion value y_t. Predicted cue utilization $\hat{\mathbf{u}}_t$ depends on the inferred cue validity $\mathbf{w}_t$. The prediction R_t depends on predicted cue utilization $\hat{\mathbf{u}}_t$ and effective cue vector $\mathbf{z}_t$.

the inference of $\mathbf{v}_t$ as $\mathbf{w}_t$. As before, the response model consists of the responses, effective cue patterns, and cue utilization. In this context, utilization is referred to as predicted utilization $\hat{\mathbf{u}}_t$, to distinguish it from the utilization in Figure 2. Predicted utilization does not depend on previous utilization, but on the current inferred cue validity $\mathbf{w}_t$, and changes in $\hat{\mathbf{u}}_t$ are taken to depend completely on changes in $\mathbf{w}_t$.

The models in Figures 2 and 3 can be made equivalent by removing the links from $\mathbf{z}_{t-1}$ and $\mathbf{y}_{t-1}$ to $\mathbf{w}_t$, and assuming identity of $\mathbf{w}_t$ and $\hat{\mathbf{u}}_t$. However, learning models which incorporate these links are of main interest here. We will describe such models in the next section.

Learning in the Weather Prediction Task

Up to now, the models have been described in general terms. We will now apply them to an MCPL task which has received much attention over recent decades, especially in the field of neuroscience. In the so-called Weather Prediction Task (Knowlton, Squire, & Gluck, 1994), the objective is to predict the state of the weather (sunny or rainy) on the basis of four "Tarot cards" (cards with geometrical patterns). On each trial t, one to three cards are presented, and the agent is asked whether the weather will be sunny or rainy. Each cue pattern $\mathbf{x}$ is associated with the state of the weather with a different probability. These probabilities are given in Table 1. The environment is constructed so that each card is associated with the outcome with a different probability. For example, the probability of rainy weather is .2 over all the trials on which card 1 is present, .4 for trials on which card 2 is present, .6 for trials on which card 3 is present, and .8 for trials on which card 4 is present. Thus, two cards are predictive of rain, one strongly (card 4), one weakly (card 3), and two cards are predictive of sun, one strongly (card 1), one weakly (card 2). A more detailed description of the weather prediction task can be found in Lagnado et al. (2006).

Table 1: Learning environment in the Weather Prediction Task. P(sun|pattern) denotes the actual conditional probability of sunny weather, while m_z is the predicted probability (see text).

Pattern	Cards present	$\mathbf{z}$	Total	P(pattern)	P(sun|pattern)	m_z
A	4	0001	19	0.095	0.895	0.891
B	3	0010	9	0.045	0.778	0.641
C	3,4	0011	26	0.130	0.923	0.936
D	2	0100	9	0.045	0.222	0.359
E	2,4	0101	12	0.060	0.833	0.821
F	2,3	0110	6	0.030	0.500	0.500
G	2,3,4	0111	19	0.095	0.895	0.891
H	1	1000	19	0.095	0.105	0.109
I	1,4	1001	6	0.030	0.500	0.500
J	1,3	1010	12	0.060	0.167	0.179
K	1,3,4	1011	9	0.045	0.556	0.641
L	1,2	1100	26	0.130	0.077	0.064
M	1,2,4	1101	9	0.045	0.444	0.359
N	1,2,3	1110	19	0.095	0.105	0.109
Total			200	1.00		

Environment

The environment consists of the criterion Y_t and the cue patterns $\mathbf{x}_t$. The model of the environment consists of the criterion, effective cue patterns $\mathbf{z}_t$, and cue validity weights $\mathbf{v}_t$. The environment is stationary, so that cue validity is independent of trial t, and $\mathbf{v}_t = \mathbf{v}_{t+1} = \mathbf{v}$, for all t. The criterion is dichotomous, and $P(Y_t = y|\mathbf{z}_t)$ is constant and a Bernouilli distribution,

$$P(Y = y|\mathbf{x}) = \mu_x^y (1 - \mu_x)^{1-y}, \tag{14}$$

where we use the scoring of $y = 1$ for rainy, and $y = 0$ for sunny weather. Since the cues are categorical variables, with g being the logit-function, there will exist a function h and vector $\mathbf{v}$ such that Equation 4 will hold[1]. The resulting model can be recognised as a logistic regression model. For simplicity, we will let $\mathbf{z}$ be a binary vector of dimension 4, in which the elements are binary dummy variables for the main effects of the four cues. In other words, the outcome is assumed to be a function of the additive combination of the individual cue validities. The resulting values for $\mathbf{z}$ can be found in Table 1, and the associated cue validities are $\mathbf{v} = (-2.10, -.58, .58, 2.10)^\mathsf{T}$.

