Blocking Requires Uncertainty about Novel Cues

Pei-Pei Liu (peiliu@notes.cc.sunysb.edu)
Department of Psychology
Stony Brook, NY 11729-2500 USA

Christian C. Luhmann (christian.luhmann@stonybrook.edu)
Department of Psychology
Stony Brook, NY 11729-2500 USA

Abstract

Blocking is a well-studied learning phenomenon in which previous learning inhibits subsequent learning about novel cues. Existing models provide different explanations for blocking and predict different beliefs about novel cues early in the second phase of blocking. Two experiments examined learners' beliefs when first encountering novel cues. The results suggest that the introduction of the novel cue in the second phase of a blocking paradigm adds uncertainty and that learners entertain the possibility that novel cues are preventative. A novel computational account is proposed to explain people's beliefs, because existing models cannot fully account for these findings.

Keywords: learning, blocking, uncertainty, Bayesian inference

Blocking was first reported by Kamin (1969) and is one of the most intensively studied phenomena in the field of learning (Pineno & Miller, 2007). In a blocking paradigm, participants first learn that the presence of a single cue event (cue A) is always followed by a certain outcome (O). Subsequently, cue A is paired with another, novel event (cue X), and this pair is followed by the same outcome (O). Despite the fact that cue X is always followed by the outcome, learners do not learn to associate cue X with the outcome. Learning about cue X is said to be blocked by the previous experience with cue A.

Blocking was first observed in conditioning experiments using non-human animals, and thus theories of conditioning have been developed to explain this phenomenon. According to one of the most famous proposals (Rescorla & Wagner, 1972), learning only occurs when outcomes are unexpected. The first phase of the blocking paradigm is designed so that learners come to expect the outcome when they encounter cue A. In the second phase, when cue A is paired with the novel cue X, the Rescorla-Wagner model suggests that learners will expect the same outcome that followed cue A. Given that this is exactly what follows the A-X pair, there is no surprise and thus no learning takes place. Thus, Rescorla-Wagner predicts that nothing is learned about the relationship between cue X and the outcome because there is never an opportunity to do so.

Attentional accounts (Mackintosh, 1975a; Kruschke, 2001), on the other hand, argue that blocking occurs because participants learn that cue X does not predict any change in the outcome (beyond cue A itself) and thus come to ignore cue X. Kruschke and Blair (2000) reported evidence consistent with this proposal. They first presented participants with a traditional blocking procedure. Subsequently, another novel cue was added to the A-X pair, which was then followed by a novel outcome. Their results showed that there was less learning about cue X even for the new, novel outcome; a finding inconsistent with Rescorla-Wagner. Such results suggest that the blocking procedure caused learners to ignore cue X, which subsequently prevented learners from learning about the relationship between cue X and the novel outcome. In addition, Mackintosh (1975b) conducted a blocking study in which rats first experienced a light repeatedly paired with a shock. Later, the rats received a single trial on which light and tone were presented together and followed by a shock. When tested, the rats' behavior indicated that they (weakly) expected the shock to follow the tone presented by itself. That is, the rats learned something about the novel cue after receiving only a single compound A-X trial. This suggests that, contrary to Rescorla-Wagner, learning in a blocking paradigm proceeds normally, at least during the first trial of the second phase, and that at least one compound trial may be required before blocking occurs.

Although the blocking paradigm involves a sequence of trial-by-trial presentations, previous studies have largely focused on learners' expectations at the end of the entire blocking procedure (the Mackintosh, 1975b study is a rare exception). Given that the prominent theories of blocking all make predictions about the trial-by-trial dynamics that underlie blocking, direct measurement of these dynamics seems to be an efficient way to distinguish between the competing theories.

In the present study, we investigate blocking by focusing on participants' beliefs about the novel cue X. According to the Rescorla-Wagner model, participants' expectations about the outcome the first time they encounter the A-X pair should be identical to their expectations at the end of the first phase when confronted with A alone. In contrast, the attentional accounts predict that, participants must gradually learn to ignore the novel cue X. Thus, the first time participants observe the A-X pair, their expectations about the outcome should be less certain than their expectations about cue A alone.

