Predictiveness and Reward Effects on Attention can be Explained by a Single Mechanism

Samuel Paskewitz and Matt Jones

University of Colorado, Boulder

Abstract

In order to learn efficiently, organisms must learn how to distribute their attention to the available cues. Traditionally, most experiments on attention learning have involved discrete outcomes (e.g. no food vs. one food pellet, or category A vs. category B). A basic finding is that cues receive attention in proportion to how well they predict such outcomes. However, more recent research has shown an apparently independent effect of outcome value on attention (Le Pelley, Mitchell, & Johnson, 2013), in which cues associated with large rewards receive more attention than those associated with small rewards. It has been suggested that a separate *derived attention* mechanism - in which attention is based directly on association strength - is necessary to explain this result (Le Pelley, Mitchell, Beesley, George, & Wills, 2016). As our primary experimental contribution, we use modified versions of this design to replicate the value effect and show that it can be reversed by manipulating the rewards given for incorrect choices. Our simulations show that CompAct - a model in which cues compete for attention on the basis of their relative predictiveness - can account for both of our empirical results. The derived attention theory, in contrast, incorrectly predicts that cues associated with large rewards will always receive more attention. We conclude that we do not need separate mechanisms to account for predictiveness effects and value effects on attention.

Keywords: learning, attention, learning models, outcome value, learned attention

Predictiveness and Reward Effects on Attention can be Explained by a Single Mechanism

[1] There is a great deal of evidence that learning is shaped by selective attention. To make decisions, organisms must be able to predict outcomes on the basis of cues in the environment. Many models suppose that this is accomplished by direct learning of associations between cues or sets of cues and outcomes or action-contingent outcomes. Some theories additionally assume a second sort of learning of *attention*, i.e. which cues one ought to learn about and which to ignore. Little or nothing will be learned about cues that are ignored, while cues that are attended can rapidly develop strong associations.

As an example, consider a house cat that sometimes is scolded when it jumps up on the counter but sometimes is not. The cat dislikes being scolded and hence wants to discriminate between these cases. It notices that when its owner is in the room it is scolded, whereas when its owner is absent it gets off scot free. However, the time of day is not strongly correlated with scolding. Suppose the cat now tries to learn under what circumstances it can get away with sharpening its claws on the furniture, and suppose its first several scratching attempts happen during the day while the owner is gone, and the cat gets off scot free. If the animal pays attention to time of day, meaning it focuses on learning an association from this cue, then after the first several attempts it might acquire a belief that it is safe to scratch the couch during daylight, at which point its owner might yell at it while eating breakfast. However, if our hypothetical feline has learned from its experience with jumping on the counter to pay attention to whether its owner is present and not to time of day, then it is likely to rapidly learn that it can safely scratch away whenever the owner is not in the house.

One must be careful to distinguish attention to a cue from its associations with particular responses or outcomes. Experiments designed to examine learned attention

[1]Samuel Paskewitz and Matt Jones, Department of Psychology and Neuroscience. This research has been supported by AFOSR grant FA9550-14-1-0318. Address correspondence to mcj@colorado.edu (MJ) or sapa1760@colorado.edu (SP).

therefore often have a two-stage structure, much like the cat example given above (e.g., Mitchell, Griffiths, Seetoo, & Lovibond, 2012). In the first stage, the participant learns to use the cues to choose among a certain set of responses. In the second stage, the same cues provide information about a *new* set of responses. Theories relating attention to associability predict that experience in the first stage will cause the participant to pay more attention to some cues than to others, and that this learned attention will transfer to the second stage, affecting the rate at which each cue acquires associations with the new responses. That is, cues that gain more attention during stage 1 will acquire stronger associations during stage 2, all else being equal. The association strengths acquired in stage 2 can be assessed using test stimuli at the end of the experiment, providing an index of learned attention from stage 1. We describe such experiments below.

Although this general phenomenon is well established, there remains the question of what exactly determines which cues the learner will come to attend. That is, what statistical properties of the initial learning environment (i.e., stage 1 of the transfer design above) determine which cues receive attention, and what cognitive mechanisms govern this process? One of the major theories is that attention is based on cues' *predictiveness*. To paraphrase Sutherland and Mackintosh (1971), the predictiveness theory states that attention to a cue is strengthened when it consistently enables the organism to make correct predictions about important future events. Under this theory, attentional learning can be seen as a form of inductive reasoning: the learner assumes that cues that have been good predictors in the past will continue to be good predictors in the future (Hume, 1748/2007). The predictiveness theory of attention is not only intuitively plausible but is supported by a wide range of experimental evidence and has been embodied in a variety of mathematical learning models; these are summarized below.

Recently however, Le Pelley and colleagues have suggested that attention is also determined by the strength of cue associations (Le Pelley et al., 2016). This claim is supported by experiments manipulating the magnitude of rewards, rather then simply whether reward is present or absent. Based on these findings, they propose a new attentional mechanism whereby the association learning rate for a cue is directly proportional to the magnitude of its current associations. They call this mechanism *derived attention*, because attention weights are derived directly from associations rather than being learned via a separate mechanism. Thus this is a single-process model: there is no attention learning separate from association learning. This could also be called the "rich get richer" theory: cues that already have strong associations more easily strengthen those associations or acquire new ones.

This paper presents new experimental results and simulations that challenge the derived attention theory while supporting a particular competitive version of the predictiveness theory. We conclude that these competitive predictiveness models provide a parsimonious explanation for many attentional learning phenomena, including those based on outcome value, and thus that it is not necessary to have separate mechanisms for attention learning based on predictiveness and based on reward magnitude.

Attentional Effects With Discrete Outcomes

Investigators of learned attention have mostly used tasks in which the participant tries to predict discrete outcomes. For example, these may be category labels (Lupyan, Rakison, & McClelland, 2007), fictitious diseases (Kruschke, 1996) or food rewards in the case of experiments with non-human animals (Sutherland & Holgate, 1966). A key feature of these designs is that the reward (whether food, money or satisfaction of finding the right answer) is all or nothing: choices are either rewarded or not rewarded, and the rewards do not vary in value. Thus one is tempted to describe "predictiveness" in terms of conditional probabilities, an approach that cannot easily be extended to more general cases. The following subsections review the main empirical phenomena and modeling approaches for discrete-outcome designs.

Correlation Effects. Learners pay more attention to cues that are correlated with outcomes than to cues that are not correlated with outcomes. We call this a *correlation effect*. Correlation effects have been demonstrated in both human (Le Pelley & McLaren, 2003; Mitchell et al., 2012) and non-human animals (Sutherland & Holgate, 1966). We shall describe a single representative experiment (Le Pelley & McLaren, 2003, summarized in Table 1) that has a very similar structure to others that we discuss later. The basic idea is that participants learn to pay attention to cues that are correlated with outcomes and that this transfers to a new learning task with separate response options.

On each trial in this representative experiment, the participant is shown a stimulus consisting of two discrete cues, such as pictures or words. The objective is to predict what category label or response is appropriate for the current stimulus (i.e., cue configuration). After making a response, the participant is told what the correct answer was and whether he or she was correct. This is thus a fully supervised learning task. Stimuli are shown in a random order, and training is continued either for a fixed number of trials or until the participant's performance has passed some criterion (e.g., giving the correct response to each stimulus twice, with no intervening errors).

In the first stage (see Table 1), participants learn to predict categories using four cues that are correlated with the outcomes (A, B, C and D) and four non-correlated cues (V, W, X and Y). In the second stage, each of the previously correlated cues is paired with a previously non-correlated cue with which it had not previously appeared. Note that all cues are equally correlated with the new outcomes (III and IV). New cues (E-L) are also introduced in this stage to make the task more difficult and for comparison. In the final stage, the participants' associations between the cues and the stage 2 outcomes are tested.

The primary result is that in stage 2, participants form strong associations with cues that were correlated with the stage 1 outcomes (A, B, C and D) but only weak associations with those that were not correlated with stage 1 outcomes (V, W, X and Y) (compare rows 1-2 to rows 3-4 of Table 1, rightmost column). All of these cues were equally correlated with the stage 2 outcomes, so this difference must be the result of *attentional* learning in stage 1, which leads to faster learning for the more strongly attended cues in stage 2.

The correlation effect is the cornerstone of the predictiveness theory of attention: cues receive attention to the degree to which they predict outcomes (Mackintosh, 1975; Sutherland & Mackintosh, 1971). This theory is certainly plausible, but needs to be explicated before it can be rigorously tested. How does an organism determine to what degree a cue is "predictive"? We cannot rely merely on our intuitions about cue predictiveness: we need mathematical models to formalize these ideas and develop testable predictions.

Modeling Attention Effects in Learning. We can make our thinking about attention more precise by translating our theories into mathematical models. The models that we consider are all based on a form of the Rescorla-Wagner model (Rescorla & Wagner, 1972), which we briefly describe next. We then describe how attention can modify this basic model. Finally, we describe theories of how attention itself could be learned by predictiveness calculations or alternatively as a direct function of reward associations.

We follow the standard approach of starting with the Rescorla-Wagner model (Rescorla & Wagner, 1972) and adding a selective attention mechanism. The Rescorla-Wagner model operates in the following manner. At the beginning of each trial, the learner is presented with a stimulus vector ($\mathbf{s}$) in which cues that are absent are represented as 0 and cues that are present are represented as 1. For example, the stimulus "light + tone" could be $\mathbf{s} = (1, 1)$ while "no light + tone" would then be $\mathbf{s} = (0, 1)$.

The stimulus vector ($\mathbf{s}$) is then combined with a matrix of association weights (W) to form a vector ($\mathbf{z}$) that predicts the value of each possible action:

$$\mathbf{z} = W\mathbf{s}, \text{ i.e. } z_k = \sum_i W_{ki} s_i \text{ for all actions } k. \tag{1}$$

Each association weight (W_{ki}) represents the amount that cue i contributes to the predicted reward for action k.

The learner then decides which action to take based on their estimated values. In most of the experiments we consider here, the participant is given full feedback about each response option. Thus the particular response rule we use does not affect learning, and some other plausible one could be used instead. To determine choice probabilities we use a softmax function (Bridle, 1990; Luce, 1959) with a parameter ξ that determines how deterministic decisions are:

$$P(\text{action } k) = \frac{e^{\xi z_k}}{\sum_j e^{\xi z_j}}. \tag{2}$$

After choosing an action, the learner is given feedback in the form of a reward vector ($\mathbf{r}$). For example, in a category learning task we might have $\mathbf{r} = (1, 0)$ if the stimulus belongs to category 1 and $\mathbf{r} = (0, 1)$ if it belongs to category 2. The actual rewards ($\mathbf{r}$) are compared to the predicted rewards ($\mathbf{z}$) to form a vector of prediction errors (δ):

$$\delta = \mathbf{r} - \mathbf{z}. \tag{3}$$

These prediction errors drive learning by modifying the association weights:

$$\Delta W = \lambda \delta \mathbf{s}^T, \text{ i.e. } \Delta W_{ki} = \lambda \delta_k s_i \text{ for all actions } k \text{ and cues } i, \tag{4}$$

where the learning rate parameter $\lambda \in (0, 1)$ determines how much new learning impacts the association weights. This rule adjusts the association weights to reduce the amount of prediction error. If a cue is absent ($s_i = 0$) then nothing is learned about it. If a cue is present ($s_i = 1$) and the reward for action k was greater than expected ($r_k > z_k \Leftrightarrow \delta_k > 0$) then the corresponding association weight (W_{ki}) increases, but if the reward was less than expected ($\delta_k < 0$) then the association weight decreases.

Now we shall describe how one can augment this model with attention. For each cue i, we assume that there is a non-negative *attention weight* (a_i). These attention weights act to magnify or reduce the effective sizes of the cues that make up the stimulus vector. In each of the above equations, we replace s_i with $a_i s_i$. Thus cues that receive more attention have a greater impact on predictions

$$z_k = \sum_i W_{ki} a_i s_i \tag{5}$$

and have faster effective learning rates

$$\Delta W_{ki} = \lambda \delta_k a_i s_i = (a_i \lambda) \delta_k s_i. \tag{6}$$

In some models, attention affects only the learning rate ("associability") of cues (Equation 6), but not prediction (Equation 5) (e.g., Mackintosh, 1975). It is hopefully clear why such models would produce similar results. Indeed, in many cases a model in which attention affects only learning can be easily "translated" (via reparameterization)[2] into an equivalent model in which attention affects both learning and prediction (see the Appendix for one example). We prefer to model attentional effects on both prediction and learning, because then Equation 6 follows from Equation 5 by gradient descent on squared error. That is, the effect of attention on learning rate can be seen as a rational consequence of its effect on prediction.

Attention has been viewed as a limited resource by some researchers (Kruschke, 2001), but not others (Mackintosh, 1975). The modeling framework we have just described can accommodate both views. So far we have not specified how the attention weights (a_i) interact with each other or change with experience. Letting the attention weights be independent models the hypothesis of unconstrained attention, whereas forcing them to compete models the hypothesis of limited attention (e.g., Kruschke, 2001). In the next two subsections, we shall present two alternative models of the predictiveness theory that differ in exactly this respect and show that the one with limited attention can explain a broader range of data.