We should note that the model is not entirely correct, and the predicted values $m_z = g^{-1}(\mathbf{z}^\mathsf{T}\mathbf{v})$ only approximate μ_x. In the correct model, $\mathbf{z}$ also contains dummy variables for all cue interactions, resulting in a vector of dimension 14 rather than 4. For practical purposes, the resulting 10 additional parameters were deemed unnecessary (and possibly detrimental, as the increase in parameters will decrease the reliability of their estimates, and result in generally poorer inferences in a limited data set). As can be seen in Table 1, the approximation with $\mathbf{z}$ is generally

---

[1]In general, for a binary cue vector $\mathbf{x}$, h can be a function which maps $\mathbf{x}$ to a set of 2^J dummy variables (J is the dimension of $\mathbf{x}$), representing an intercept, the main effects, and all possible (two-way, three-way, etc.) interactions between the cues. The resulting saturated model has the same number of validity parameters as possible cue patterns and can fit any set of conditional probability distributions $P(Y|\mathbf{x})$ for these possible patterns $\mathbf{x}$ perfectly.

quite close, but fails somewhat for the relatively rare patterns B, D, K and M. This is due to the fact that the model assumes $P(Y|\text{pattern B}) = P(Y|\text{pattern K})$ and $P(Y|\text{pattern D}) = P(Y|\text{pattern M})$, while in reality, they are different (e.g., the effect of card 2 is stronger in the presence of cards 1 and 4 than in their absence). Finally, as the deviations from the predictions m_z are not drastic, and there is much evidence to suggest that simple additive rules of cue integration are a default strategy (Cooksey, 1996; Einhorn, Kleinmuntz, & Kleinmuntz, 1979; Hastie & Dawes, 2001), we expect $\mathbf{z}$ to be adequate to model the responses.

Learning process

The objective for the agent is to infer cue validity $\mathbf{v}$, and then base responses on these inferences. There are many ways in which cue validity can be inferred. As depicted in Figure 3, a restriction we place is that observations of the cues and criterion affect inferences only once. In other words, we assume an on-line learning process, such that learning consists of updating inference $\mathbf{w}_t$ on the basis of y_t and $\mathbf{z}_t$. We will describe two learning strategies in more detail, a Bayesian and an associative procedure.

Bayesian learning.

On-line learning has a natural representation in Bayesian terms. In Bayesian learning, a learner represents the uncertainty regarding $\mathbf{v}$ as a probability distribution over $\mathbf{v}$. After each trial t, this probability distribution is updated by incorporating the new evidence presented on that trial. Letting $P_t(\mathbf{v})$ denote the prior probability distribution, the posterior distribution after trial t is the product of the prior distribution $P_t(\mathbf{v})$ and the likelihood $P(y_t|\mathbf{v}, \mathbf{z}_t)$ of $\mathbf{v}$ for observations y_t and $\mathbf{z}_t$, i.e.

$$P_{t+1}(\mathbf{v}) = P(y_t|\mathbf{v}, \mathbf{z}_t)P_t(\mathbf{v})/K \tag{15}$$

where the proportionality constant $K = \int P(y_t|\mathbf{v}, \mathbf{z}_t)P_t(\mathbf{v})d\mathbf{v}$ ensures that $P_{t+1}(\mathbf{v})$ is a probability distribution. The likelihood is given in Equation 14. If the agent has to give a point estimate $\mathbf{w}_{t+1}$ of $\mathbf{v}$, (s)he may give the posterior expectation of $\mathbf{v}$, or the maximum a posteriori (MAP) value of $\mathbf{v}$. When we represent an agents' learning state in terms of inference $\mathbf{w}_t$, we choose the first option.

Associative learning.