Experiment 1A: The First Phase 2 Trial

Method

Eight undergraduate students at Stony Brook University participated for partial course credit and were instructed to learn about how several medications were related to allergic reactions. To do so, participants were provided with a set of hypothetical medical records, each of which included information about what medication the patient had taken and whether the patient developed an allergic reaction. Figure 1 illustrates an example of a trial in Experiment 1A. The various medications were each represented by a different color. The allergic reaction was represented by a vertical thermometer-like bar that was either green and only partially filled (to represent no allergic reaction), or red and more than half-filled (to represent allergic reaction). Although the presence of an allergic reaction was always represented as a half-filled meter, participants were not explicitly told whether there could be different degrees of allergic reactions. This was done so as to allow participants to assume additivity (Beckers, De Houwer, Pineno, & Miller, 2005).

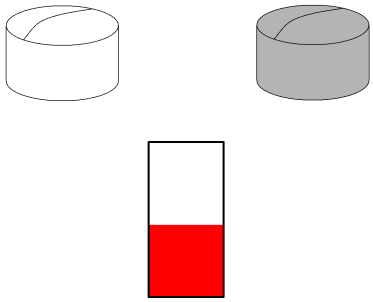

Figure 1. A example of a trial used in Experiment 1.

The trial sequence included two different types of trials (see Table 1). Medication A was always followed by an allergic reaction and medication Z was always followed by the absence of an allergic reaction. Each type of trial was repeated 8 times (randomly ordered). On each trial, the medication was presented on the upper half of the screen for 2000 ms. The effect was then presented on the bottom half of the screen for 1000 ms.

Table 1. Design of Experiment 1A and 1B.

	Phase 1	Phase2	Test
Exp 1A	A+/Z-		A→? AX→?
Exp 1B	A+/Z-		A→? X→?
	A+/Z-	AX+/CD+/EF-/YZ-	AX→?

At the end of the sequence, participants were presented with a new patient that had either taken medication A or the pair of medications, A and X. Participants were asked to judge how likely this new patient was to develop an allergic reaction on a scale from 1 (definitely will not develop an allergic reaction) to 8 (definitely will develop an allergic reaction).

Each participant completed two sequences, each of which was identical except for the colors of the medications used. After one sequence participants were asked to judge medication A and after the other they were asked to judge the pair of medication A and X. The order of the two sequences was counterbalanced across participants. Before starting, participants completed a brief sequence of 4 practice trials to familiarize them with the task.

Results

Participants' judgments were converted from the original scale into probabilities ranging from 0 to 1, where 0 represented the patient definitely not having an allergic reaction and 1 representing the patient definitely having an allergic reaction. The judged probability (shown in Figure 2) for the patient taking medications A and X (M = .67, SD = .21) was significantly lower than that for the patient taking only medication A (M = .92, SD = .18, $t(7) = 2.94$, $p < .05$). In addition, judgments for medication A ($t(7) = 6.78$, $p < .001$) and judgments for medication X ($t(7) = 2.31$, $p = .054$) were greater than chance.

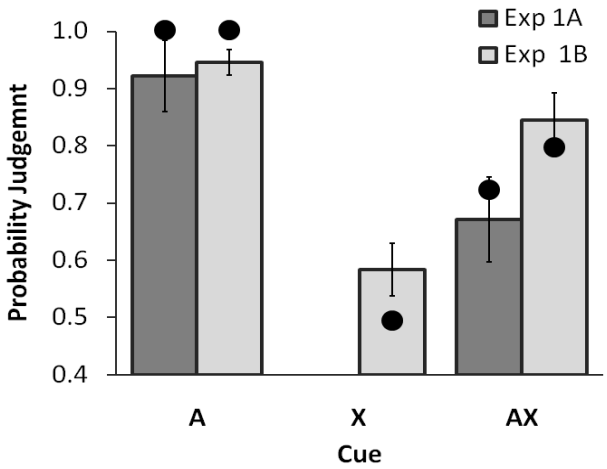

Figure 2. Results from Experiment 1A and 1B. Error bars represent standard error. Filled circles indicate the behavior predicted by our Bayesian account (see text for details).