Modeling Attention Based on Predictiveness. We now have a formal theory of how attention affects learning. However, we still need to know where attention itself comes from. If we simply assign higher attention weights to more "predictive" cues by hand then our modified Rescorla-Wagner model will produce the correlation effects

[2]If one sets attention in the "learning plus prediction" model to be the square root of attention in the "learning only" model then the models behave in the same way. This exact equivalence holds only if attention is independent of the current stimulus vector (e.g. it does not hold for CompAct, a model described below).

described above: the correlated cues will develop larger association weights and have a greater effect on behavior. However, for a complete theory we need to develop an *attention learning rule* that somehow automatically calculates how "predictive" each cue is and uses this information to determine the cues' attention weights (a_i).

One standard approach to deriving incremental learning rules, including ones for learning attention (e.g., Kruschke, 2001), is by gradient descent on squared error. This is a mathematical procedure to gradually determine what set of attention weights yields the most accurate predictions of outcomes. After each trial, the model asks itself "how could I have adjusted attention to reduce my error?" This is the same principle used in multi-layer neural networks to obtain the backpropagation algorithm (Rumelhart, Hinton, & Williams, 1986). Thus, the "predictiveness" of a cue for a given model can be defined as the extent to which paying more attention to that cue makes the model's predictions more accurate in the current task. Rather than having attention directly determined by some objective notion of predictiveness, learning by gradient descent defines predictiveness with respect to the current knowledge state of the model, and predictiveness determines *changes* in attention.

Perhaps the simplest such learning rule arises when the attention weights (a_i) are independent of each other and of which stimuli are present on a given trial, as in the basic model of Equation 5. In that case, gradient descent gives us:

$$\begin{aligned} \Delta a_i &= -\mu \frac{\partial}{\partial a_i} \sum_k \tfrac{1}{2}\delta_k^2 \\ &= \mu \sum_k \delta_k \cdot W_{ki} s_i \end{aligned} \tag{7}$$

where $\mu \in (0, 1)$ is an attention learning rate.

With a little interpretation, we can see that this learning rule captures a sort of predictiveness. Suppose that δ_k is negative, i.e. the magnitude of outcome k was less than expected. If $W_{ki}s_i$ was also negative, then cue i foretold this outcome and hence should receive more attention, which indeed it does under Equation 7 (negative $\times$ negative =

positive). On the other hand if $W_{ki}s_i$ was positive then cue i made the prediction worse, and hence attention to i decreases (negative $\times$ positive = negative). This suggests that we re-write Equation 7 in the following equivalent form:

$$\Delta a_i = \mu \sum_k \operatorname{sign}(\delta_k W_{ki} s_i)|\delta_k||W_{ki}s_i|. \tag{8}$$

If $\operatorname{sign}(\delta_k W_{ki} s_i) = 1$ then cue i was a "good predictor" for outcome k, because its contribution to the prediction served to reduce the prediction error, and hence it receives more attention. The reverse is true when $\operatorname{sign}(\delta_k W_{ki} s_i) = -1$, i.e. when cue i was a "bad predictor" because its contribution increased prediction error. The magnitude of the increase or decrease in attention reflects the overall amount of prediction error ($|\delta_k|$), and this makes sense: large errors require large shifts in attention to correct them. The amount of attention change also depends on the size of cue i's prediction ($|W_{ki}s_i|$). We can also see that this attention learning rule (Equation 7 or 8) produces attention weights that reflect the covariance between unexpected outcomes (δ_k) and cue predictions ($W_{ki}s_i$). It is thus reasonable to interpret $\sum_k \delta_k \cdot W_{ki}s_i$ as the predictiveness of cue i on a particular trial. Note that this notion of predictiveness does not reflect any purely objective statistics about the learner's environment (e.g. conditional probabilities of outcomes given cues, or correlations between cues and outcomes), but rather the relationship between cues and the learner's own performance (i.e., prediction error).

It should be clear that this learning rule will produce the correlation effects described above. Consider any of the non-correlated cues in stage 1 (refer to Table 1), e.g. cue V. Because of the symmetry of the task, $|\delta_k||W_{kV}s_V|$ will be similar on A + V and B + V trials, but $\operatorname{sign}(\delta_k W_{kV} s_V)$ for each k will alternate between these two trial types. Thus any attention gained by cue V on one trial will be wiped out on another. In contrast, a correlated cue such as A will consistently gain attention: $\delta_k W_{kA} s_A$ has the same sign (positive, for each k) on A + W and A + V trials, leading to a consistent increase in attention. We thus see that Equation 7/8 formalizes our notion of learning attention based on “predictiveness”.

The Highlighting Effect and Competitive Predictiveness. When we add a competitive element to attention learning (i.e. assume that attention is limited), the predictiveness principle gains the ability to explain a far broader range of phenomena, beyond correlation effects. One of these is *highlighting*, a phenomenon showing that the cues present on a trial influence each other's attention learning (Kruschke, 1996, 2009). As an illustration of this phenomenon, we shall consider Experiment 2 from Kruschke (1996), which is outlined in Table 2.

This highlighting task has three types of cue: “perfect early” (PE_1 and PE_2), “perfect late” (PL_1 and PL_2) and “imperfect” (I_1 and I_2). Each of the “perfect” cues indicates a single correct response. For example, whenever the participant sees cue PE_1, the correct response is always E_1. The imperfect cues are paired with the perfect ones in such a way that either of two responses could be correct for each imperfect cue. For example, $I_2 + PL_2$ means response L_2 is correct, but $I_2 + PE_2$ indicates response E_2. The perfect early, perfect late, imperfect structure is duplicated, so we have I_1 and I_2, etc.

For the first half of the task, participants are only shown $PE_1 + I_1$ and $PE_2 + I_2$ trials (hence the “early” designation). In the second stage, the late cues ($PL_1 + I_1$ and $PL_2 + I_2$) are added to the early ones, with all trial types having equal frequency. The key finding is that when given the test stimulus $PE_1 + PL_1$, participants mostly choose L_1 (the outcome associated with PL_1) despite having received more training trials with cue PE_1 (and the same pattern is observed with regard to PE_2 and PL_2). This result suggests that PL_1/PL_2 received more attention than PE_1/PE_2, leading to the formation of stronger associations.

What would cause this difference in attention? The simple predictiveness model (Equation 7) cannot account for it: the early and late cues have the same perfectly predictive relationship with their respective responses. Do we need to abandon this model of attention learning, or can it be modified so that it explains highlighting?

We could modify our simple predictiveness model so that attention changes are determined by the predictiveness (defined in the sense of Equation 7) of the cue in question *relative to the predictiveness of other cues present at the same time.* This idea captures the intuition of attention as a limited resource over which cues must compete. In particular suppose that, if one cue is a bad predictor, then attention to its companions tends to increase, and if that cue is a good predictor then attention to its companions tends to decrease:[3]

$$a_i \text{ increases with } \sum_k \delta_k W_{ki} s_i \text{ and} \tag{9}$$

$$a_i \text{ decreases with } \sum_k \delta_k W_{kj} s_j \text{ for all } j \neq i. \tag{10}$$

Note that Equation 9 is just the simple predictiveness learning rule described above (see Equation 7). However, Equation 10 introduces a new competitive element to attention learning.

Kruschke (2001) has implemented this competitive learning rule in a model called EXIT, which reproduces highlighting and other related phenomena. However, EXIT also has a complex set of additional mechanisms, viz. exemplar-mediated attention and rapid attention shifts. We therefore prefer to present a new model called CompAct ("Competitive Activation"), which implements only the competitive attention rule. In addition to its simplicity, CompAct better aligns with the other models we consider, making the theoretical comparisons we want to do more transparent.

CompAct belongs to the class of attention-modified Rescorla-Wagner models we define above with Equations 5 and 6. In CompAct, the attention weights (a_i) depend on a set of pre-attention weights (η_i) as well as on the current stimulus (i.e., cue compound), and they act to normalize the stimulus vector:

$$a_i = \frac{\eta_i}{\sum_j |\eta_j s_j|} \Rightarrow \sum_i |a_i s_i| = 1. \tag{11}$$

Thus cues compete for activity (by which we mean $a_i s_i$), which is always normalized to 1.

[3]Recall that $s_j = 0$ for absent cues, so the only cues relevant to attention change are those present in the current trial.

The attention weights are a function of the whole stimulus, i.e. a_i depends on the values of s_j for all other cues j.

This competition between cues for activation also leads to a competitive learning rule for attention when we derive the latter using gradient descent. The learning rule for attention is based not on a directly, but rather on the pre-attention weights (η) from which a is derived. This learning rule (derived in the Appendix) is

$$\begin{aligned}\Delta\eta_i &= -\mu\frac{\partial}{\partial\eta_i}\sum_k \tfrac{1}{2}\delta_k^2 \\ &= \mu\frac{1}{\sum_j |\eta_j s_j|}\sum_k\left[(1-|s_i|a_i)\delta_k W_{ki}s_i - |s_i|\sum_{j\neq i} a_j(\delta_k W_{kj}s_j)\right],\end{aligned} \tag{12}$$

Because of the way a is defined, $1-|s_i|a_i \geq 0$ and $|s_i|a_j \geq 0$. Thus this learning rule embodies the principle of competitive predictiveness expressed in Equations 9 and 10.

If there are only two cues and $s_1 = s_2 = 1$, then $1 - a_1 = a_2$, so Equation (12) reduces to

$$\Delta\eta_1 = \mu\frac{a_2}{\sum_j |\eta_j|}\sum_k \delta_k(W_{k1} - W_{k2}). \tag{13}$$

Thus, attention updating is proportional to the difference between the two cues' association weights. The contribution of each outcome k is to shift attention toward the cue with the greater association weight if prediction error (δ_k) is positive, and vice versa if prediction error is negative.

The competitive attention learning rule embodied in CompAct (Equation 12) can explain highlighting. (What follows is essentially a simplified version of the explanation in Kruschke 2001 using the EXIT model.) Looking at plots of simulated prediction errors and association weights helps us to see this more clearly (Figure 1a). In the first half of the task (trials $1-75$), the I cues form strong associations with the "early" outcomes, i.e. $I_1 \rightarrow E_1$ and $I_2 \rightarrow E_2$ (Figure 1a, top left panel). When I_1 and I_2 are paired (respectively) with PL_1 and PL_2 in the second half of the task, the learner thus incorrectly predicts that E_1 or E_2 is the correct response (see bottom right panel). The difference between the PL cue association (blue line) and I cue association (maroon line) is large and negative, while prediction error (black line) is large and negative. We thus see that attention to the PL cue increases (see Equation 13), while attention to the I cue decreases. Using Kruschke's terminology, we can say that the incorrect $I_1 \rightarrow E_1$ and $I_2 \rightarrow E_2$ associations *highlight* cues PL_1 and PL_2. Attention to PL_1 and PL_2 is thus higher than attention to PE_1 and PE_2 (Figure 1b), and this makes predictive strength ($W_{ki}a_i s_i$) greater for the PL cues (c.f. Equation 5). Consequently, the model chooses the responses associated with the PL cues when given conflicting cue pairs at the end of the experiment.

This principle of competitive predictiveness embodied in CompAct and EXIT can explain not only correlation and highlighting effects, but many other learning phenomena. These include learned inattention after blocking (Kruschke, 2001) and even so-called "Pearce-Hall" effects, in which what seem to be the *less* predictive cues receive more attention (Haselgrove, Esber, Pearce, and Jones 2010; the simulation result will be demonstrated in a forthcoming paper). The competitive predictiveness principle used by EXIT and CompAct is thus a simple and parsimonious explanation for many attentional learning phenomena in the domain of discrete outcomes.

Value Effects on Attention

The "Alien Slot Machine" Effect. The theory of competitive predictiveness described above seems to be sufficient for explaining attentional effects in tasks with discrete outcomes. However, there is an important class of value-driven effects on attention that some researchers have claimed to be inexplicable by any predictiveness mechanism (Le Pelley et al., 2016, 2013). In visual search tasks, cues associated with large rewards capture visual attention (Anderson, Laurent, & Yantis, 2011), even when this impairs performance (Le Pelley, Pearson, Griffiths, & Beesley, 2015). Le Pelley et al. (2013) have demonstrated such a value effect in a learning setting, using a modified classification task similar to those described above. In brief, cues associated with large rewards received more attention (in stage 2) than did those associated with small rewards, even though all of these cues were equally good predictors of the correct response. Their "alien slot machine" task worked as follows.

The goal of a classification task is usually expressed in terms of maximizing accuracy rather than obtaining the most possible reward. However, a two-category classification task can also be described as a reward learning task with two actions (viz. "classify as Category A" and "classify as Category B") with the participant receiving a reward of 1 unit (or any other positive value) for correct classification decisions and a reward of 0 units (or some other fixed lesser value, e.g. -1) for incorrect classification decisions. The outcome coding used in Table 1 illustrates how this can be done for each stage of the correlation effect experiment described above.