In associative learning, a learner adapts associations between cues and criterion on the basis of observations. The Rescorla-Wagner model (Rescorla & Wagner, 1972) is a long-time favourite for associative learning, and is equivalent to a single-layer feed-forward network, where the weights are adapted according to the LMS (or delta) rule (Gluck & Bower, 1988). We use a somewhat different representation than Gluck and Bower (1988) with a logistic rather than linear activation function. This has the advantage that the model provides direct estimates of the (conditional) probability of the criterion, rather than having to transform the activation later to derive these probabilities. The associative model learns by a gradient descent on the cross-entropy error (Bishop, 1995), resulting in the following rule for trial-by-trial updates of inferred cue validity:

$$\mathbf{w}_{t+1} = \mathbf{w}_t + \eta_t(y_t - m_{t,z})\mathbf{z}_t, \tag{16}$$

in which η_t stands for the *learning rate*, and $m_{t,z}$ is the current inference of the expected value of the criterion

$$m_{t,z} = g^{-1}(\mathbf{z}_t^T \mathbf{w}_t), \tag{17}$$

with g^{-1} denoting the inverse of the (logit) link function. Note that if an identity link function were used, Equation 16 would be identical to the weight updates of Gluck and Bower (1988). The effect of the logit link function is that the absolute weights can grow without bound, while the predictions $m_{t,z}$ are still on the same scale as the criterion Y_t.

Under certain conditions[2], it can be shown that repeated application of Equation 16 will result in inferences $\mathbf{w}_t$ which converge to $\mathbf{v}$ as $t \to \infty$ (Robbins & Monro, 1951). A simple scheme obeying these conditions is $\eta_t = \eta/t$. Another option, which does not guarantee convergence, is a constant learning rate $\eta_t = \eta$, for which Equation 16 becomes the LMS rule (Gluck & Bower, 1988). With such a constant learning rate, $\mathbf{w}_t$ will fluctuate around $\mathbf{v}$ as $t \to \infty$, with variance roughly proportional to η. Although inference with a constant learning rate is not consistent in the statistical sense (the variance of estimations does not approach 0 as $t \to \infty$), it is preferable over a decreasing learning rate when the environment is non-stationary (i.e., when $\mathbf{v}_t$ changes with t).

Response process

Having inferred the cue validity weights, the agent must form a prediction. Under squared loss, the (subjectively) optimal prediction would be $r_{t,z}^* = m_{t,z}$, as in Equation 8. However, since the agent must predict the state of the weather, and not the probability of a state of the weather, this optimal response is not an allowed prediction. For responses $R_t = \{0, 1\}$ in the same set as the criterion, squared loss is identical to 0-1 loss, and the optimal prediction is as in Equation 11, which can be restated as

$$r_{t,z}^* = \begin{cases} 0 & \text{if } m_{t,z} < .5 \quad (\text{or equivalently } \mathbf{z}_t^T \mathbf{w}_t < 0) \\ 1 & \text{otherwise} \end{cases}. \tag{18}$$

Since the optimal prediction is a deterministic function of $\mathbf{z}_t$ and $\mathbf{w}_t$, the distribution $P(R_t|\mathbf{z}_t)$ for an optimally responding agent is degenerate (all probability mass is on a single point). This is somewhat problematic for our model; therefore, we let the variability of predictions approach 0, rather than being 0 exactly. More precisely, we assume that responses are made by a variant of Luce's (1959) choice rule

$$P(R_t = 1|\mathbf{z}_t) = \frac{m_{t,z}^\lambda}{m_{t,z}^\lambda + (1 - m_{t,z})^\lambda}, \tag{19}$$

in which parameter λ determines the *consistency* of responses. The effect of this parameter is depicted in Figure 4. For a rational agent, $\lambda \to \infty$, so that the relation between $m_{t,z}$ (the inferred probability of the criterion) and the probability of the response approaches a step function. The distribution $P(R_t|\mathbf{z}_t)$ is still a Bernouilli distribution, with parameter $\rho_{t,z} = g^{-1}(\mathbf{z}_t^T \hat{\mathbf{u}}_t)$. Finally, it is easy to prove (e.g. Speekenbrink, Channon, & Shanks, 2007) that the assumed response function results in a linear relation between inferred validity and predicted utilization

$$\hat{\mathbf{u}}_t = \lambda \mathbf{w}_t. \tag{20}$$

For a rational agent, absolute utilization weights are thus infinitely larger than the absolute inferred validity weights. We will assume a similar response rule for actual individuals, but with λ possibly not approaching infinity.

---

[2]These conditions are (1) $\lim_{t \to \infty} \eta_t = 0$, (2) $\sum_{t=1}^\infty \eta_t = \infty$, and (3) $\sum_{t=1}^\infty \eta_t^2 < \infty$.
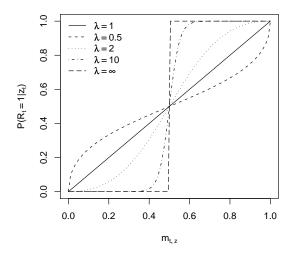
Figure 4. Effect of response consistency λ on the relation between $m_{t,z}$, the (inferred) expected value of the criterion, and $P(R_t = 1|\mathbf{z}_t)$, the probability of response.

