Title
Is the blocking effect sensitive to causal model? It depends how you ask

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 40(0)

Authors
Don, Hilary J
Livesey, Evan J

Publication Date
2018

Is the blocking effect sensitive to causal model? It depends how you ask

Hilary J. Don (hdon7006@uni.sydney.edu.au)
School of Psychology, Brennan McCallum Building,
University of Sydney, NSW 2006 AUS

Evan J. Livesey (evan.livesey@sydney.edu.au)
School of Psychology, Brennan McCallum Building,
University of Sydney, NSW 2006 AUS

Abstract

Cue competition effects in human contingency learning appear to be sensitive to the causal nature of cue-outcome relationships. While blocking effects are reliably demonstrated in scenarios where cues are presented as causes of outcomes, several studies have failed to find blocking in scenarios where cues are presented as effects of outcomes, a finding that is typically taken as evidence for the involvement of controlled reasoning processes in cue competition. These studies typically measure blocking with continuous causal ratings about individual cues. Previous studies have found that sensitivity to causal model may depend on how the test question is phrased. In contrast, the current study tested the sensitivity of blocking to causal scenarios across different formats of the same test question. Participants completed a causal learning task with instructions suggesting either a predictive (i.e. cue causes outcome) or diagnostic (cue is caused by outcome) cue-outcome relationship. Participants were then asked about the likelihood of outcomes occurring by either giving a continuous rating of each outcome or a discrete choice about the most likely outcome. When measured by continuous ratings of individual cues, blocking was evident in predictive, but not diagnostic scenarios. However, when measured by discrete choice or using a compound *negation* test, blocking was robust and insensitive to causal scenario. The results suggest that contributions of predictive memory and causal reasoning to cue competition effects may depend substantially on the type of measure used.

Keywords: causal model; cue competition; blocking; test sensitivity

Introduction

Causal model theories propose that people use their knowledge about causal relationships to guide their learning and judgements of events (Waldmann & Holyoak, 1992; Rehder, 2003). While there is no question that people are capable of reasoning about causality, there is continuing debate regarding how readily and effectively people use this kind of knowledge in learning and cognitive tasks, and how much of performance is a result of simpler associative memory processes (e.g. Mitchell, De Houwer & Lovibond, 2009). This debate has often centred on the phenomenon of cue competition, in which the causal status of predictive cues is ambiguous because they appear simultaneously.

One well-known example of cue competition is the blocking effect (Kamin, 1969). In a contingency learning task, blocking is said to occur if judgements about the strength of a cue-outcome relationship are weaker when that cue is presented simultaneously with a second cue that is already established as a good predictor of the outcome (see Shanks, 2007 for a review). That is, if cue A is repeatedly paired with an outcome followed by compound cues AB paired with the same outcome, then outcome recall or judgements about the relationship between B and the outcome are reduced. The effect is usually observed by comparison to a control cue (e.g. D), which has been trained in compound with another cue that has not been pre-trained.

Some theories attribute blocking to a deficit in learning. According to most associative models, for instance, the outcome is already well predicted by A, and therefore there is little prediction error on AB trials to support learning about B (Rescorla & Wagner, 1972). In other words, learning about B is “blocked” by prior learning about the A association.

Other accounts of cue competition emphasise reasoning processes at test, based on a consideration of the causal structure of the cue-outcome relationship. Waldmann & Holyoak (1992) suggested that the causal structure imposed by the task should influence the observation of cue competition effects like blocking. In their study, participants in two groups responded to cue-outcome trials in the same manner, but the cover story describing the cues differed. In a predictive group, cues were described as possible *causes* of the outcome, and in a diagnostic group, cues were described as possible *effects* of the outcome (see Figure 1). As participants received the same presentation of cues and outcomes, associative theories should predict competition between cues in both groups. Instead, Waldmann & Holyoak (1992) found competition among cues when they were described as causes, but not when they were described as effects. This suggests that people use their understanding of causal structure during the task. In a predictive structure, if cue A is established as a cause of the outcome, the causal status of B is unclear, and is therefore a redundant cue. Reasoning about causation in this kind of causal model is therefore conducive to observing blocking effects. In comparison, in a diagnostic scenario, there is no reason to assume that B is not an effect of the outcome if cue A is already an established effect of the outcome. Cue B therefore still has useful diagnostic value, leading to a judgment for cue B that is just as strong as for control cues. In other words, multiple causes should compete for associative strength, but multiple effects should not (Waldmann & Holyoak, 1992).