Discussion

The results from Experiment 1A suggest that participants did not fully expect the allergic reaction to follow medications A and X even though they had a strong expectation that the effect would follow medication A by itself. In fact, participants' judgments about the AX pair were surprisingly close to chance. This lack of certainty is inconsistent with the predictions of the Recorla-Wagner model, which suggests that learners' expectations about the outcome when confronted with medications A and X should be identical to their expectations when confronted with medication A alone. Our results suggest that participants did not consider cue X to be redundant (as predicted by Rescorla-Wager). Instead, the addition of a novel cue appears to have added uncertainty to the outcome.

In contrast, our participants' behavior is quite consistent with the predictions of the attentional accounts (e.g., Kruschke, Mackintosh). Because nothing is known about the novel medication X, the attentional accounts assume that attention directed to this cue can only hurt the accuracy of learners' predictions. This decrement in accuracy then leads learners to ignore the novel cue on subsequent presentations so that their expectations are eventually based almost entirely on beliefs about medication A.

Experiment 1B: The Second Phase 2 Trial

Skeptics might argue that participants' uncertainty in Experiment 1A was due to task demands. Because participants first encounter medication X at the same time that they are asked to make the probability judgments, they may have over-emphasized the lack of information about medication X. For example, participants might have felt self-conscious expressing absolute certainty in the face of overtly novel information.

In Experiment 1B, participants again observed patients taking medication A and developing allergic reactions. Unlike Experiment 1A, however, participants in Experiment 1B then observed a single patient taking both medications A and X and developing an allergic reaction. The purpose of including this last patient taking both A and X was to essentially give participants direct feedback about what outcome would follow medications A and X as well as to justify high degrees of confidence in AX probability judgments. In addition, this variation will allow us to investigate whether one observation of AX is enough for learners to fully expect the outcome.

In addition, Experiment 1B asked participants to make probability judgments for a patient taking medication X alone. This allowed us to test an additional prediction of the attentional account. Specifically, these models (as well as Rescorla-Wagner) assume that learners believe novel cues to have a strength of zero (no influence). By eliciting probability judgments of medication X alone, we can investigate whether the uncertainty seen in Experiment 1A is due to the novel cue having no influence.

Method

Twenty-one undergraduate students at Stony Brook University participated for partial course credit. The same stimuli and procedure from Experiment 1A were used. The design was identical to Experiment 1A except for the following changes. First, six different types of trials were used (see Table 1). Second, there were 3 sequences of trials (run separately in counter-balanced order): two consisting of Phase 1 only (as in Experiment 1A), and the third including an additional 4 trials from Phase 2 of a traditional blocking design (see Table 1). In addition to one trial on which AX was followed by the outcome, another pair of medications (C and D) was followed by the allergic reaction and is the control condition traditionally employed in a blocking design (C and D were never presented alone). There were also two pairs of medications (E and F, Y and Z negative controls) followed by the absence of the allergic reaction to control for the base rate of the presence of the allergic reaction (at 50%). Without the negative controls, participants might believe that the allergic reaction would follow any medication. Each type of trial was only presented once in Phase 2. 3) Following the two Phase-1-only sequences, participants were asked to judge either medication A or medication X. At the end of the Phase-2 sequence, participants were asked to judge the A-X pair of medications.

Results

Participants' judgments were again converted into probabilities as in Experiment 1A. Participants again judged the probability of an allergic reaction to be lower for patients taking medications A and X (shown in Figure 2) ($M = .85$, $SD = .22$) than for those patients taking medication A alone ($M = .95$, $SD = .10$, $t(20) = 2.41$, $p = .03$). The judged probability of an allergic reaction was also lower for patients taking medication X alone ($M = .58$, $SD = .21$) than for patients taking medication A alone ($t(20) = 8.12$, $p < .001$) and for patients taking medications A and X ($t(20) = 3.92$, $p < .001$). Judgments for medication A and for the A-X pair were significantly greater than chance ($ps < .001$), whereas judgments for medication X were not ($t(20) = 1.81$, $p = .09$).