Application

Lagnado et al. (2006, Experiment 1) collected data from 16 healthy human subjects on the Weather Prediction Task, and used rolling regression analysis to estimate rational inference and cue utilization. Here, we will re-analyse their data using the dynamic lens and rational models.

Are Bayesian and associative learning rational?

The two learning models have different properties. Here, we will address how they compare, assuming a maximising response strategy. To derive inferences $\mathbf{w}_t$ from Bayesian learning, we discretisized the parameter space (i.e., the domain of $\mathbf{v}$) as a multidimensional grid with 7^4 equally spaced points, and then applied exact Bayesian inference on this grid[3]. A uniform prior $P_1(\mathbf{v})$ was used. Such a uniform prior is optimal when no prior information about $\mathbf{v}$ is present, and results in rapid convergence of $\mathbf{w}$ to $\mathbf{v}$. For the rational version of the associative model with decreasing learning rate, we used a learning rate of $\eta_t = 5.66/t$, where the constant (5.66) was derived as 1 over the largest eigenvalue of the expected Hessian matrix, evaluated at the optimal parameter vector. This is an optimal learning rate for the environment (e.g., LeCun, Bottou, Orr, & Miller, 1998). For the associative model with constant learning rate, we used a value of $\eta_t = .32$ (this value resulted in the best predictions in a simulation study with 500 randomly generated WP tasks of 200 trials each). For both associative models, the weights $\mathbf{w}_1$ were initialized to 0.

We used the three learning models to form inferences $\mathbf{w}_t$ in 1000 simulated Weather Prediction Tasks of 200 trials each. These inferences were then used to form responses by a maximising strategy, as in Equation 18. We then compared the expected performance of the rational agents between the different models. For each set of trials, expected performance is computed

---

[3]We used grid points -3,-2,-1,0,1,2 and 3 for each cue.

as

$$\frac{1}{T}\sum_{t=1}^{T} r_t P(Y = 1|\mathbf{x}_t) + (1 - r_t)[1 - P(Y = 1|\mathbf{x}_t)]. \tag{21}$$

The maximum performance in the Weather Prediction Task is .83, which is only obtained by giving the optimal response on each trial. The mean expected performance was .81 ($SD = .01$) for the Bayesian model, .80 ($SD = .02$) for the associative model with decreasing learning rate, and .81 ($SD = .01$) for the associative model with constant learning rate. These values are all rather close to the maximum performance, and while the difference between the models is not large, a repeated measures ANOVA shows it is significant, $F(2, 1998) = 326.56$, $p < .001$. Paired-sample t-tests show that the Bayesian model performs significantly better than the associative model with decreasing or constant learning rate, $t(999) = 21.65$, $p < .001$, and $t(999) = 21.37$, $p < .001$, respectively, while the associative model with constant learning rate performed better than the one with decreasing learning rate, $t(999) = 13.16$, $p < .001$. A maximising Bayesian thus outperformed a maximising associative learner in this environment. It is somewhat surprising that the constant learning rate outperformed the decreasing learning rate in the associative model. This is due to the fact that, in a few cases, relatively improbable cue-criterion pairs at the start of the task pushed the model to a suboptimal solution. Due to the decreasing learning rate, it then takes relatively long to move to the optimal solution. Thus, although convergence to the optimal solution is guaranteed, and the optimal learning rate should result in an optimal rate of convergence, actual convergence can be very slow due to "adverse" initial observations. Other schemes, in which the learning rate only decreases after an initial period of constant learning, can help avoid such problems.

Model fitting

Out of the three learning strategies, Bayesian learning was best. Here, we will compare the different strategies in how well they describe participants' learning. We fitted several versions of each model to the data.