Le Pelley et al. (2013) modified this design by making one action yield more reward (when correct) than the other (see Table 3). Participants were told that they were playing an alien slot machine that paid out in two different currencies, one of which was worth 150 times as much as the other (represented as actions I and II, respectively). The machine paid out only if the participant guessed the correct currency.

After stage 2 training with two new currencies (which had unknown and hence from the participants' perspective equal values), Le Pelley and colleagues found that participants' responses were more influenced by the cues that had been associated with high value in stage 1 (A and B) than those that had been associated with low value (C and D). Thus, even though all these four cues were perfect predictors for the category labels (currency names) in stage 1 (as well as the amount of reward), it can be inferred that the cues associated with the more valuable outcomes received more attention. This conclusion was supported by eye-tracking data: participants spent more time looking at cues A and B than at C and D during stage 2 training.

Derived Attention Can Explain Value and Correlation Effects, but Not Highlighting. Le Pelley and colleagues have a simple explanation for the alien slot machine effect, viz. that attention is derived directly from current cue associations (Le Pelley et al., 2016). This *derived attention* model is based on the same extended Rescorla-Wagner framework as above (Equations 1, 3, 4 and 6) and makes a very simple proposal for the determinants of attention. On each trial, the attention vector is obtained via the current association weights. The simplest form of this is to make attention equal to the total magnitude of all the cue's association weights (i.e., summed over all responses). We set a lower bound ($a_{\min}$) to ensure that each cue receives at least some attention, because otherwise the model can never learn about a cue that starts with null association weights. Thus the derived attention theory can be embodied in the following formula for attention weights:[4]

$$a_i = \max\left\{a_{\min}, \sum_k |W_{ki}|\right\} \tag{14}$$

Under this model, a cue will tend to receive more attention if it predicts a large outcome. This explains the attentional capture effects described above (if we assume subjects learn cue-reward associations in visual search tasks and that a also controls visual attention), as well as the alien slots effect. Derived attention also predicts correlation effects: cues that do not consistently predict outcomes will acquire small association weights and hence receive little attention. However, derived attention does not explain highlighting: the E cues in that design will have stronger associations and receive more attention than the L cues (simply by virtue of receiving more training trials), leading to the opposite of the empirical result (see Figure 2).

[4]In the original formulation, attention affects only learning and not prediction (Le Pelley et al., 2016). In order to more easily compare the derived attention model to other models, we prefer to simulate a version in which attention also affects prediction (as in Equation 5). We show in appendix C that it actually does not matter: both versions are equivalent.

Summary: Do We Need Different Mechanisms to Explain Different Phenomena?

Based on the findings reviewed thus far, it appears that we have two different attention mechanisms that each explain different learning phenomena. Both competitive predictiveness and derived attention can explain correlation effects, which are the most basic phenomena of learned attention. However, it is not obvious how a cue that is perfectly correlated with a large outcome is more predictive than a cue that is also perfectly correlated with a small outcome, so how can any sort of predictiveness theory explain value effects such as those in the alien slot machine experiment? On the other hand, derived attention can easily explain these value effects, but it cannot account for highlighting. Le Pelley et al. (2016) thus suggest that a hybrid model with both types of attentional mechanism may be the best solution.

We, however, shall show that it is sufficient to assume only one mechanism, viz. competitive predictiveness. We first describe experiments (1-a and 1-b) that replicate the basic alien slot machine phenomenon under some minor modifications of the task. Next, in Simulation 1 we show that this resulted is predicted by CompAct and the simple predictiveness model, as well as the derived attention model. We then examine a new reward structure in which the difference in reward between correct and incorrect responses is greater for low value cues than for high value cues. In Simulation 2, we show that the derived attention model and CompAct make contrary predictions with regard to this new experimental design. The derived attention model (as well as the simple predictiveness model) predicts that participants will still attend more to the high value cues, but CompAct predicts that mechanisms analogous to those used by Kruschke to explain highlighting (Kruschke, 2009) will draw their attention to the *low value* cues instead. We confirm CompAct's seemingly paradoxical prediction in Experiments 2-a and 2-b. Our conclusion is that derived attention is neither necessary nor sufficient to explain value effects on attention, whereas competitive predictiveness is sufficient to explain both these and many other phenomena.

Experiment 1-a

Experiment 1-a sought to replicate the alien slot machine effect of Le Pelley et al. (2013), using a modified task design that enabled us to separately manipulate rewards associated with correct and incorrect responses. This affordance becomes critical in Experiments 2-a and 2-b below. To do this, we modified the cover story so that, instead of predicting which of two currencies the slot machine would pay on each trial, the subject's task was to select among two alternative buttons on the machine (see Figure 3). Each button held a different reward, both of which were revealed after the subject made a selection. This modification also enabled rewards to be balanced between the two responses, as explained below.

Table 4 shows the design of the experiment. The key features of Le Pelley et al.'s (2013) design are retained: In stage 1, two of the relevant cues (A and B) predict large rewards and two (C and D) predict small rewards. In stage 2, these cues are equally predictive of rewards of equal magnitude. Thus the critical test trials (A + C and B + D) provide a test of how the reward structure in stage 1 affected attention during stage 2. In particular, if subjects learn to attend to high value cues during stage 1, then on these test trials they should select the responses associated with those cues in stage 2.

Method

Participants. Participants were recruited using Amazon Mechanical Turk ($n = 30$). They were compensated with a base payment of \$1.67 plus a bonus payment of up to \$2.50 based on performance. The task required between 10 and 30 minutes to complete for most participants.

Apparatus and Materials. The experiment was written in Javascript and managed using PsiTurk (https://psiturk.org/). Data analysis was performed using R with the *ggplot2* package for making graphs (https://cran.r-project.org/web/packages/ggplot2).

Design. We modified the design of Le Pelley et al. (2013) in several ways, as shown in table 4. First, we arranged payoffs such that the two high value cues (A and B) were not associated with the same action. Instead, one high value cue (A) and one low value cue (C) indicated that action I gave the best payout, while cues B (high value) and D (low value) were associated with action II. The purpose of this arrangement was to equalize the average value of each action, which simplifies the theoretical analysis and prevents unmotivated participants from merely choosing a single action on all trials (as they might if one action had a much greater average payoff than the other). To do this, we discarded the part of the cover story from Le Pelley et al. (2013) in which the responses were identified with alien currencies of different values, and instead asked subjects to select between two buttons displayed on the slot machine. This also freed us to change the reward for incorrect responses to 1 instead of 0 in order to encourage participants to process the value of each outcome and not merely whether they were "right" or "wrong". We also changed the other payoffs: participants received 100 points for a correct answer on high value trials (indicated by cue A or B) and 2 points for a correct answer on low value trials (cue C or D). We thus refer to this reward structure as a 100/1 vs. 2/1 design, whereas experiment 1 of Le Pelley et al. (2013) (Table 3) can be called a 150/0 vs. 1/0 design.

An unfortunate result of this new reward structure was that the "irrelevant" cues (V, W, X and Y) conveyed information about which choice was better on any given trial. This made the design less comparable to Le Pelley et al.'s (2013) original, in which these cues were completely uninformative. Our design still featured cues associated with large rewards (A and B) compared to cues associated with small rewards (C and D), the central feature of the original alien slots design. Nevertheless, we refined our design in Experiment 1-b (see below).

Procedure. Participants were given instructional text before each stage. They were told to imagine themselves as space explorers playing a series of "alien slot machines" for "space credits" that would be converted into real world cash at the end of the experiment.

On each trial, the participant saw a schematic slot machine, with a pair of cues and text prompting him or her to choose from two buttons corresponding to two keys on the computer keyboard (see Figure 3, left). After making a response, the participant was told how much money he or she won for that response and how much money he or she would have won for making the other response, in slightly faded text (see Figure 3, right). These reward amounts were presented both numerically and by a picture of that number of coins.

Each cue was a picture of an "alien fruit" labeled with a nonsense word. The fruits and their names were designed to be as distinct from each other as possible, particularly with regard to fruit color (pilot experiments showed that some participants tended to group purple fruits together, for example). Fruit pictures and names were randomly assigned to act as cue A, cue B etc., separately for each participant. Each trial type within a stage occurred once, in random order, before being repeated. The different slot machine in each stage was represented by a slightly different background color.

In order to ensure that we analyzed data only from people who learned the task, we added performance criteria for the various stages and a very simple introductory stage. This tutorial stage is shown in Table 5 and was the same for all experiments in this report. After completing the tutorial, participants moved on to the main experiment. In order to continue past the tutorial, stage 1 and stage 2, participants needed to reach a criterion level of performance (respectively 8, 8 and 10 consecutive correct choices). We thus only had test trial data from participants who learned the correct response contingencies in each stage of the experiment.

In the final (test) stage, participants were told that they would be tested on the machine from stage 2, winning a base prize of 500 Space Credits for each correct answer. For some of the cue pairs ("old concordant" and "new concordant"), both cues had been paired with the same correct response. These were used as a check that participants still remembered stage 2 associations and were properly following instructions.

The critical test stimuli were the "discordant" pairs A + C and B + D, for which the stage 2 reward structure did not clearly indicate a correct response. In these pairs, one cue had predicted high value and the other low value during stage 1, and the cues were associated with opposing responses during stage 2. Participants who paid equal attention to the high value cues (A and B) and low value cues (C and D) should not favor one response over the other. However, a participant who attended more to the high value cues would form stronger A $\rightarrow$ III and B $\rightarrow$ IV associations than C $\rightarrow$ IV or D $\rightarrow$ III associations. This would be reflected in their responses to the discordant stimuli: they would tend to choose response III when shown A + C and response IV when shown B + D (as indicated in the "predicted" column of Table 4). Participants who paid more attention to the low value cues would show the opposite pattern. We thus used responses on the discordant test stimuli as an indicator of how participants distributed their attention.

In addition to recording participants' choices in the test stage, we collected a confidence rating for each choice. We tried to make participants feel more invested in their confidence ratings by giving them a monetary value. This procedure was explained in a short tutorial prior to the test stage. After meeting the performance criterion in stage 2, participants were told they could now "hedge their bets" after choosing an action by reserving a portion of their winnings to be paid in case they had made the wrong choice. For example, if the participant chose to reserve 30%, then the payoff for that trial would be 70% of the reward for the chosen action and 30% of the reward for the non-chosen action (see Figure 4 for screenshots). They could reserve 0, 10, 20, 30, 40 or 50 percent of their winnings this way, thereby yielding six confidence levels, with a choice to reserve less money interpreted as indicating greater confidence. Participants completed eight further trials of stage 2 using this confidence rating system before using it in the test stage.

The test stage was run for two blocks, with each stimulus presented once per block in random order. Participants were told that they had entered a "special bonus round" in which they would be tested on their knowledge of the machine. No feedback was given during this stage, to prevent further learning. After completing the entire experiment, each participant was told how much he or she had won and asked to rate their engagement in the task and its difficulty on scales from 0 to 10.

In addition to the base pay, each participant was paid a bonus equal to the proportion of possible space credits won times a multiplier based on the stage of the task reached (\$0.25, \$0.50, \$0.75, or \$2.50). For example, a participant who completed the entire experiment and won 75% of all possible space credits would receive a bonus of $\$2.50 \times 0.75 \approx \$1.88$. The payments were based on the *proportion* of space credits won (total divided by maximum possible) in order to avoid punishing participants who learned the task quickly.

Analysis. Responses on concordant test trials were coded as $+1$ (correct) or -1 (incorrect). Responses on discordant test trials were coded as $+1$ (consistent with greater attention to A and B) or -1 (consistent with greater attention to C and D). Confidence was coded on a scale from 1 to 6, with 6 being the most confident (i.e., reserving 0% of winnings). The response and confidence values were multiplied to produce a response-confidence score on each trial. Scores from each discordant trial type (A + C and B +D) were averaged, and these averages were added together to form an aggregate measure of attentional focus, ranging from -12 to 12. A positive total score for discordant stimuli indicated greater attention to high value cues (A and B), whereas a negative score indicated greater attention to low value cues (C and D). Scores from old concordant trials, new concordant trials with high value cues, and new concordant trials with low value cues were treated similarly, to obtain measures of adherence to the task (see Table 4 for an explanation of these trial types).

Results and Discussion

Out of 30 participants, 4 failed to meet the learning criterion in stage 0, 14 passed stage 0 but failed stage 1, and 12 passed all stages. The results presented here are from those 12 participants. The discordant scores were greater than zero on average, indicating greater attention to the high value cues (A and B) than to the low value cues (C and D) ($M = 5.04$, $SD = 7.26$, $d = 0.69$, $t_{11} = 2.40$, $p < .05$). Scores for the concordant test stimuli were also positive, confirming that the participants applied what they had learned in stage 2 to the test stage (old concordant: $M = 9.79$, $SD = 4.27$, $d = 2.29$, $t_{11} = 7.94$, $p < .0001$; new A/B concordant: $M = 7.50$, $SD = 6.07$, $d = 1.24$, $t_{11} = 4.28$, $p < .01$; new C/D concordant: $M = 5.54$, $SD = 5.02$, $d = 1.10$, $t_{11} = 3.82$, $p < .01$). Choice proportions in the test stage are given in Table 8.