Discussion

The results in Experiment 1B demonstrated that participants did not expect the effect to follow the A-X pair of medications as much as they expected the effect to follow medication A alone even after observing the effect following A-X. This finding is, again, inconsistent with Rescorla-Wagner's explanation of blocking, which suggests that the effect should be fully expected at the beginning of Phase 2. This also suggests that task demands are unlikely to explain the results of Experiment 1A. These judgments are consistent with the attentional accounts because the certainty about the allergic reaction following medication A was decreased by the novel medication X.

However, participants' judgments about medication X alone suggest that when encountering a novel cue, participants expected the allergic reaction at approximately chance levels. This finding is inconsistent with attentional accounts which assume that novel cues have an initial strength of zero. Cues with strengths of zero produce outcomes 0% of the time whereas participants expected the novel cue to produce the effect 58% of the time. Thus, we are left with results that are decidedly inconsistent with Rescorla-Wagner and not entirely consistent with the attentional account. Here we put forth our own suggestion in an attempt to reconcile these findings and make additional, novel predictions.

A Bayesian Approach to Blocking

Our proposal relies on two important principles. First, we assume that beliefs about influence are represented as probability distributions, not point estimates. That is, each potential strength value has some probability of being the true value. This stands in sharp contrast to predominant

psychological models (e.g., Rescorla-Wagner, Kruschke, etc.). Second, following recent empirical work with both human (Beckers et al., 2005) and non-human animals (Beckers, Miller, De Houwer, & Urushihara, 2006), we assume causal additivity. That is, learners expect that two generative cues will jointly produce an effect that is greater than that produced by either alone. Although in our experiment there were only two levels for the effect (reaction/no reacction), the allergic reaction representation clearly implies a continuous outcome dimension required for additivity.

We assume that cues can be fully described by their influence, q (e.g., causal power, Cheng, 1997) which can range from 1.0 (strongly generative) to -1.0 (strongly preventative) with values of 0.0 indicating no influence. In addition, we assume that the only influence on the allergic reaction is exerted by the medications taken by that patient.

For simplicity, we only model the second phase of the blocking paradigm here. Thus, we assume that medications that were reliably paired with the allergic reaction during the first phase (e.g., medication A) have maximal generative influence (i.e., $P(E|A) = q_A = 1.0$). In contrast, we assume that learners are completely uncertain about the influence of the novel medication introduced in the second phase. Specifically, we represent uncertainty using the generic priors described by Lu, et al. (2008):

$$P(q_X) \propto \theta e^{-\alpha(1-q_X)} + (1-\theta)e^{-\alpha q_X} \quad (1)$$

To exhibit a preference for sparseness (i.e., that causes are unlikely to have weak influences), we set α=4. In addition, we assume that learners expect causes to be slightly more likely to exert generative influences than preventative influences (thus, θ is set to 0.65).

When asked to predict the probability of allergic reaction in a patient taking the known medication A and the novel medication X, we assume that cues combine their influence in the manner of a noisy-OR gate (Glymour, 1998). Thus, if medication X is generative (that is, if q_X>0),

$$P(E|A,X) = q_A + q_X - q_A q_X \quad (2)$$

If medication X is preventative (that is, if q_X<0),

$$P(E|A,X) = q_A(1 + q_X) \quad (3)$$

Because the strength of medication X is uncertain, we assume that participants' probability judgments reflect the following quantity:

$$P(E|A,X) = \int_{-1}^{1} P(E|q_A,q_X)P(q_X)dq_X \quad (4)$$

The result of this computation is shown in Figure 2. Note that the critical finding of Experiment 1A is mirrored in the behavior of the model. The model suggests that the pair of medications, A and X, is not guaranteed to produce an allergic reaction even though medication A alone is. This is because the uncertainty concerning the influence of medication X leaves open the possibility that medication X could prevent an allergic reaction that medication A alone would otherwise produce. This possibility means that the probability of the allergic reaction in the presence of medications A and X should necessarily be lower than in the presence of medication A alone.