For the Bayesian learning model, we fitted versions with different prior distributions on the weights **v**. All prior distributions were multivariate Normal, $\mathbf{v} \sim N(0, \Sigma)$, with the covariance Σ a diagonal matrix with elements σ. The first model ($B.0$) was identical to the rational Bayesian model used earlier, and had an uniform prior over **v**, i.e. $\sigma \to \infty$. In the second model, $B.\sigma$, we estimated the prior standard deviation σ as a single free parameter. In the third model, $B.\sigma_i$, σ was estimated separately for each individual. In all these models, we fixed the response scaling parameter to $\lambda = 1$. We thus assumed a probability matching strategy, rather than a maximising strategy. Probability matching is often found in research with real subjects (Shanks, Tunney, & McCarthy, 2002), so this is likely to describe participants' responses better than a pure maximising strategy. Moreover, the models are more distinguishable under probability matching. However, we also fitted variants of each model in which the response scaling parameter λ_i was estimated as a free parameter for each individual. These models are referred to as before, but with a λ_i added (e.g., $B.0$ becomes $B.0.\lambda_i$, $B.\sigma$ becomes $B.\sigma.\lambda_i$, etc.). As before, the Bayesian models were estimated by discretising the parameter space. The parameters σ and λ were estimated by maximum likelihood[4].

---

[4]We maximised the likelihood of responses R_{it} following from a Bayesian agent i learning about criterion Y_{it} and responding according to the response process in Equation 20.

Table 2: Model fits: pseudo-R^2 and Bayesian Information Criterion (BIC). Best fitting models are signalled in bold.

#par	Model	pseudo-R^2	BIC	Model	pseudo-R^2	BIC	Model	pseudo-R^2	BIC
0	$B.0$	0.321	3010.51	$A.0$	0.222	3452.49	$Ac.0$	0.341	2922.98
16	$B.0.\lambda_i$	0.360	2967.14	$A.0.\lambda_i$	0.291	3273.68	$Ac.0.\lambda_i$	0.387	**2846.08**
1	$B.\sigma$	0.332	2971.92	$A.\eta$	0.261	3286.06	$Ac.\eta$	0.343	2921.72
17	$B.\sigma.\lambda_i$	0.374	**2913.68**	$A.\eta.\lambda_i$	0.295	**3262.25**	$Ac.\eta.\lambda_i$	0.388	2852.85
16	$B.\sigma_i$	0.354	2992.42	$A.\eta_i$	0.290	3275.89	$Ac.\eta_i$	0.370	2925.38
32	$B.\sigma_i.\lambda_i$	0.383	2995.35	$A.\eta_i.\lambda_i$	0.314	3300.86	$Ac.\eta_i.\lambda_i$	0.397	2933.39

We also fitted several versions of the associative learning model. As mentioned earlier, we used a version with a decreasing learning rate, referred to as A, and one with a constant learning rate, referred to as Ac. We fitted the rational versions described earlier, with learning rate $\eta_t = 5.66/t$ for model $A.0$ and $\eta_t = .32$ for model $Ac.0$. We also fitted models, $A.\eta$ and $Ac.\eta$, in which a single learning rate parameter was estimated as a free parameter, as well as models $A.\eta_i$ and $Ac.\eta_i$, in which the learning rate was estimated freely for each individual. In all these models, we assumed probability matching, but we also fitted versions in which the consistency parameter λ_i was estimated for each participant. As before, parameters η and λ were estimated by maximum likelihood.

The results of the model fitting are given in Table 2. This table contains the values of the McFadden pseudo-R^2 (e.g., Dobson, 2002) and the Bayesian Information Criterion (BIC, Schwarz, 1978). The pseudo-R^2 represents the the proportional improvement in the log-likelihood of the fitted model over a minimal model, and values between .20 and .40 are usually taken to indicate good fit. The pseudo-R^2 does not take the number of estimated parameters into account. Since the versions of the models differ in this respect, the BIC provides a better criterion for comparison. The pseudo-R^2 of all models is in the acceptable range. However, the associative model with decreasing learning rate is clearly outperformed by the other two models. When we compare the Bayesian learning models with the BIC, we see that $B.\sigma.\lambda_i$ fitted best. The estimated parameter $\sigma = .77$ indicates that individuals learned slower than a Bayesian learner with a uniform prior. The mean response consistency λ_i was 1.41, with a standard deviation of .65. For the associative model with decreasing learning rate, the best fitting model was $A.\eta.\lambda_i$. The estimated parameter $\eta = 3.41$ indicates slower learning than the optimal value (5.66). The mean of λ_i was 1.07, with a standard deviation of 0.47. For the associative model with a constant learning rate, the best fitting model was $Ac.\lambda_i$. Hence, there was no strong evidence that participants' learned at a different rate than optimal (indeed, the estimated learning rate in model $Ac.\eta.\lambda_i$ was $\eta = .28$, which is only slightly below the optimal $\eta = .32$). The mean of λ_i was 1.39, with a standard deviation of 0.70.