In conclusion, this experiment successfully replicated Le Pelley et al.'s (2013) alien slot machine effect, using somewhat different methods and rewards and a different participant population. Aside from independently confirming that result, Experiment 1-a shows that our modified methods are appropriate for investigating outcome value effects on learning.

Experiment 1-b

Experiment 1-b was a nearly exact replication of Experiment 1-a, with one modification to the irrelevant cues in stage 1 so that they would not carry any information about the correct response or rewards (see Table 6).

Method

Participants. Participants were recruited using Amazon Mechanical Turk ($n = 60$). They were compensated with a base payment of \$2.00 plus a bonus payment of up to \$2.50 based on performance. The task required between 10 and 30 minutes to complete for most participants.

Apparatus and Materials. These were identical to those used in Experiment 1-a.

Design. This was a revised version of the 100/1 vs. 2/1 design used in experiment 1-a with half as many "irrelevant" cues. This ensures that those cues (X and Y) truly are irrelevant in that they do not indicate which action will result in a better payout. The design is outlined in table 6.

Procedure. Aside from the altered design described above, the methods were identical to those of experiment 1-a.

Results and Discussion

Out of 60 participants, 10 successfully completed the entire task (17 failed at stage 0, 25 at stage 1 and 6 at stage 2). Discordant scores were positive on average, indicating greater attention to the high value cues ($M = 4.50$, $SD = 6.00$, $d = 0.75$, $t_9 = 2.49$, $p = .032$). As before, scores for the concordant test stimuli were also positive (old concordant: $M = 10.15$, $SD = 3.96$, $d = 2.56$, $t_9 = 8.11$, $p < .0001$; new A/B concordant: $M = 7.75$, $SD = 3.57$, $d = 2.17$, $t_9 = 6.87$, $p < .0001$; new C/D concordant: $M = 7.35$, $SD = 5.21$, $d = 1.41$, $t_9 = 4.46$, $p = .002$), indicating good adherence to the task. Choice proportions in the test stage are given in table 8.

We thus replicated the value effect obtained in Experiment 1-a, giving us more confidence that this is a real effect (especially given the small sample sizes). Both Experiments 1-a and 1-b can be viewed as conceptual replications of Le Pelley et al.'s (2013) finding that when the reward for one correct response is greater than that for the other, and rewards for incorrect choices are small and equal, cues associated with high values receive more attention.

Simulation 1

Having replicated the alien slot machine effect (Le Pelley et al., 2013), we wished to determine whether a derived attention mechanism was necessary to explain it. We therefore simulated the simple predictiveness model and CompAct, along with the derived attention model, on a wide array of parameter values to determine whether they were capable of behaving in the same way as real participants in Le Pelley et al. (2013) as well as our Experiment 1-b. This exhaustive approach was practical because each of the models has only one or two parameters relevant to learning. The choice consistency parameter ξ

does not affect learning when full reward feedback is given (as it is in these tasks), and merely scales choice probabilities.

Method

Simulations were performed in R (https://cran.r-project.org). The models were defined by the equations given in the introduction: Equations 2-7 for simple predictiveness, Equations 2-6 and 11-12 for CompAct, and Equations 2-6 and 14 for derived attention. Because ξ merely scales choice probabilities, we fixed it at 2 for both models. The two other parameters of the simple predictiveness model and CompAct (λ and μ) are by definition restricted to the range $[0, 1]$: we divided the interval from 0.05 to 0.95 evenly by intervals of 0.05, giving us a grid of 361 points. We used this same range of λ values for the derived attention model and let $a_{\text{min}} = 0.1$. The initial attention weights for the simple predictiveness and derived attention models were set at $a_i = 0.1$, and CompAct was initiated with $\eta_i = 1$.

We generated 10 random trial sequences for each experiment design, with 122 trials in stage 1 and 24 trials in stage 2, and used these for all three models. The number of trials in each stage matched Le Pelley et al.'s (2013) design, and was similar to what participants experienced on average in our experiments. In both the simple predictiveness and derived attention models, association weights can grow so large that they present computational difficulties and cause model predictions to change dramatically from one run to another. To ameliorate this tendency, we divided reward values by 100 in our simulations and capped attention weights at a maximum value of 1.5. The unknown reward values in stage 2 of Le Pelley et al.'s (2013) experiments were simulated as 1 (i.e. 2/3 of the higher valued currency in stage 1).

To determine whether each model qualitatively behaves like human participants, we examined choice probabilities on A + C and B + D test trials. We obtained an index of attentional preference for the high value cues (A and B, positive values) over the low value cues (C and D, negative values) using the following formula:

$$\text{choice index} = P(III|A+C) + P(IV|B+D) - P(IV|A+C) - P(III|B+D) \tag{15}$$

where $P(III|A+C)$ is the probability of choosing action III given stimulus A + C, etc. These probabilities were taken directly from the softmax choice function (Equation 2). This choice index is the same as the aggregate measure of attentional focus used to analyze the human data, except that the simulated data are not multiplied by confidence ratings (see below). For each model and experiment design, we report the highest and lowest choice index (averaged across all 10 runs) among all tested parameter values, along with the accompanying parameter values and the attention weights at the end of stage 1 (a_i for the simple predictiveness and derived attention models, and η_i for CompAct). We can thus see whether the models are highly flexible—able to fit various patterns of data—or make firm predictions that do not depend on particular parameter values.

Results

When simulated on the original alien slot machine design (Le Pelley et al., 2013), all three models behaved in the same fashion as human learners, paying more attention to the high value cues (A and B), and exhibiting a positive choice index under all parameter values. Details are given in Table 9.

The same result was obtained when the models were simulated on our Experiment 1-b. All three models preferred the actions associated with the high value cues (A and B), under all parameter values, just as the human learners did (see results reported below). Again this was explained by the models' greater attention to the high value cues. Details are presented in Table 10.

For both experimental designs, each of the models behaved qualitatively in the same way as human participants (paying more attention to cues A and B than to C and D) across the entire range of parameter values. Thus Le Pelley et al.'s (2013)'s alien slot machine effect is a definite prediction of all three models.

Discussion

Le Pelley and colleagues claim that a derived attention mechanism is needed to account for the alien slot machine effect (Le Pelley et al., 2016). In this simulation we show that this is not the case: both competitive (CompAct) and non-competitive predictiveness models strongly predict greater attention to the high value cues (A and B). In both models, attention to cue i increases as a function of the term $\sum_k \delta_k W_{ki} s_i$ (Equations 7 and 12). The association weights (W_{ki}) and prediction errors (δ_k) are larger on trials with the high value cues than those with the low value cues, driving larger increases in attention. In conclusion neither the original alien slots experiment (Le Pelley et al., 2013) nor our Experiments 1-a and 1-b allow us to distinguish which model is correct.

Simulation 2

As shown in Simulation 1, all three of the models which we are considering (simple predictiveness, CompAct and derived attention) predict the alien slot machine effect demonstrated by Le Pelley et al. (2013) and replicated in our Experiments 1-a and 1-b. We therefore require a new experimental design, about which the models make different predictions. Such a design is summarized in Table 7. We test the model predictions obtained in this section in Experiments 2-a and 2-b.

Our new design is summarized in Table 7. It differs from that of Experiment 1-b in that the reward for an incorrect choice on high value trials (cues A and B) is increased to 95, while the reward for correct choices on low value trials (cues C and D) is increased to 50. Thus the difference in reward between the correct and incorrect actions is small on high value trials ($100 - 95 = 5$) and large on low value trials ($50 - 1 = 49$). We call this design "100/95 vs. 50/1".

These changes to the reward structure cause our theories to make different predictions as to which cues will receive the most attention. The derived attention model predicts that the cues with larger associations in stage 1 (A and B) will receive more attention, which will carry over to affect learning in stage 2 (just as in Le Pelley et al. 2013 and our Experiments 1-a and 1-b). However, if attention is based on competitive predictiveness, we expect the larger difference between rewards for the two response options on C and D trials to yield more attention to those cues. Observe that in Experiments 1-a and 1-b (and the experiments of Le Pelley et al.) the difference between rewards is confounded with the overall magnitude of reward, making it impossible to determine which factor is driving attention.

Why would the competitive predictiveness theory predict that greater difference between rewards leads to more attention? The key is the irrelevant cues X and Y. Due to the symmetry of the experimental design, cues X and Y cannot distinguish which response is correct. That is, their prediction ($W_{ki}s_i$) will be approximately the same for both the correct and incorrect responses (see the maroon curves in the upper half of Figure 5a). On high value trials, both of the rewards are so large and close together that the prediction (δ_k) is large and positive, at least early in learning (Figure 5a, top two panels). Thus $\delta_k W_{ki} s_i$ is positive for X and Y, i.e. they are "good predictors". This causes these cues to compete for attention with the truly relevant cues (A and B), reducing the rate at which the latter gain attention. In other words, on high value trials, most of the predictive work to be done is anticipating the overall reward level for both actions, as opposed to differentiating the rewards between actions. Cues X and Y are just as suited to this job as are A and B.

During low value trials, on the other hand, the reward for an incorrect choice is only 1. Cues X and Y substantially over-predict this, causing $\delta_k W_{ki} s_i$ to be negative and hence X and Y to be "bad predictors" and lose attention (see Figure 5a, bottom right panel). This lost attention to X and Y is gained by C and D, much like in the highlighting experiment described above (c.f. Figure 1). Meanwhile, prediction error for the correct action rapidly decreases to nearly zero, so cues X and Y do not gain attention via this response, even though they have large association weights (Figure 5a, bottom left panel). In other words, most of the predictive work on low-value trials lies in differentiating the outcomes for the two actions, which X and Y cannot do but C and D can.

In conclusion, according to the competitive predictiveness theory cues X and Y "highlight" cues C and D, but they compete for attention with A and B. The simulations confirm this. These competitive dynamics are not a factor for the simple predictiveness or derived attention models, which predict (in contrast to the competitive predictiveness theory) that the high value cues (A and B) will receive more attention than the low value cues (C and D).

Method

Methods were identical to those in Simulation 1, except that the task design was taken from Table 7.

Results and Discussion

The simple predictiveness and derived attention models paid more attention to the high value cues (A and B), whereas CompAct paid more attention to the low value cues (C and D). As shown in Table 11, these results were obtained over the entire range of parameter values. That is, each model behaved in qualitatively the same way, regardless of parameter values. Thus the task design in Table 7 is effective in distinguishing competitive predictiveness from the other two theories.

Experiment 2-a

Experiment 2-a used the same task design as Simulation 2 (Table 7), to test the contrasting model predictions demonstrated in that simulation. In particular, we test whether humans participants behave like the simple predictiveness and derived attention models, paying attention to the cues associated with high overall reward value (A and B), or like CompAct, paying more attention to the cues associated with greater reward difference between the responses (C and D).

Method

Participants. Participants were recruited using Amazon Mechanical Turk ($n = 60$). They were compensated with a base payment of \$2.00 plus a bonus payment of up to \$2.50 based on performance. The task required between 10 and 30 minutes to complete for most participants.

Apparatus and Materials. These were identical to those used in the other experiments.

Design. The experimental design is described in simulation 2 and summarized in Table 7.

Procedure. The procedure was the same as in the other experiments.

Results and Discussion

Out of 60 participants, 15 successfully completed the entire task (19 failed at stage 0, 22 at stage 1 and 6 at stage 2). Discordant scores were negative at a marginal level of significance, suggesting greater attention to the low value cues (C and D) ($M = -2.33$, $SD = 4.74$, $d = 0.49$, $t_{14} = -1.91$, $p = .077$). This result, if replicated, provides evidence for the competitive predictiveness principle as opposed to simple predictiveness or derived attention. Accurate learning and understanding of the test procedure were indicated by results from old concordant ($M = 9.47$, $SD = 4.77$, $d = 1.98$, $t_{14} = 7.68$, $p < .001$), new A/B concordant ($M = 3.60$, $SD = 5.22$, $d = 0.69$, $t_{14} = 2.67$, $p = .018$) and new C/D concordant ($M = 5.10$, $SD = 6.15$, $d = 0.83$, $t_{14} = 3.21$, $p = .006$) trials. Choice proportions in the test stage are given in Table 8.

In summary, we found preliminary evidence that participants paid more attention to the low value cues (C and D) than to the high value cues (A and B). This is directly contrary to the predictions of the simple predictiveness and derived attention models, but can be explained by competitive predictiveness (as implemented in CompAct). The 100/95 vs. 50/1 reward structure causes the irrelevant cues X and Y to compete for attention with cues A and B on high value trials but highlight cues C and D on low value trials. This effect of outcome value on attention can thus be explained by the same mechanism that Kruschke (2009) has proposed underlies highlighting effects.