In Experiment 1B, learners observed a single patient take medications A and X and develop an allergic reaction. Upon observing this patient, learners are assumed to update their beliefs by employing Bayes rule (e.g., $P(q_X|D) = P(q_X) \cdot P(D|q_X)$). When making this update, however, we assume that learners believe that the allergic reaction that has followed the pair of medications, A and X, is impossible if each medication is exerting a generative influence (if they did, a stronger reaction would be expected, Beckers et al., 2005). That is, the allergic reaction observed in these patients was produced by one of the two medications, but not both (rather than Equation 2):

$$P(E|A,X) = q_A(1-q_X) + q_X(1-q_A) \quad (5)$$

Once this update is made, however, participants' probability judgments are assumed to again be made according to Equation 2-4. Figure 2 shows the simulated results of Experiment 1B. Note that the model predicts less than certain probability judgments even after observing the pair of medications, A and B, followed by an allergic reaction. Nonetheless, the judged probability increases in light of this additional data.

This model can be contrasted with previous models of blocking. For example, associative models assume that novel cues have no influence (i.e., q=0). This is clearly inconsistent with Experiment 1B. When making predictions in the presence of a strong, generative cue (A) and a novel cue (X), Rescorla-Wagner assumes that strengths sum (e.g., 1+0). Thus, learners' predictions should reflect absolute certainty in the presence of the effect. The essential problem with this is the false precision embodied in the strength estimates attributed to novel cues. By endowing these novel cues with a strength of zero, it is inconceivable that these cues could be preventative. Under our account, learners are not nearly this confident in their prior beliefs.

The critical innovation of our model is leaving open the possibility that novel cues could be preventative. Indeed, if the novel cue introduced in the second phase was guaranteed to be generative, the model's predictions would be nearly indistinguishable from those of Rescorla-Wagner. Thus, the model predictions illustrated in Figure 2 rely strongly on learners' prior beliefs about what sorts of influences are permissible. Experiment 2 is designed to provide a strong test of this prediction.

Experiment 2: The Influence of Prior Beliefs

In Experiment 1, learners were not certain that the allergic reaction would follow the A-X medication pair, despite strong expectations that the allergic reaction would follow medication A on its own (Experiment 1A), and even after observing a patient taking medications A and X and developing a reaction (Experiment 1B). According to our model, this is because people believe that X may either be generative or preventative. For example, medications may terminate the presence of symptoms (a preventative influence), but may also produce side effects (a generative influence). According to our proposal, however, minimizing the possibility of preventative influence should

reduce the uncertainty observed in Experiment 1. To test this prediction, Experiment 2 manipulated the plausibility of preventative influence by using different sets of stimuli. The Prevention-Possible condition used the same medication stimuli used in Experiment 1 because medications can either prevent or cause allergic reactions (e.g., as a side effect). The Prevention-Unlikely condition instead used food stimuli because food may cause allergic reactions but it is much less likely that food can prevent allergic reactions.

Method

Forty-two undergraduate students at Stony Brook University participated for partial course credit and were randomly assigned to one of the two conditions (Prevention-Possible/Prevention-Unlikely), with 22 of them in the Prevention-Possible condition. The stimuli in the Prevention-Possible condition were identical to Experiment 1. In the Prevention-Unlikely condition, the medications were replaced with pictures of food (e.g., mushrooms, eggs, bread, cheese, and peanuts). The procedure and design were identical to Experiment 1A.