Overall, model $Ac.0.\lambda_i$ fitted the data best. This corroborates an earlier finding in an experiment with amnesic and control groups (Speekenbrink et al., 2007), where the associative model with a constant learning rate also fitted best (although there, we did not include a Bayesian model with estimated prior variance). The model fits are not informative as to the reason for this better fit. Since inference $\mathbf{w}_t$ converges to $\mathbf{v}$ for both the Bayesian model and associative model with decreasing learning rate, but not the associative model with a constant learning rate, the better fit of the latter could be due to this difference. To get more insight into

the reasons for relative misfit, we compare the predicted cue utilization to the unconstrained estimates of cue utilization.

Cue utilization

The dynamic generalized linear model was used to estimate the utilization weights. The parameters $\mathbf{u}_t$ in this model were estimated by MCMC analysis, using the WinBUGS software (Spiegelhalter, Thomas, Best, & Lunn, 2003)[5].

The estimated cue utilization profiles are depicted in Figure 5. As can be seen, most individuals adapted their utilization rather quickly to the environment. Moreover, the weights diverge in the expected direction, and weights for strongly predictive cards (1 and 4) are more extreme than those for weakly predictive cards (2 and 3). When we compare the cue utilization weights to the cue validity weights (-2.10, -.58, .58, and 2.10 for the four cues respectively), we see that for most participants, utilization weights are more extreme than cue validity, reflecting a tendency towards a maximising strategy.

To gain some more insight into the validity of the different learning models, we can compare the estimated cue utilization weights to those predicted by the learning models. In Figure 6, we have done so for the best (1), closest to average (10) and worst performing (7) individuals. For the best and average individuals, the utilization profiles predicted by best fitting models $B.\sigma.\lambda_i$ and $Ac.0.\lambda_i$ are quite close to each other, and reasonably close to the estimated utilization. The profiles predicted by $A.\eta.\lambda_i$ are further removed. For the worst performing individual, all predicted profiles are very close to each other. This is mainly due to the low consistency parameter λ_i, which was estimated around .5 in all three models. This individual hardly appeared to use cards 3 and 4, and, while predictive of rain, used card 2 as predictive of sun (cue validity is negative, but utilization positive).

In Figure 7, we averaged the estimated and predicted utilization profiles over participants. We also included the average rational inference $\overline{\mathbf{w}}_t$, as derived from models $B.0$, $A.0$ and $Ac.0$. Comparing the estimated utilization to the rational inference, we see that at the end of the task, utilization weights were more extreme than rational inference, again indicating a tendency towards maximising (i.e., consistency $\lambda > 1$).

Interestingly, the model-predicted utilization was closer to the rationally inferred validity than to estimated utilization. While the consistency parameter allows predicted utilization to be more extreme than the inferred validity, and hence closer to the estimated utilization weights, on average, the estimated consistency parameter failed to match the two. As such, the models predict responses to be more variable than estimated. There are a number of possible reasons for this. First, we should note that due to the logit link function, differences on the higher ends of the scale affect the probability of responses less than differences nearer the middle of the scale (i.e., nearer 0). For example, the difference between $u = 1$ and $u = 2$ corresponds to a difference in probability of .15, while a difference between $u = 8$ and $u = 7$ corresponds to a difference in probability of 5.76×10^{-4}. As such, matching observed to predicted utilization at the start of the task is more important than matching them at the end of the task. That said, an interesting possibility is that consistency was not constant, as assumed, but increased over time. Comparison