Experiment 2-b

Experiment 2-b was an attempt to independently replicate the result of experiment 2-a (more attention to the low value cues C and D) with a larger sample.

Method

Participants. Participants were recruited using Amazon Mechanical Turk ($n = 139$). They were compensated with a base payment of $2.00 plus a bonus payment of up to $2.50 based on performance. The task required between 10 and 30 minutes to complete for most participants.

Apparatus and Materials. These were identical to those used in the other experiments.

Design. The design was identical to that of Experiment 2-a.

Procedure. The procedure was the same as in the other experiments.

Results and Discussion

Fifty-seven out of 139 participants successfully completed the whole experiment (18 failed at stage 0, 49 at stage 1 and 10 at stage 2). Mean discordant scores were negative ($M = -1.76$, $SD = 5.23$, $d = 0.34$, $t_{56} = -2.55$, $p = .014$). Scores for the old concordant ($M = 9.39$, $SD = 4.44$, $d = 2.12$, $t_{56} = 15.97$, $p < .0001$), new AB concordant ($M = 3.75$, $SD = 5.10$, $d = 0.73$, $t_{56} = 5.55$, $p < .0001$) and new CD concordant ($M = 4.84$, $SD = 5.95$, $d = 0.81$, $t_{56} = 6.14$, $p < .0001$) test item pairs were all positive. Choice proportions in the test stage are given in Table 8.

We thus successfully replicated the result of Experiment 2-a, viz. that participants pay more attention to the low value cues (C and D) in the 100/95 vs. 50/1 design. This result provides evidence for competitive predictiveness and against derived attention.

General Discussion

Summary of Findings

The present work resolves an important tension in previous work on attention learning, between findings of attention driven by cue predictiveness and of attention driven by association strength. As evidence for the latter, Le Pelley et al. (2016) point to the result that cues associated with larger rewards received more attention in a certain associative learning task (Le Pelley et al., 2013). They proposed that theories of attention learning based on a principle of predictiveness must thus be supplemented by a form of attention derived directly from association strength. We challenge this argument on two grounds, one empirical and the other theoretical. After replicating their "alien slots" effect (Experiments 1-a and 1-b), found using our 100/95 vs. 50/1 experimental design that this effect can be reversed (Experiments 2-a and 2-b). Thus attention is not merely based on the overall size of rewards, but is affected by the difference between the rewards obtained after different responses.

Second, our simulations showed that, contrary to initial intuitions, both non-competitive and competitive predictiveness models can produce the original alien slots effect (greater attention to cues associated with larger rewards) without any additional derived attention mechanism. However, the non-competitive predictiveness and derived attention models could not account for our Experiment 2 results, leaving the competitive predictiveness theory (embodied by CompAct and EXIT) as the only one capable of explaining all of our findings. We confirmed through further simulations (reported in the Appendix) that all three models can produce correlation effects, but only the competitive predictiveness model can explain highlighting effects. This suggests that highlighting and the results of the various alien slots experiments are produced by the same mechanism, viz. attention learning through competitive predictiveness.

Methodological Considerations

The present experiments are subject to a significant methodological concern: only about 40% of participants were able to complete the task and thus provide usable data. Aside from the expense and inconvenience, this raises the question as to whether our findings apply only to particularly motivated or capable participants. One reason to expect that our results hold more generally is that we replicated the original alien slots effect (Le Pelley et al., 2013) in Experiments 1-a and 1-b using the same inclusion criteria as in Experiments 2-a and 2-b. Le Pelley et al. ran participants for a fixed number of trials and included all subjects in their analysis; thus our performance criterion method does not appear to have qualitatively affected the results. Because of these differences in procedure, it is not clear whether our participants performed worse than those in Le Pelley et al.'s initial study. The average of their participants' choice accuracy did improve over training, but only reached a final level of around 60-80% in different conditions. A simple way to test whether our results hold generally would be to collect test trial data from all participants (not just those who reached a criterion level of performance), running them for a fixed number of trials like in Le Pelley et al.'s study.

The stage 1 task may be difficult because it requires the participant to keep track of many cues at once. Task difficulty has been found to increase with the number of cues (Collins, Brown, Gold, Waltz, & Frank, 2014). Other researchers have reported low performance in a correlation effect design similar to that given in table 1 (Mitchell et al., 2012). They solved the problem by breaking up stage 1 training into separate subtasks, giving participants alternating blocks with only half of the stimuli. Such a procedure could be implemented in our experiments by providing different distractor cues for high value and low value trials. In other words, instead of pairing C and D with X and Y (as in table 6 or 7), we would use new cues V and W. Simulations of CompAct and the derived attention model (not reported here) show that this would not qualitatively change their predictions. The theoretical analysis of these "separate irrelevant cue" designs would be similar to that of our existing ones, but simpler. It is worth noting that none of the models examined here predicts that performance will drop with the number of cues present in the experiment or in any stage thereof (as opposed to the number of cues present on a single trial): this suggests a separate line of investigation relating associative learning to working memory. Relatedly, we might also observe higher overall performance if we give memorable labels to the response options (Lupyan et al., 2007).

Theoretical Analysis

As our simulations make clear, models of attention learning based on predictiveness (viz., CompAct and EXIT, as well as the simple predictiveness model) reproduce the original alien slot machine effect as well as our modified replications (Experiments 1-a and 1-b). We thus do not need to add any derived attention mechanism to these models in order to account for value effects on attention. This is a nice outcome: theoretical simplicity is always desirable.

This result may seem paradoxical: the high value cues (A and B) and low value cues (C and D) all consistently indicate which action is best, so one is tempted to consider them all equally predictive. The problem with this reasoning is that "predictiveness" needs to be more precisely defined. When we deal with discrete outcomes such as "action 1 is better than action 2" or "the stimulus belongs to category 1", it is easy to define the predictiveness of cues in terms of conditional probabilities or mutual information between the cue and outcome. However, defining predictiveness in terms of probabilities breaks down when either the outcome or cue becomes a continuous variable, and the former is the case in alien slot machine experiments. Indeed, this simple notion of predictiveness is not even fully adequate for discrete outcomes, as it ignores redundancy between cues (as in blocking) and interaction effects (e.g. in "exclusive or" category structures).

Predictiveness cannot be defined solely in terms of the environment, because models differ in how they combine cues to generate expectations. The optimal allocation of attention for a model built on one principle may be very different from that for a model built on another principle. After all, a cue is "predictive" only to the degree to which it helps to form accurate predictions of outcomes, which depends on how those predictions are generated. For example, in some exemplar models, predictions are based on summed similarity to previous stimuli (cue combinations), and attention works by moderating how stimulus dimensions contribute to the similarity calculation (Nosofsky, 1986). This is quite different from the models presented here, and might lead to different optimal attention allocations, even for the same task design.

For the predictiveness models considered here, based in the Rescorla-Wagner architecture, we have proposed a common notion of predictiveness in terms of error reduction. Specifically, the learning rule for attention is defined by (stochastic) gradient descent on squared prediction error ($\frac{1}{2}\sum_k \delta_k^2$). The resulting update shifts the attention weights toward those values that minimize prediction error on the previous trial (and hence hopefully on future trials as well). It thus seems reasonable to say that the attention shifts that come from gradient descent are based on predictiveness. Note, however, that this predictiveness is relative to the current knowledge state of the learner, rather than to some objective description of the environment. Thus for example in blocking (Kamin, 1968) a cue might be "non-predictive" from the learner's perspective, even if objectively it is perfectly correlated with the outcome.

In a modified Rescorla-Wagner model (Equation 5), the Chain Rule for computing derivatives implies that the gradient will always include the following factor (see Equations 7 and 12):

$$-\frac{\partial}{\partial a_i}\left(\frac{1}{2}\sum_k \delta_k^2\right) = \sum_k \delta_k \cdot W_{ki} s_i. \tag{16}$$

The term on the right-hand side is related to the covariance between cue predictions ($W_{ki}s_i$) and unexpected outcomes (δ_k), which can naturally be interpreted as the degree to which cue i predicts things (relative to the learner's current knowledge).

Thus it seems reasonable to define cue i's *predictiveness* as $\sum_k \delta_k \cdot W_{ki} s_i$ in any modified Rescorla-Wagner model. This clarifies what we mean when we say that changes in attention are determined by predictiveness. This notion of predictiveness can be applied equally well to situations involving discrete outcomes (e.g., highlighting) or continuous outcomes (e.g., alien slots). In the simple predictiveness model, this is the sole determinant of attention learning (Equation 7). When we compute the gradient based on CompAct's more complex competitive attention rule (Equation 11) we naturally obtain a learning rule based on competitive predictiveness, i.e. the relative size of $\sum_k \delta_k \cdot W_{ki} s_i$ for different cues (Equation 12). This predictiveness term also shows up in the learning rules for other, even more complex models based on the same principles (e.g., Kruschke, 2001).

This analysis also suggests that derived attention and simple predictiveness are not so different as they might appear: both are driven by the magnitude of association weights (W_{ki}). For simplicity, consider a case in which there is only one outcome being predicted (so we can omit the subscript k and refer to association weights by w_i for each cue i). Also assume that the cue-reward contingencies remain constant, so that the sign of w_i tends to remain the same. Then according to the derived attention model, changes of attention to cue i are determined by

$$\Delta a_i = \Delta|w_i| = |w_i^{t+1}| - |w_i^t| = \text{sign}(w_i)\Big(w_i^{t+1} - w_i^t\Big) = \text{sign}(w_i)\Delta w_i \propto \delta \; \text{sign}(w_i) s_i. \quad (17)$$

The attention learning rule of the simple predictiveness model differs only by a factor of $|w_i|$:

$$\Delta a_i \propto \delta w_i s_i = \delta \; \text{sign}(w_i)|w_i| s_i = |w_i|\Big(\delta \; \text{sign}(w_i) s_i\Big). \quad (18)$$

The simple predictiveness and derived attention models therefore have very similar attention learning rules. This helps to explain why these two models make the same qualitative predictions in all the experiments considered here (see Tables 9-13), including ones where their predictions are empirically supported (correlation effects, Le Pelley et al. 2013, Experiments 1-a and 1-b) and ones where they are not (Experiments 2-a and 2-b, highlighting).

Thus the critical contrast is not between derived attention and predictiveness, but rather between predictiveness with competition (as in CompAct) or without competition (as in the simple predictiveness model). It is *competition* which allows CompAct and EXIT to explain highlighting and the result of our Experiment 2 (the 100/95 vs. 50/1 design), as well as other phenomena.

Thus our main contribution is to show that attention learning based on predictiveness can explain value-based effects, contrary to previous claims (Le Pelley et al., 2013), and furthermore that competitive predictiveness provides a unified explanation for a wide variety of attentional phenomena in associative learning. Although these links between theory and data are new, the idea that attention is based on competitive predictiveness is not (Kruschke, 2001; Mackintosh, 1975). What is special about our new "competitive activation" model (CompAct) compared to other models that embody this principle?

CompAct can be seen as a simplified version of Kruschke's EXIT model (Kruschke, 2001). In EXIT, attention to cues is based on a complex exemplar-mediated similarity comparison, meaning that similar stimulus vectors will receive similar attention weights. Thus the simple notion of "attention to a cue" is not present in EXIT: attention depends on the stimulus in question and could differ for different parts of the stimulus space. There is some evidence that such effects exist (Aha & Goldstone, 1992), but they are not encountered in most learning experiments. One might feel therefore that CompAct's mechanism is preferable, as it makes theoretical analysis far simpler. EXIT also shifts attention by a large amount on each trial prior to association learning, and then retains only a small amount of that shift as permanent attention learning. In CompAct, all attention shifts constitute permanent learning, and they occur simultaneously with association learning. So far, we have not found that EXIT's rapid attention shifts make any difference to model predictions, although the issue bears further study. Finally, EXIT includes a parameter which determines the amount of competition between cues by adjusting the norm according to which the stimulus vector is normalized (CompAct always uses the L^1 norm). This mechanism can also be incorporated in CompAct, but for this paper have decided to keep attentional competition fixed, for the sake of simplicity. In all of our simulations so far, we have found that CompAct can explain everything that EXIT does using only the principle of competitive attention. We conjecture that EXIT's additional mechanisms (exemplar-mediated attention and rapid attention shifts) are unnecessary, but this question requires more thorough study. At any rate, CompAct's simplicity makes it easier to analyze why it behaves as it does in any given task.