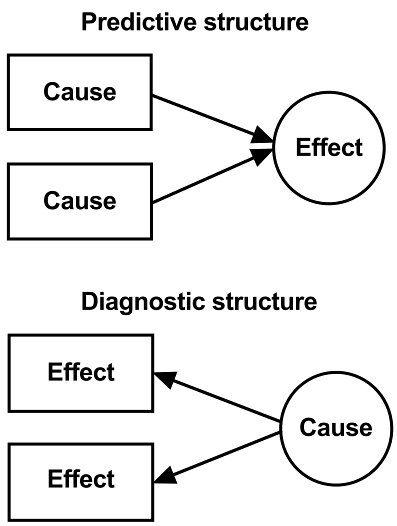

Figure 1: A schematic of a predictive (cues cause outcome) and diagnostic (cues are caused by outcome) model structure where rectangles represent cues and circles represent outcomes.

The demonstration of asymmetry in cue competition in predictive and diagnostic models has been taken as key evidence for causal model theories. Causal model effects have been replicated in several cue competition paradigms (e.g., Blanco, Baeyens & Beckers, 2014; Luque, Cobos, and López, 2008; Waldmann, 2000; 2001). However, other studies have failed to find this asymmetry (e.g. Arcediano et al., 2005; Shanks & Lopez, 1996), and several researchers have suggested causal model effects are dependent on particular task conditions. For example, one study found that causal model effects were only evident when the relevance of the causal structure was specifically emphasised to participants (Lopez, Cobos and Caño, 2005). Matute, Arcediano and Miller (1996) found that the wording of the test question was a critical factor. Asymmetry in cue competition was observed using causality based questions, but not using predictive or diagnostic value questions (which showed competition in both groups), or contiguity questions (which failed to show competition in either group).

As Thorwart and Livesey (2016) noted, studies investigating causal model effects typically measure cue competition using continuous causal ratings about individual test cues, and evidence for causal model effects have often emerged when ratings are close to ceiling. Therefore the absence of clear differences in ratings for blocked and control cues described as effects rather than causes may in part be due to differences in the way participants use the scale in different conditions. Indeed, Tangen and Allan (2004) found an influence of causal scenario when learning was measured by ratings during the test phase, but found no influence of causal scenario on trial-by-trial discrete predictions of the presence or absence of an outcome during training. They suggested that participants base their outcome predictions on associative strength between events, while ratings require participants to consider not only the association between events, but also the causal status of the cue, or the causal relationship between cues and outcomes. Therefore continuous ratings may invite reasoning about causality more than other measures, and the use of rating scales may be an important condition for these causal model effects to emerge. While prior studies have shown that causal model-effects depend on the nature of the test question (e.g. Matute et al., 1996), it is not yet known whether the format of the test is important, when the nature of the test question is held constant. Other cue competition effects, like highlighting and the inverse base-rate effect (Medin & Edelson, 1988; Kruschke, 1996; Don & Livesey, 2017) rely almost exclusively on discrete choice test measures and negation test trials. Negation tests are compound trials composed of two cues that were paired with different outcomes during training, in order to determine which cue has greater control over responding.

This experiment therefore aimed to assess the sensitivity of blocking to causal model across different test measures. Specifically, we compared causal model effects in continuous ratings and discrete outcome choice. Comparing differences in responses elicited by causal ratings and discrete-choice outcome recall questions has proven useful in teasing apart the influence of controlled and automatic processes in blocking and learned predictiveness (Mitchell et al., 2006; Shone, Harris & Livesey, 2015). However, unlike these studies, here we changed the format of the question but essentially asked for the same judgement. As blocking has previously been demonstrated using a negation-style test measure (Kruschke, Kappenman & Hetrick, 2005), we also included negation trials to assess whether they are also sensitive to causal model effects in the same way as single-cue test trials.

In this experiment, participants completed a modified allergist task with instructions indicating either a predictive or diagnostic relationship between cues and outcomes, manipulated between subjects, while maintaining the same cue and outcome stimuli between groups. In the predictive scenario, food cues were described as causes of different disease outcomes, identified by a name (e.g. Marshall's disease), and the task was to determine which foods were causing which disease. In the diagnostic scenario, food cues were chosen by different people and thus the outcome (the person to whom the meal belongs) effectively causes the cue. Foods were described as being part of that person's meal (e.g. Marshall's meal), and the task was to determine which meal belonged to which person. Learning for cues was then assessed with either continuous outcome likelihood ratings, or discrete outcome choice. There were two kinds of test trials; single-cue trials, and negation trials. We expect causal model to influence blocking on single-cue test trials when using continuous causal ratings. We are interested in whether this effect is also evident when using discrete choice or when testing blocking using a compound of cues with competing predictions.

Method

Participants

Eighty-eight students from the University of Sydney (61 female, mean age = 19.41, *SD* = 3.33) participated in the experiment in return for monetary compensation ($15/hr). Participants were randomly allocated to each condition (*n* = 22).