---

[5]As a prior distribution for each cue utilization weight u_j, we used an (independent) zero-mean Normal distribution with a variance of 1. As a prior distribution for $\mathbf{\Sigma}$, the covariance matrix of $\mathbf{e}_t$, we used an inverse Wishart distribution with 16 degrees of freedom and covariance matrix $\mathbf{V} = \mathrm{diag}(1, 1, 1, 1)$. For all estimations, we ran three parallel Markov chains, with a burn-in period of 4000 and a further 20,000 draws for the actual estimation.
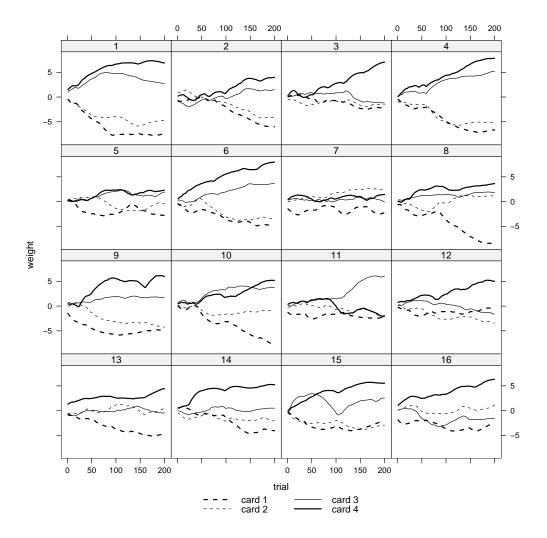
Figure 5. Cue utilization weights for each of the 16 participants. Strongly predictive cards (1 and 4) have thick lines, weakly predictive cards (2 and 3) thin lines. Cards predictive of sun (1 and 2) have solid lines, cards predictive of rain (3 and 4) broken lines.
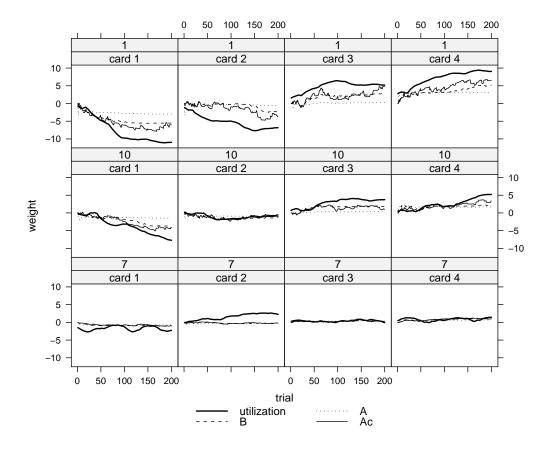
Figure 6. Estimated and predicted cue utilization weights for the best (top), closest to average (middle) and worst (bottom) performing participants. Participants' performance was computed by Equation 21. Predicted utilization weights were derived from the best fitting Bayesian model (B), best fitting associative model with decreasing learning rate (A) and best fitting associative model with constant learning rate (Ac) (in the main text, these are referred to as $B.\sigma.\lambda_i$, $A.\eta.\lambda_i$, and $Ac.0.\lambda_i$ respectively).

of the individual estimated and predicted utilization showed that, in most cases, they matched quite well in the first part of the task, but that later on the task, estimated utilization started to diverge from the predicted utilization. In those cases, the observed responses indicated a pure maximising strategy[6]. At the start of the task, this pure maximising was clearly not evident. Such possible changes in consistency deserve further attention, although their identification will probably require strong constraints on the change process.

Comparing predicted utilization to rationally inferred validity, we see that the shapes of the lines match more closely for the associative model with constant learning rate than for the other models. While rational inference by the Bayesian model and associative model with de-

---

[6]While a pure maximising strategy corresponds to infinite utilization weights, the estimated utilization weights are evidently not infinite. However, an absolute utilization weight of 11 (e.g., participant 1, cue 1, Figure 5), is practically infinite, as the predicted probability of a suboptimal response to card 1 is approximately only 1 in 100,000. The random-walk assumption, which predicts relatively smooth trial-by-trial changes in utilization, keeps the estimated utilization within bounds.
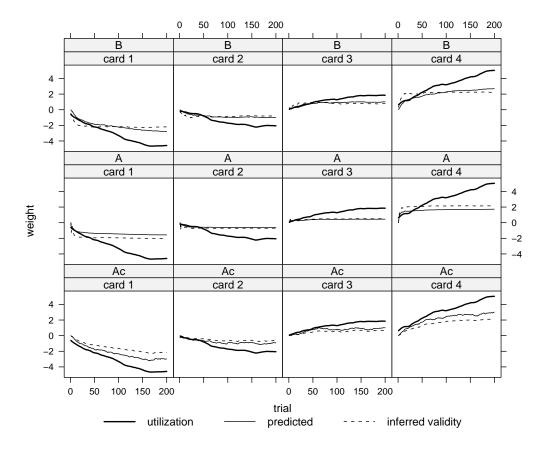
Figure 7. Mean estimated cue utilization weights (utilization), mean predicted cue utilization weights (predicted) and rationally inferred cue validity weights (inferred validity), for the Bayesian model (B), associative model with decreasing learning rate (A) and constant learning rate (Ac).

creasing learning rate show quick convergence, rational inference from the associative model with constant learning rate is more "linear" like the shape of the estimated utilization curves. By decreasing the prior variance in the Bayesian model, learning becomes slower (and more "linear"). Although the associative model with constant learning rate fitted participants' responses best, the difference with the best fitting Bayesian model was not large. This is also evident from Figure 7, which shows that the predicted utilization from both models is rather similar. As such, the apparent gradual learning could be due to participants being conservative Bayesian learners, or associative learners geared towards a non-stationary environment. At present, there is little to distinguish between these two accounts. However, environments encountered by participants outside the laboratory will often be (slowly) changing. Hence, if learning is geared towards non-stationary environments, this can be given a rational explanation.