Mackintosh's (1975) model also embodies the competitive predictiveness principle, and thus ought to be compared to CompAct. The principal difference is that Mackintosh's model uses cue-specific prediction errors to learn associations, rather than the combined prediction errors used by CompAct (and EXIT). These cue-specific prediction errors are used to determine cues' "predictiveness" and hence the learning rule for attention. In Le Pelley et al.'s (2016) implementation of Mackintosh's ideas, the learning rule for attention is as follows (adapted to our notation):

$$\Delta a_i = \mu \sum_k \left(\left| r_k - \sum_{j \neq i} w_{kj} s_j \right| - |r_k - w_{ki} s_i| \right). \tag{19}$$

Comparing this to CompAct's learning rule (Equation 12), we see that the two models implement the same competitive principle in somewhat different ways (Kruschke, 2001, has previously pointed out this similarity in reference to EXIT). Nevertheless, in simulations we have found that Mackintosh's model—implemented as in Le Pelley et al. (2016)—cannot produce highlighting effects or the complex value effect observed in our Experiment 2. Analyzing this result here would take us too far afield, but it does suggest that CompAct provides a better implementation of the competitive predictiveness idea.

Conclusions

Correlation effects, highlighting and many other learning effects can be explained by the theory that cues compete for attention based on how "predictive" they are, with predictiveness being related to the correlation between cue predictions and unexpected outcomes. We have shown here that this principle (as embodied in CompAct) can also explain simple value effects on attention, in which cues associated with large rewards receive more attention. Models in which attention is derived directly from association strengths can explain these value effects, but not highlighting. We have replicated a simple value effect experimentally, and also demonstrated that the effect depends on relative rewards between alternative response options in a manner predicted by competitive predictiveness (in particular by CompAct) but not by derived attention. We thus conclude that the competitive predictiveness principle embodied in CompAct is sufficient to account for a broad array of learning phenomena, and that we do not require any additional derived attention mechanisms.

References

Aha, W., & Goldstone, R. (1992). Concept learning and flexible weighting. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, 534-539.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, *108*(25), 10367–10371. doi: 10.1073/pnas.1104047108

Bridle, J. S. (1990). Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimates of parameters. *Advances in Neural Information Processing Systems*, *2*, 211-217.

Collins, A., Brown, J., Gold, J., Waltz, J., & Frank, M. (2014). Working memory contributions to reinforcement learning impairments in schizophrenia. *Journal of Neuroscience*, *34*, 13747-13756.

Haselgrove, M., Esber, G. R., Pearce, J. M., & Jones, P. M. (2010). Two kinds of attention in Pavlovian conditioning: evidence for a hybrid model of learning. *Journal of Experimental Psychology. Animal Behavior Processes*, *36*(4), 456–470. doi: 10.1037/a0018528

Hume, D. (1748/2007). *Enquiry concerning Human Understanding.* Oxford, UK: Oxford University Press.

Kamin, L. J. (1968). "attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior, 1967: Aversive stimulation* (p. 9-31). University of Miami Press.

Kruschke, J. K. (1996). Base rates in category learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*(1), 3–26.

Kruschke, J. K. (2001). Toward a Unified Model of Attention in Associative Learning. *Journal of Mathematical Psychology*, *45*(6), 812–863. doi: 10.1006/jmps.2000.1354

Kruschke, J. K. (2009). Chapter 5 Highlighting. In *Psychology of Learning and Motivation* (Vol. 51, pp. 153–185). Elsevier. doi: 10.1016/S0079-7421(09)51005-5

Le Pelley, M. E., & McLaren, I. P. L. (2003). Learned associability and associative change in human causal learning. *Quarterly Journal of Experimental Psychology: Section B*, *56*(1), 68.

Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N., & Wills, A. J. (2016). Attention and associative learning in humans: An integrative review. *Psychological Bulletin*, *142*(10), 1111–1140. doi: http://dx.doi.org.colorado.idm.oclc.org/10.1037/bul0000064

Le Pelley, M. E., Mitchell, C. J., & Johnson, A. M. (2013). Outcome value influences attentional biases in human associative learning: Dissociable effects of training and instruction. *Journal of Experimental Psychology: Animal Behavior Processes*, *39*(1), 39–55. doi: 10.1037/a0031230

Le Pelley, M. E., Pearson, D., Griffiths, O., & Beesley, T. (2015). When goals conflict with values: Counterproductive attentional and oculomotor capture by reward-related stimuli. *Journal of Experimental Psychology: General*, *144*(1), 158.

Luce, D. (1959). *Individual choice behavior.* Wiley.

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language Is Not Just for Talking: Redundant Labels Facilitate Learning of Novel Categories. *Psychological Science*, *18*(12), 1077–1083.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*(4), 276–298. doi: http://dx.doi.org.colorado.idm.oclc.org/10.1037/h0076778

Mitchell, C. J., Griffiths, O., Seetoo, J., & Lovibond, P. F. (2012). Attentional mechanisms in learned predictiveness. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*(2), 191.

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

Rescorla, R. A., & Wagner, A. R. (1972). 3 A Theory of Pavlovian Conditioning:

Variations in the Effectiveness of Reinforcement and Nonreinforcement. *Classical conditioning II: Current research and theory*, *2*, 64–99.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986, October). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. Retrieved 2018-08-07, from `https://www.nature.com/articles/323533a0` doi: 10.1038/323533a0

Sutherland, N. S., & Holgate, V. (1966). Two-cue discrimination learning in rats. *Journal of Comparative and Physiological Psychology*, *61*(2), 198–207. doi: 10.1037/h0023144

Sutherland, N. S., & Mackintosh, N. J. (1971). *Mechanisms of Animal Discrimination Learning*. Academic Press. (Google-Books-ID: 5c_YBAAAQBAJ)

Appendix

Simulations of Correlation and Highlighting Effects

We simulated CompAct, the simple predictiveness model and the derived attention model on experimental designs that produce correlation and highlighting effects: Le Pelley and McLaren (2003) as in Table 1, and Kruschke (1996) as in Table 2. Simulation details were the same as reported above, save that we capped attention weights at a maximum value of 1 rather than 1.5. This was necessary to prevent the behavior of the derived attention model from fluctuating wildly at larger values of λ. The reward for correct responses was coded as 1 and the reward for incorrect responses as 0.

For quantifying the correlation effect in our simulation of Le Pelley and McLaren (2003), we created a correlation effect index defined by

$$\text{correlation effect index} = P(III|A+C) + P(IV|B+D) - P(III|V+X) - P(IV|W+Y)$$

where $P(III|A+C)$ is the probability of choosing response III given the test stimulus $A+C$, etc. Higher values of this index indicate greater attention to the predictive cues (A, B, C and D) relative to the non-predictive cues (V, W, X and Y). The results are given in Table 12. Observe that all three models produce a correlation effect under all parameter values.

We created a choice index for the highlighting effect (Kruschke, 1996, Experiment 2) defined by

$$\text{highlighting index} = P(L_1|PE_1+PL_1)+P(L_2|PE_2+PL_2)-P(E_1|PE_1+PL_1)-P(E_2|PE_2+PL_2)$$

where $P(E_1|PE_1 + PL_1)$ is the probability of choosing response E_1 when presented with the compound cue $PE_1 + PL_1$, etc. Positive values of the highlighting index indicate a highlighting effect. The results are given in Table 13. Note that CompAct is the only one of the three models simulated that can produce a highlighting effect.

Derivation of CompAct's Attention Learning Rule

Attention is learned by gradient descent on squared error. Using the chain rule, we see that

$$\begin{aligned}
\frac{\partial}{\partial \eta_i}\frac{1}{2}\sum_k \delta_k^2 &= \sum_k \delta_k \frac{\partial}{\partial \eta_i}(r_k - z_k) \\
&= -\sum_k \delta_k \frac{\partial}{\partial \eta_i} z_k \\
&= -\sum_k \delta_k \frac{\partial}{\partial \eta_i}\sum_l W_{kl} a_l s_l \\
&= -\sum_k \delta_k \sum_l W_{kl} s_l \frac{\partial a_l}{\partial \eta_i}. \qquad (20)
\end{aligned}$$

We now must calculate $\frac{\partial a_l}{\partial \eta_i}$:

$$\begin{aligned}
\frac{\partial a_l}{\partial \eta_i} &= \frac{\partial}{\partial \eta_i}\left[\left(\sum_j |\eta_j s_j|\right)^{-1}\eta_l\right] \\
&= \left[\frac{\partial}{\partial \eta_i}\left(\sum_j |\eta_j s_j|\right)^{-1}\right]\eta_l + \left(\sum_j |\eta_j s_j|\right)^{-1}\frac{\partial \eta_l}{\partial \eta_i} \\
&= \left[-\left(\sum_j |\eta_j s_j|\right)^{-2}\frac{\partial}{\partial \eta_i}\sum_j |\eta_j s_j|\right]\eta_l + \left(\sum_j |\eta_j s_j|\right)^{-1} I_{\{i=l\}}
\end{aligned}$$

$$= \left(\sum_j |\eta_j s_j|\right)^{-1} \left[I_{\{i=l\}} - \left(\sum_j |\eta_j s_j|\right)^{-1} |s_i|\eta_l \right], \tag{21}$$

where $I_{\{i=l\}}$ equals unity if $i = l$ and is zero otherwise.

Substituting this result into Equation 20, we see that

$$\frac{\partial}{\partial \eta_i} \frac{1}{2} \sum_k \delta_k^2 = -\sum_k \delta_k \sum_l W_{kl} s_l \frac{\partial a_l}{\partial \eta_i}$$

$$= -\sum_k \delta_k \sum_l W_{kl} s_l \left(\sum_j |\eta_j s_j|\right)^{-1} \left[I_{\{i=l\}} - \left(\sum_j |\eta_j s_j|\right)^{-1} |s_i|\eta_l \right]$$

$$= -\left(\sum_j |\eta_j s_j|\right)^{-1} \sum_k \delta_k \left[W_{ki} s_i - \sum_l W_{kl} s_l \left(\sum_j |\eta_j s_j|\right)^{-1} |s_i|\eta_l \right]$$

$$= -\left(\sum_j |\eta_j s_j|\right)^{-1} \sum_k \delta_k \left[W_{ki} s_i - |s_i| \sum_l W_{kl} a_l s_l \right]$$

$$= -\left(\sum_j |\eta_j s_j|\right)^{-1} \sum_k \delta_k \left[(1 - |s_i| a_i) W_{ki} s_i - |s_i| \sum_{l \neq i} W_{kl} a_l s_l \right]$$

$$= -\left(\sum_j |\eta_j s_j|\right)^{-1} \sum_k \left[(1 - |s_i| a_i) \delta_k W_{ki} s_i - |s_i| \sum_{l \neq i} \delta_k W_{kl} a_l s_l \right]. \tag{22}$$

Multiplying by $-\mu$, we see that we have obtained the learning rule given in equation 12.

Equivalence of Different Versions of the Derived Attention Model

In this paper, we simulate the derived attention model with the assumption that attention (a) controls both learning (Equation 6) and expectancy (Equation 5). This makes it easier to compare derived attention to the simple predictiveness model and CompAct. However, Le Pelley and colleagues (Le Pelley et al., 2016) assume that attention does not affect expectancy, i.e. their simulations use Equation 1 to determine the model's expectations of reward, rather than Equation 5. We do not want the reader to be distracted by this difference, as it is irrelevant to the issues we investigate here.

In fact, as we shall show, these two versions of the derived attention model are formally equivalent. That is, if we set their attention and association weights such that they initially make the same predictions, then they will continue to make the same predictions after experiencing any sequence of stimuli and rewards.

First, we must determine under what conditions the two versions of the model make the same predictions. We shall designate the model in which attention controls learning and expectations Model 1, and the one in which attention controls only learning Model 2. Quantities associated with Model 2 shall be distinguished by having a tilde (~) placed over them. Recall that the expectation of reward in Model 1 is given by

$$z_k = \sum_i W_{ki} a_i s_i \tag{23}$$

and in Model 2 by

$$\tilde{z}_k = \sum_i \tilde{W}_{ki} s_i. \tag{24}$$

Thus we have $z_k = \tilde{z}_k$ for all possible $\mathbf{s}$ if and only if

$$\tilde{W}_{ki} = W_{ki} a_i. \tag{25}$$

We can accomplish this at the start of learning by choosing appropriate initial association weights, e.g. $W_{ki} = \tilde{W}_{ki} = 0$. Initializing weights at zero can be seen as representing a lack of knowledge about novel stimuli. Note that equality of expectancy ($z_k = \tilde{z}_k$) also implies equality of prediction error ($\delta_k = \tilde{\delta}_k$).

Now assume that Equation 25 is satisfied at the beginning of a trial. It will also be satisfied after that trial if the following condition holds:

$$\tilde{W}_{ki} + \Delta\tilde{W}_{ki} = (W_{ki} + \Delta W_{ki}) a_i \Leftrightarrow \Delta\tilde{W}_{ki} = \Delta W_{ki} a_i \Leftrightarrow \lambda \delta_k \tilde{a}_i s_i = (\lambda \delta_k a_i s_i) a_i \tag{26}$$

which is satisfied if

$$\tilde{a}_i = a_i^2. \tag{27}$$

Now let $a_i = \sum_k |W_{ki}|$ and $\tilde{a}_i = \sum_k |\tilde{W}_{ki}|$ as in Equation 14 (here we ignore the lower bound $a_{\min}$). Using Equation 25, we see that

$$\tilde{a}_i = \sum_k |\tilde{W}_{ki}| = \sum_k |W_{ki} a_i| = a_i \sum_k |W_{ki}| = a_i^2$$

and thus Equation 27 is satisfied. In other words, if both models start with null association weights (or any other values satisfying Equation 25), then after any sequence of stimuli and outcomes they will always be related by Equations 25 and 27, which in turn implies they will always generate the same expected outcomes ($\mathbf{z} = \tilde{\mathbf{z}}$) and hence the same predicted response probabilities.