Apparatus and Stimuli

The experiment was programmed using PsychToolbox for Matlab (Kleiner et al., 2007). Participants were tested in individual cubicles up to four at a time. Experimental stimuli included 28 300 x 300 pixel images of foods, accompanied by a label in blue text. Food cues included *apple*, *avocado*, *bacon*, *banana*, *beef*, *bread*, *broccoli*, *carrots*, *cheese*, *cherries*, *chicken*, *chocolate*, *coffee*, *corn*, *eggs*, *fish*, *lemon, mango*, *milk*, *mushrooms*, *pasta*, *peach*, *peas*, *peanuts*, *pineapple*, *prawns*, *rice*, and *strawberries*. In both conditions, the outcomes were identified by one of four names, *Marshall*, *Scarlet*, *Florence*, and *Jakob*. These names indicated the names of different diseases in the predictive scenario (e.g. Scarlet's disease), or the person to whom the meal belonged in the diagnostic scenario (e.g. Scarlet's meal). For each participant, foods and names were randomly allocated to the cues and outcomes, respectively, shown in Table 1.

Design

The experiment was a 2 x 2 design with causal scenario (predictive vs. diagnostic) and test measure (ratings vs. discrete choice) as between-subject factors. Training contingencies are shown in Table 1. In Phase 1, four pre-trained cues were each paired with one of four possible outcomes. In the second phase, the pre-trained cues were presented in compound with a novel cue (the blocked cue), and paired with the same outcome. Four compound control cues were also presented in this phase. Compound and single cue filler trials were included in Phase 1 and Phase 2, respectively, so that participants experienced both individual and compound cues in each phase. Two different kinds of test trials were included. First, participants made outcome likelihood responses for each cue presented individually. Here, blocking would be indicated by reduced accuracy in responding for the correct outcome for the blocked cues compared to the control cues. Second, four negation trials were included, which were composed of one blocked cue and one control cue presented in compound, which had each been paired with a different outcome in Phase 2. These trials determine which cue has greater control of responding. Thus, blocking is indicated by greater responses for the outcome previously paired with the control cue than the outcome previously paired with the blocked cue.

Table 1: Experiment design.

TRAINING		TEST	
Phase 1	Phase 2	Trial type	
A_1 - 1	A_1B_1 - 1	Blocked	B_1 - B_4
A_2 - 2	A_2B_2 - 2	Control	C_1 - C_4
A_3 - 3	A_3B_3 - 3		D_1 - D_4
A_4 - 4	A_4B_4 - 4		
	C_1D_1 - 1		
	C_2D_2 - 2	Negation	B_1D_3, B_2D_4,
	C_3D_3 - 3		B_3D_2, B_4D_1
	C_4D_4 - 4		
E_1F_1- 1	G_1 - 1	Filler	E_1 - E_4
E_2F_2 - 2	G_2 - 2		F_1 - F_4
E_3F_3 - 3	G_3 - 3		G_1 - G_4
E_4F_4 - 4	G_4 - 4		

Note: Letters refer to individual food cues. Numbers refer to different outcomes.

Procedure

At the start of the experiment, participants were given instructions that outlined the causal scenario of the task. In the predictive condition, participants were given the following instructions:

> *In this task, you are asked to imagine that you work for the food safety industry. You have discovered that people are suffering from several different diseases after eating certain contaminated foods. Your job is to learn which foods are causing which disease by observing a number of meals that have been eaten, and the kind of disease that occurs.*

Participants in the diagnostic scenario were instructed:

> *In this task, you are asked to imagine that you work for a catering company that delivers meals to different people. You have been given a number of meals to deliver to several regular customers. However, the orders have all been mixed up. Your job is to learn which meal belongs to which person, from the food it contains.*

On each trial, two food cues appeared on the upper half of the screen, and participants were required to select one of the four outcomes presented below. When an outcome was selected, the options disappeared and corrective feedback was provided while the food cues remained on the screen. These trials occurred in an identical manner for both causal scenario conditions. There were five blocks of randomly presented training trials in each phase, with each trial type presented twice per block. The position of compound cues on the screen (left vs. right) was counterbalanced within each block. The transition between Phase 1 and Phase 2 occurred without a break.