Conclusion

In this chapter, we introduced the dynamic lens model and a general framework to study learning in multiple cue probability learning tasks. As in the original lens model, the objective

is to compare cue utilization to rational utilization and the structure of the environment. By allowing utilization and validity to change over time, the dynamic lens model offers a fine-grained overview of the learning process, allowing for a precise comparison of observed and ideal learning dynamics.

Applying the methods to the Weather Prediction Task, we showed that participants learned to predict the criterion rather well. This learning was best described by an associative learning model with a constant learning rate, and participants appeared to learn at an optimal rate for the task. Furthermore, the average response consistency showed participants responded in a better way than probability matching. However, using the dynamic lens model to obtain relatively unconstrained estimates of cue utilization, we saw that participants' responses were more indicative of a maximising strategy than predicted from the learning models. As such, the response part of these models failed to match predicted and estimated utilization. A possible reason for this discrepancy is that participants' became more consistent as the task progressed. The discrepancy was only evident from a comparison between the unconstrained estimates of utilization and those predicted from the rational learning models. This illustrates the advantages of combining the dynamic lens model and rational analysis.

The sorts of interesting question our framework raises are: what happens in a changing environment, how many cues can individuals learn about, how do results from tasks with a discrete criterion, such as the weather prediction task, generalise to tasks with a continuous criterion, such as forecasting share price? These questions will be addressed in future work.

References

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Lawrence Erlbaum.

Bishop, C. M. (1995). *Neural networks for pattern recognition.* Oxford, UK: Oxford University Press.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*(3), 193–217.

Cooksey, R. W. (1996). *Judgment analysis.* San Diego: Academic Press.

Dobson, A. J. (2002). *An introduction to generalized linear models* (2nd ed.). Boca Raton: Chapman & Hall.

Durbin, J., & Koopman, S. J. (2001). *Time series analysis by state space methods.* Oxford: Oxford University Press.

Einhorn, H. J., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment. *Psychological Review, 86*, 465–485.

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

Hammond, K. R., Wilkins, M. M., & Todd, F. J. (1966). A research paradigm for the study of interpersonal learning. *Psychological Bulletin, 65*, 221–232.

Hastie, R., & Dawes, R. M. (2001). *Rational choice in an uncertain world.* Thousand Oaks: Sage.

Kelley, H., & Friedman, D. (2002). Learning to forecast price. *Economic Inquiry, 40*, 556–573.

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning & Memory, 1*, 106–120.

Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and strategy in multiple cue learning. *Journal of Experimental Psychology: General, 135*, 162–183.

LeCun, Y., Bottou, L., Orr, G. B., & Miller, K. R. (1998). Efficient backprop. In G. B. Orr & K. R. Miller (Eds.), *Neural networks: Tricks of the trade* (pp. 9–50). Berlin: Springer-Verlag.

Luce, D. R. (1959). *Individual choice behavior.* New York: Wiley.

McCullagh, P., & Nelder, J. A. (1983). *Generalized linear models.* London: Chapman & Hall.

Oaksford, M., & Chater, N. (Eds.). (1998). *Rational models of cognition.* Oxford: Oxford University Press.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning ii: Current research and theory* (pp. 64–99). New York: Appleton-Century-Cross.

Robbins, H., & Monro, S. (1951). A stochastic approximation method. *Annals of Mathematical Statistics*, *22*, 400–407.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of statistics*, *6*, 461–464.

Shanks, D. R., Tunney, R. J., & McCarthy, J. D. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making*, *15*, 233–250.

Speekenbrink, M., Channon, S., & Shanks, D. R. (2007). Learning strategies in amnesia. *Neuroscience and Biobehavioral Reviews*, in press.

Spiegelhalter, D., Thomas, A., Best, N., & Lunn, D. (2003, January). *WinBUGS version 1.4 user manual.* Cambridge. (Available from http://www.mrc-bsu.cam.ac.uk/bugs)

West, M., & Harrison, J. (1997). *Bayesian forecasting and dynamic models* (2nd ed.). New York: Springer.