Results

Participants' judgments were again scaled to range between 0 and 1, and these judgments are shown in Figure 3. We conducted a 2 (condition: Prevention-Possible/Prevention-Unlikely) by 2 (judgment: cue A/ cues A and X) ANOVA with repeated measures on the latter factor. This analysis yielded significant main effects of judgment ($F(1, 40) = 9.82$, $p < .01$) and of condition ($F(1, 40) = 245.87$, $p < .001$). The interaction between condition and judgment was also significant ($F(1, 38) = 4.54$, $p < .05$). Participants in the Prevention-Possible condition judged the effect to be less likely to follow cue A and X than participants in the Prevention-Unlikely condition ($t(20) = 2.97$, $p < .01$). In contrast, the two groups did not differ on their judgments of cue A alone ($t(20) = .98$, $p = .34$). In the Prevention-Possible condition, the judged probability of the effect following the A-X pair was lower than the probability for cue A alone ($t(21) = 2.99$, $p < .01$). There was no such difference in the Prevention-Unlikely condition ($t(19) = 1.28$, $p = .22$).

Discussion

Experiment 2 showed that the possibility of a novel cue being preventative influenced participants' probability judgments. When novel cues could have a preventative effect, pairing a known, generative cue (cue A) with a novel cue (cue X) made it less probable that the outcome would follow the pair of cues. Consistent with Experiment 1, this is because participants allow for the possibility that the outcome would be prevented by the novel cue. However, when it was unlikely that novel cues could exert a preventative influence, learners believed that the A-X pair had to produce the same outcome because there was nothing to counteract the influence of cue A.

In line with our model, the results in Experiment 2 suggest that part of the reason that participants were unsure whether the effect would follow the A-X pair was that they believed that the novel cue could potentially exert a preventative influence. When this possibility was eliminated, learners' uncertainty was also reduced. This is also evidence that learners do not treat the novel cue X as redundant, but instead attempt to infer the role of cue X as best they can.

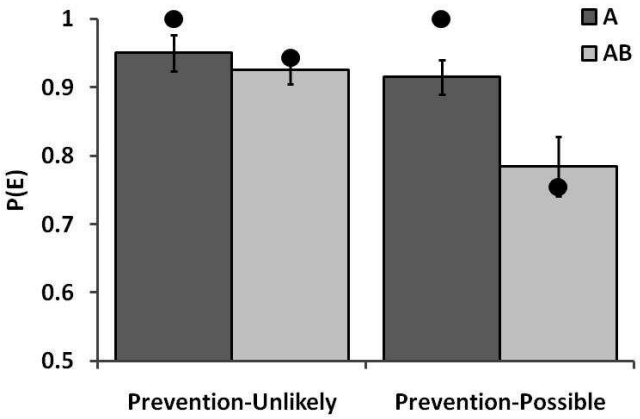

Figure 3. Results from Experiment 2. Error bars represent standard error. Filled circles indicate the behavior predicted by our Bayesian account. The Prevention-Possible condition was simulated as described in the text. The Prevention-Unlikely condition was simulated by modifying Eq. 4 such that $\theta = .9$.

General Discussion

Participants' beliefs about novel cues in a traditional blocking paradigm were evaluated in two experiments. In Experiment 1, participants were asked to judge how likely the outcome was to follow a novel cue (i.e., cue X) paired with a known, generative cue (i.e., cue A). Participants judged the outcome to be less likely to follow the combination of cues than to follow the known generative cue by itself. The results suggest that the novel cue introduced uncertainty to participants' expectations, which is inconsistent with the predictions of the Rescorla-Wagner model but consistent with attentional accounts of learning. However, participants' evaluations of the novel cue itself were inconsistent with the attentional accounts. We proposed a new, Bayesian account of blocking. Our model is able to account for the results in Experiment 1 and suggests that participants' prior beliefs about novel cues, particularly whether or not they may exert a preventative influence, play a critical role in blocking. This prediction was tested and confirmed in Experiment 2. We found that participants in the Prevention-Possible condition judged the pair of cues, A and X, to be less likely to produce the outcome than cue A alone, suggesting that the novel cue X added uncertainty to the situation. On the other hand, participants in the Prevention-Unlikely condition judged the pair of cue A and cue X to be just as likely to produce the outcome as cue A alone, suggesting that cue X did not add significant uncertainty.