Stage 1			Stage 2			Test Stage	
Cues	I	II	Cues	III	IV	Cues	Preferred
A + V	1	0	A + X	1	0	A + C	III (strong)
A + W	1	0	B + Y	0	1	B + D	IV (strong)
B + V	0	1	C + V	1	0	V + X	III (weak)
B + W	0	1	D + W	0	1	W + Y	IV (weak)
C + X	0	1	E + F	1	0	E + H	neither
C + Y	0	1	G + H	0	1	F + G	neither
D + X	1	0	I + J	1	0	I + J	III
D + Y	1	0	K + L	0	1	K + L	IV

Table 1

Design of experiment 1 from Le Pelley & McLaren (2003). Each row within each stage indicates a trial type, with + indicating a conjunction of cues (e.g., cues A and V presented together). Response options are indicated by I and II (stage 1) and III and IV (stage 2 and test stage). The rightmost two columns for stages 1 and 2 indicate which response is rewarded for each trial type. The rightmost column for the test stage shows the empirical result, namely which response subjects tend to select.

Stage 1					Stage 2					Test		
Cues	E_1	E_2	L_1	L_2	Cues	E_1	E_2	L_1	L_2	Cues	E	L
$PE_1 + I_1$	1	0	0	0	$PE_1 + I_1$	1	0	0	0	I	.80	.12
$PE_2 + I_2$	0	1	0	0	$PE_2 + I_2$	0	1	0	0	PE	.93	.01
					$PL_1 + I_1$	0	0	1	0	PL	.03	.92
					$PL_2 + I_2$	0	0	0	1	$PE + PL$	.32	.65

Table 2

Illustration of the highlighting effect (Kruschke, 1996, Experiment 2). The rightmost four columns for stages 1 and 2 indicate which response is rewarded for each trial type. The rightmost two columns for test show the empirical response proportions in that stage. Due to the symmetry of the design, the test results for I_1 are averaged with I_2, PE_1 with PE_2, etc. The critical finding is that participants tend to choose the "late" outcome (L_1 or L_2) on $PE + PL$ test trials.

Stage 1			Stage 2			Test Stage	
Cues	I	II	Cues	III	IV	Cues	Preferred
A + V	150	0	A + D	*usp.*	0	A + C	III
A + W	150	0	B + C	0	*usp.*	B + D	IV
B + X	150	0	E + F	*usp.*	0		
B + Y	150	0	G + H	0	*usp.*		
C + V	0	1					
C + W	0	1					
D + X	0	1					
D + Y	0	1					

Table 3

*Le Pelley et al. (2013) Experiment 1: Reward value biases attention. Each row within each stage indicates a trial type, with + indicating a conjunction of cues (e.g., cues A and V presented together). Response options are indicated by I and II (stage 1) and III and IV (stage 2 and test stage). The rightmost two columns for stages 1 and 2 indicate which response is rewarded for each trial type. The right column for the test stage shows the empirical result, namely which response subjects tend to select. The abbreviation "*usp.*" stands for "unspecified"; subjects were merely told these responses were correct, and they presumably treated these outcomes as positive rewards of unknown value.*

Stage 1			Stage 2			Test Stage		
Cues	I	II	Cues	III	IV	Cues	Predicted	Trial Type
A + V	100	1	A + D	100	0	A + C	III	discordant
A + W	100	1	B + C	0	100	B + D	IV	discordant
B + X	1	100	E + F	100	0	E + F	III	old concordant
B + Y	1	100	G + H	0	100	G + H	IV	old concordant
C + V	2	1				A + E	III	new A/B concordant
C + W	2	1				B + G	IV	new A/B concordant
D + X	1	2				D + F	III	new C/D concordant
D + Y	1	2				C + H	IV	new C/D concordant

Table 4

Early version of the 100/1 vs. 2/1 design, as used in Experiment 1-a. The "predicted' column lists the responses preferences predicted on the basis of previous research (Le Pelley et al., 2013), on the assumption that subjects would attend to high value cues during stage 1 and that this would affect their learning during stage 2.

Cues	Action 1	Action 2
$\Phi + \Phi$	1	0
$\Theta + \Theta$	0	1

Table 5

Design of the tutorial stage used in all of our experiments.

Stage 1			Stage 2			Test Stage		
Cues	I	II	Cues	III	IV	Cues	Predicted	Trial Type
A + X	100	1	A + D	100	0	A + C	III	discordant
A + Y	100	1	B + C	0	100	B + D	IV	discordant
B + X	1	100	E + F	100	0	E + F	III	old concordant
B + Y	1	100	G + H	0	100	G + H	IV	old concordant
C + X	2	1				A + E	III	new A/B concordant
C + Y	2	1				B + G	IV	new A/B concordant
D + X	1	2				D + F	III	new C/D concordant
D + Y	1	2				C + H	IV	new C/D concordant

Table 6

100/1 vs. 2/1 design used in Experiment 1-b. The "predicted" column lists the responses preferences predicted on the basis of previous research (Le Pelley et al., 2013), on the assumption that subjects would attend to high value cues during stage 1 and that this would affect their learning during stage 2.

Stage 1			Stage 2			Test Stage		
Cues	I	II	Cues	III	IV	Cues	Predicted	Trial Type
A + X	100	95	A + D	100	0	A + C	?	discordant
A + Y	100	95	B + C	0	100	B + D	?	discordant
B + X	95	100	E + F	100	0	E + F	III	old concordant
B + Y	95	100	G + H	0	100	G + H	IV	old concordant
C + X	50	1				A + E	III	new A/B concordant
C + Y	50	1				B + G	IV	new A/B concordant
D + X	1	50				D + F	III	new C/D concordant
D + Y	1	50				C + H	IV	new C/D concordant

Table 7

100/95 vs. 50/1 design used in Experiments 2-a and 2-b. The derived attention theory predicts greater attention to the "high value" cues (A and B) whereas competitive predictiveness models predict the opposite (hence the question marks in the "predicted" column).

	Experiment 1-a		Experiment 1-b		Experiment 2-a		Experiment 2-b	
Cues	III	IV	III	IV	III	IV	III	IV
A + C	.67	.33	.68	.32	.33	.67	.39	.61
B + D	.21	.79	.32	.68	.60	.40	.54	.46
E + F	.92	.08	.82	.18	.93	.07	.90	.10
G + H	.00	1.00	.05	.95	.13	.87	.08	.92
A + E	.79	.21	.91	.09	.57	.43	.62	.38
B + G	.17	.83	.18	.82	.27	.73	.32	.68
D + F	.75	.25	.77	.23	.80	.20	.69	.31
C + H	.29	.71	.18	.82	.27	.73	.22	.78

Table 8

Proportions of choices in the test stage of each experiment.

		Maximum Choice Index		
Model	Parameters	Choice Index	A/B Atn.	C/D Atn.
Simple Predictiveness	$\lambda = 0.35, \mu = 0.75$	1.51	1.47	0.11
CompAct	$\lambda = 0.4, \mu = 0.95$	0.59	1.60	1.11
Derived Attention	$\lambda = 0.35$	1.51	1.21	0.10
		Minimum Choice Index		
Model	Parameters	Choice Index	A/B Atn.	C/D Atn.
Simple Predictiveness	$\lambda = 0.05, \mu = 0.05$	0.08	0.36	0.10
CompAct	$\lambda = 0.05, \mu = 0.05$	0.02	1.09	1.01
Derived Attention	$\lambda = 0.05$	0.06	0.28	0.10

Table 9

Results of simulation 1 for the original alien slot machine experiment (Le Pelley et al., 2013). A/B atn. is the average of attention to cues A and B at the end of stage 1 (a for simple predictiveness and derived attention models, η for CompAct). C/D atn. is the same for cues C and D. Choice index (ranging from -2 to 2) is defined in the text; positive values indicate greater attention to cues A and B, while negative values indicate greater attention to cues C and D. Names of models consistent with empirical data are printed in bold type.

		Maximum Choice Index		
Model	Parameters	Choice Index	A/B Atn.	C/D Atn.
Simple Predictiveness	$\lambda = 0.9, \mu = 0.95$	1.53	1.00	0.12
CompAct	$\lambda = 0.4, \mu = 0.95$	0.42	1.34	1.06
Derived Attention	$\lambda = 0.3$	1.51	0.99	0.10
		Minimum Choice Index		
Model	Parameters	Choice Index	A/B Atn.	C/D Atn.
Simple Predictiveness	$\lambda = 0.05, \mu = 0.05$	0.02	0.20	0.10
CompAct	$\lambda = 0.05, \mu = 0.05$	0.01	1.04	1.01
Derived Attention	$\lambda = 0.05$	0.01	0.14	0.10

Table 10

Results of simulation 1 for Experiment 1-b. A/B atn. is the average of attention to cues A and B at the end of stage 1 (a for simple predictiveness and derived attention models, η for CompAct). C/D atn. is the same for cues C and D. Choice index (ranging from -2 to 2) is defined in the text; positive values indicate greater attention to cues A and B, while negative values indicate greater attention to cues C and D. Names of models consistent with empirical data are printed in bold type.

		Maximum Choice Index		
Model	Parameters	Choice Index	A/B Atn.	C/D Atn.
Simple Predictiveness	$\lambda = 0.15, \mu = 0.35$	1.23	1.20	0.42
CompAct	$\lambda = 0.05, \mu = 0.05$	-0.01	0.98	1.00
Derived Attention	$\lambda = 0.15$	1.33	1.13	0.27
		Minimum Choice Index		
Model	Parameters	Choice Index	A/B Atn.	C/D Atn.
Simple Predictiveness	$\lambda = 0.05, \mu = 0.05$	0.05	0.30	0.12
CompAct	$\lambda = 0.15, \mu = 0.95$	-0.92	0.41	1.51
Derived Attention	$\lambda = 0.75$	0.06	1.15	0.70

Table 11

Results of simulation 2, for Experiments 2-a/b. A/B atn. is the average of attention to cues A and B at the end of stage 1 (a for simple predictiveness and derived attention models, η for CompAct). C/D atn. is the same for cues C and D. Choice index (ranging from -2 to 2) is defined in the text; positive values indicate greater attention to cues A and B, while negative values indicate greater attention to cues C and D. Names of models consistent with empirical data are printed in bold type.

		Maximum Correlation Effect Index		
Model	Parameters	Cor. Effect Index	Pred. Cue Atn.	Non-Pred. Cue Atn.
Simple Predictiveness	$\lambda = 0.95, \mu = 0.95$	0.76	0.96	0.31
CompAct	$\lambda = 0.25, \mu = 0.95$	0.48	1.41	0.24
Derived Attention	$\lambda = 0.95$	0.70	0.99	0.60
		Minimum Correlation Effect Index		
Model	Parameters	Cor. Effect Index	Pred. Cue Atn.	Non-Pred. Cue Atn.
Simple Predictiveness	$\lambda = 0.05, \mu = 0.05$	0.01	0.20	0.15
CompAct	$\lambda = 0.05, \mu = 0.05$	0.01	1.04	0.96
Derived Attention	$\lambda = 0.05$	0.00	0.14	0.14

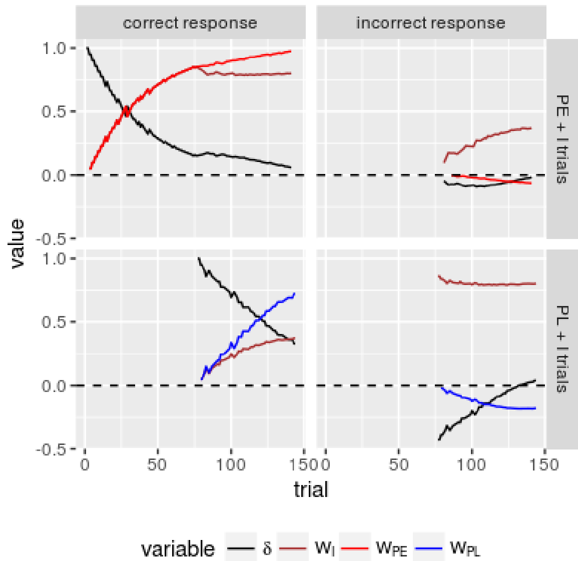
Table 12

Simulation results for the correlation effect (Le Pelley & McLaren, 2003). Pred. cue atn. is the average of attention to cues A, B, C and D at the end of stage 1 (a for simple predictiveness and derived attention models, η for CompAct). Non-pred. cue atn. is the same for cues V, W, X and Y. The correlation effect index (ranging from -2 to 2) is defined in appendix A; positive values indicate greater attention to predictive than to non-predictive cues. Names of models consistent with empirical data are printed in bold type.

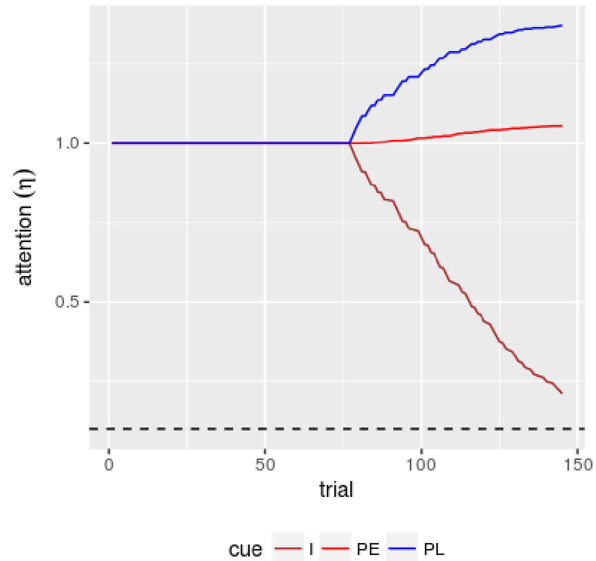
		Maximum Highlighting Index		
Model	Parameters	HL Effect Index	Early Cue Atn.	Late Cue Atn.
Simple Predictiveness	$\lambda = 0.95, \mu = 0.95$	-0.05	0.80	0.84
CompAct	$\lambda = 0.25, \mu = 0.95$	0.14	1.04	1.40
Derived Attention	$\lambda = 0.5$	0.00	1.00	0.59
		Minimum Highlighting Index		
Model	Parameters	HL Effect Index	Early Cue Atn.	Late Cue Atn.
Simple Predictiveness	$\lambda = 0.4, \mu = 0.3$	-0.62	1.00	0.48
CompAct	$\lambda = 0.05, \mu = 0.05$	-0.08	1.01	1.04
Derived Attention	$\lambda = 0.9$	-0.02	1.00	0.78

Table 13

Simulation results for highlighting (Kruschke, 1996, Experiment 2). Early cue atn. is the average of attention to cues PE_1 and PE_2 at the end of the experiment (a for simple predictiveness and derived attention models, η for CompAct). Late cue atn. is the same for cues PL_1 and PL_2. The highlighting index (ranging from -2 to 2) is defined in appendix A; positive values indicate a highlighting effect. Names of models consistent with empirical data are printed in bold type.

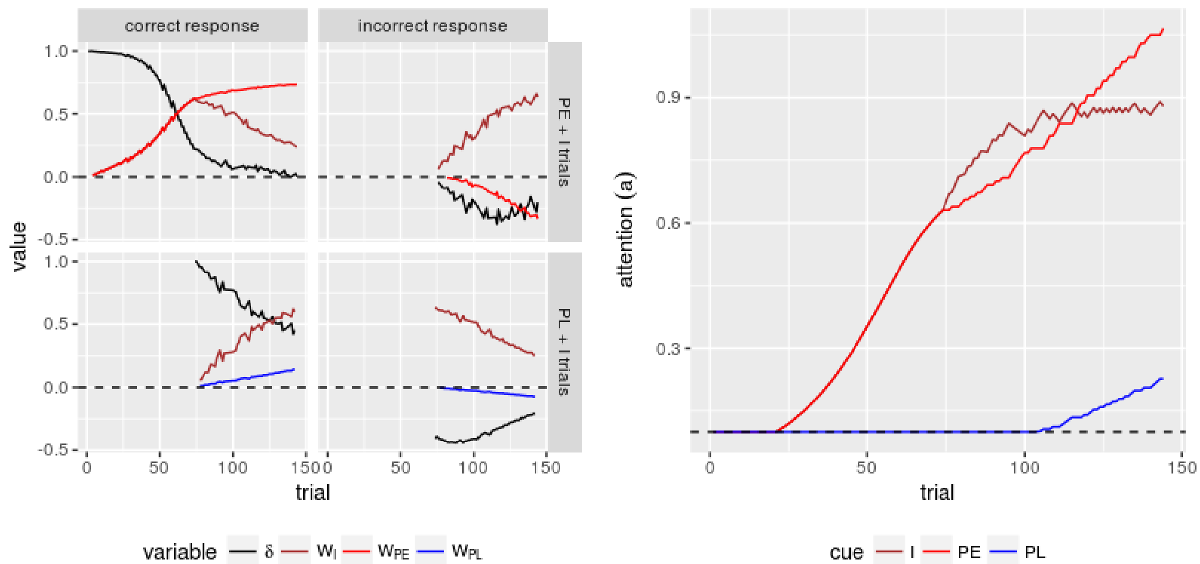


(a) Associations and prediction error



(b) Attention weights

Figure 1. Simulation of CompAct on a highlighting design (Kruschke, 1996, Experiment 2). Parameters were as follows: $\lambda = 0.1$, $\mu = 0.5$, $\xi = 2.0$. Associations and attention weights have each been combined according to the symmetry of the design: e.g. the red line in the top left panel of the left graph includes both $PE_1 \rightarrow E_1$ associations and $PE_2 \rightarrow E_2$ associations. For $PE_1 + I_1$ trials the correct response is E_1 and by "incorrect response" we mean L_1 (not E_2 or L_2); this is similar for the remaining trial types. Note that $PL + I$ trials do not occur until the second stage of the experiment (trial 76). CompAct attends most strongly to the PL cues, leading it to select response L in the critical $PE + PL$ test trials (not shown), as do human participants.



(a) Associations and prediction error (b) Attention weights

Figure 2. Simulation of the derived attention model on a highlighting design (Kruschke, 1996, Experiment 2). Parameters were as follows: $\lambda = 0.1$, $a_{\min} = 0.1$, $\xi = 2.0$. Associations and attention weights have been combined together according to the symmetry of the design: e.g. the red line in the top left panel of the left graph includes both $PE_1 \rightarrow E_1$ associations and $PE_2 \rightarrow E_2$ associations. For $PE_1 + I_1$ trials the correct response is E_1 and by "incorrect response" we mean L_1; this is similar for the remaining trial types. Note that $PL + I$ trials do not occur until the second stage of the experiment (trial 76). The derived attention model attends most strongly to the PE cues, leading it to select response E in the critical $PE + PL$ test trials (not shown), in contrast to human participants.

Figure 3. Screenshots of the task, before choice (left) and during feedback (right).

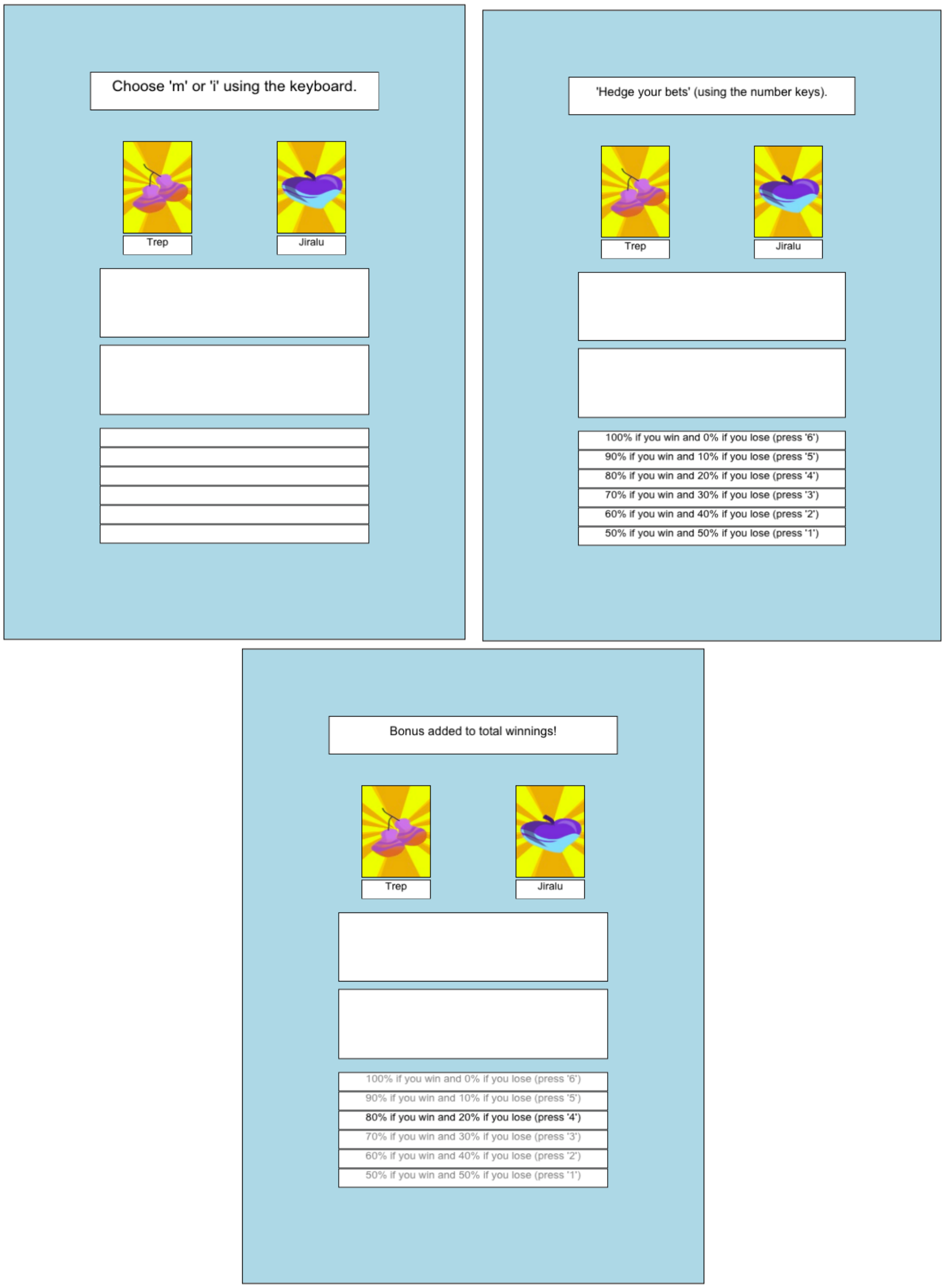
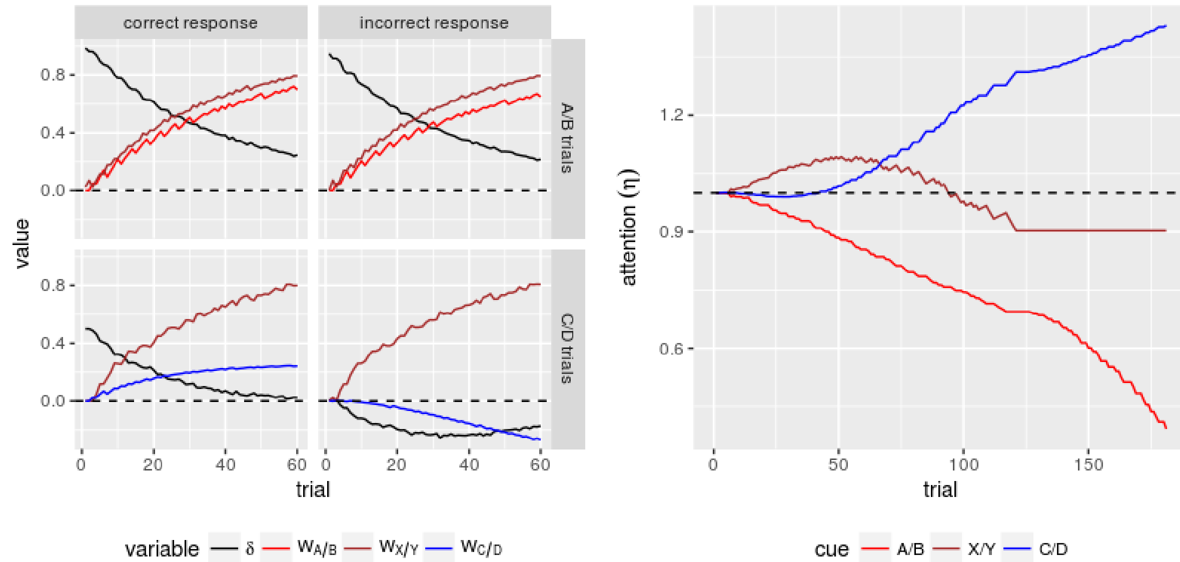


Figure 4. Screenshots of the test stage.



(a) Associations and prediction error (b) Attention weights

Figure 5. Simulation of CompAct on stage 1 of Experiment 2. Parameters were as follows: $\lambda = 0.1$, $\mu = 0.5$, $\xi = 2.00$. Reward values were divided by a factor of 100. Different associations have been combined together according to the symmetry of the design: e.g. the red line in the top left panel includes both A $\rightarrow$ I associations and B $\rightarrow$ II associations.