Our results suggest that novel cues introduce uncertainty. Existing models assume that learners believe, with great confidence, that novel cues have no influence at all, ignoring potential uncertainty. Our model, on the other hand, assumes that people's beliefs about novel cues are highly uncertain and can be represented as a probability distribution over possible strength values. Such consideration of uncertainty is crucial in explaining participants' judgments in the current studies and how blocking is acquired. Moreover, uncertainty may help to understand learning in general, not only blocking.

Several researchers (e.g., De Houwer & Beckers, 2003; Lovibond, 2003) have argued that blocking may involve controlled, inferential processes in addition to simpler, associative processes. According to this account, learners assume additivity: that if two cues each produce a certain outcome, the pair of cues should produce a stronger outcome than either does individually. Based on this assumption, learners in a blocking procedure can infer that cue X has no influence on the outcome because, if it did, the outcome would have been greater than what cue A produced by itself. Beckers, et al. (2005) investigated whether additivity was critical to blocking by providing learners with pretraining which demonstrated the additivity either held or not before having participants complete a blocking procedure. When additivity was violated, participants showed weaker blocking than when additivity held, arguably because the former could not reason using the inferential procedure outlined above.

The model proposed here reflects the additivity assumption (Eq. 5) and can thus account for such results. Critically, the current model can do so without positing two separate process (one associative and one inferential). By assuming additivity, the current model gradually learns that the novel cue, X, neither prevents nor produces the outcome. If the additivity assumption did not hold, learners could only rule out preventative influences; they would be left uncertain as to whether the novel cue had strong or weak (or no) influence, resulting in far less blocking.

Although our model is able to account for many prior findings, there is data regarding other aspects of learning that the model cannot explain. For example, our model cannot immediately account for Kruschke and Blair's (2000) finding that subsequent learning about a blocked cue is inhibited. Also, Kruschke, Kappenman, and Hetrick (2005) utilized eye-tracking to demonstrate that participants learned to literally ignore the blocked cue. So far, our model does not predict anything about the acquisition of attentional shift during learning. These datasets suggest possible future directions for modification of the model proposed here.

References

Beckers, T., De Houwer, J., Pineno, O., & Miller, R. R. (2005). Outcome additivity and outcome maximality influence cue competition in human causal learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 238-249.

Beckers, T., Miller, R. R., De Houwer, J., & Urushihara, K. (2006). Reasoning rats: Forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference. *Journal of Experimental Psychology: General, 135*, 92-102.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367-405.

De Houwer, J., & Beckers, T. (2003). Secondary task difficulty modulates forward blocking in human contingency learning. *The Quarterly Journal of Experimental Psychology, 56*, 345-357.

Glymour, C. (1998). Learning Causes: Psychological Explanations of Causal Explanation. *Minds and Machines, 8*, 39-60.

Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279-296). New York: Appleton-Century-Crofts.

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

Kruschke, J. K., & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin and Review, 7*, 636-645.

Kruschke, J. K., Kappenman, E. S., & Hetrick, W. P. (2005). Eye gaze and individual differences consistent with learned attention in associate blocking and highlighting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 830-845.

Lovibond, P. F. (2003). Causal beliefs and conditioned responses: Retrospective revaluation induced by experience and by instruction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 97-106.

Lu, H., Yuille, A., Liljeholm, M., Cheng, P. W., Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review, 115*(4), 955-984.

Mackintosh, N. J. (1975a). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82*, 276–298.

Mackintosh, N. J. (1975b). Blocking of conditioned suppression: Role of the first compound trial. *Journal of Experimental Psychology: Animal Behavior Processes, 1*, 335-345.

Pineno, O., & Miller, R. R. (2007). Comparing associative, statistical, and inferential reasoning accounts of human contingency learning. *The Quarterly Journal of Experimental Psychology, 60*, 310-329.

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