Half the participants within each causal scenario group completed a rating test phase, while the other half completed a discrete choice test phase. Participants were instructed to use the knowledge they had gained so far to respond to trials in the test phase. In the ratings groups, participants were asked to either rate the likelihood of each

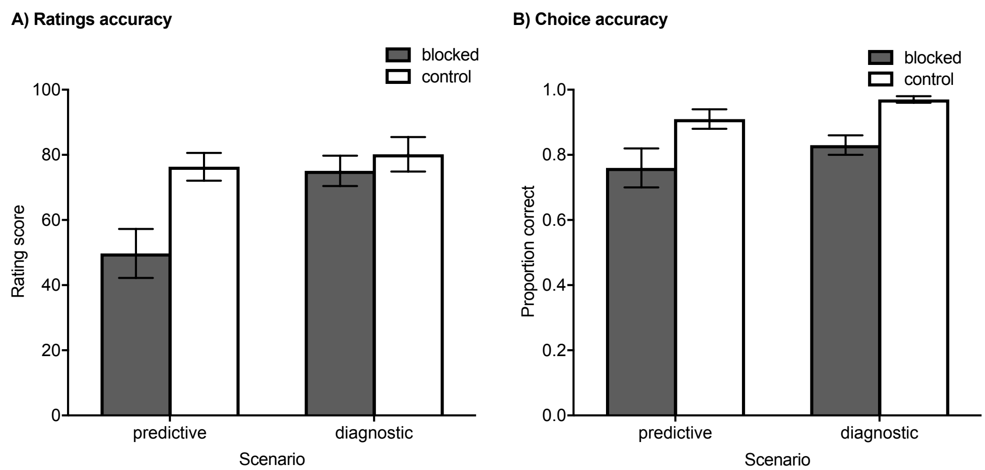

Figure 2: Accuracy for blocked and control cues presented individually at test. Panel A shows the average rating for correct outcome minus the average ratings for the incorrect outcomes in the rating groups. Panel B shows the proportion of correct outcome choice in the choice groups. Error bars show standard error of the mean.

disease, given the presented meals had been eaten, or to rate the likelihood that the meal belonged to each customer. On each trial, one or two food cues were presented on the upper half of the screen. Four rating scales were presented on the lower half of the screen, one for each outcome. Participants were required to make all four ratings, and could adjust their ratings before pressing the space bar to move to the next trial. In the choice conditions, participants were asked to select the disease they thought was most likely, given the meal had been eaten, or select which customer each meal was most likely to belong to. Outcomes appeared on the lower half of the screen in the same manner as training. Participants were also able to adjust their choice before pressing the space bar to move to the next trial. No feedback was provided in this phase. Test trials included all training cues presented individually and four negation trials, shown in random order.

Results

The measures in the rating and choice groups are on different scales, and were therefore analysed separately.

Single-cue trials

Figure 2 shows accuracy for blocked and control cues presented individually at test. To assess accuracy, a 2 x (2) mixed measures ANOVA was run with scenario as a between-subjects factor and cue (blocked vs. control) as a within-subjects factor for each test group separately.

Ratings. In the ratings group, accuracy was measured as the difference between the rating for the correct outcome and the average rating for the three incorrect outcomes. There was a significant main effect of cue $F(1,42) = 19.91$, $p < .001$, indicating a significant overall blocking effect. There was also a significant main effect of scenario $F(1,42) = 4.30$, $p = .044$, indicating greater overall accuracy in the diagnostic scenario. However this was qualified by a significant cue x scenario interaction $F(1,42) = 9.21$, $p = .004$, in which there was a stronger blocking effect in the predictive scenario than the diagnostic scenario. Further analysis of simple effects revealed a significant blocking effect in the predictive scenario $F(1,21) = 20.06$, $p < .001$, but not in the diagnostic scenario, $F(1,21) = 1.70$, $p = .207$. This replicates the classic sensitivity of blocking to causal model found using single cues and continuous ratings.

Choice. In the choice group, accuracy was measured as the proportion of correct outcome choice. There was a significant main effect of cue, indicating an overall blocking effect, $F(1,42) = 18.49$, $p < .001$, but no significant effect of scenario, $F(1,42) = 2.65$, $p = .111$, and critically, no interaction, $F < 1$.

Negation compound trials

Figure 3 shows responding to negation trials at test.

Ratings. In the ratings group, we compared participants' ratings for the outcome paired with the blocked cue to ratings for the outcome paired with the control cue. There was a significant main effect of cue, indicating overall higher ratings for the outcome paired with the control cue than the outcome paired with the blocked cue, $F(1,42) = 10.37$, $p = .002$. Although the blocking effect appears weaker in the diagnostic group, there were no significant effects of scenario, Fs < 1. There was a significant blocking effect in both the predictive, $F(1,21) = 13.57$, $p = .001$, and diagnostic, $F(1,21) = 9.32$, $p = .006$, groups.

Choice. In the choice group, we compared the proportion of choice of the outcome paired with the blocked cue and paired with the control cue. There was a significant main effect of cue, such that there was greater choice of the outcome paired with the control cue than the blocked cue, indicating a blocking effect, $F(1,42) = 25.40$, $p < .001$. There were no significant effects of scenario, highest $F(1,42) = 1.08$, $p = .305$.

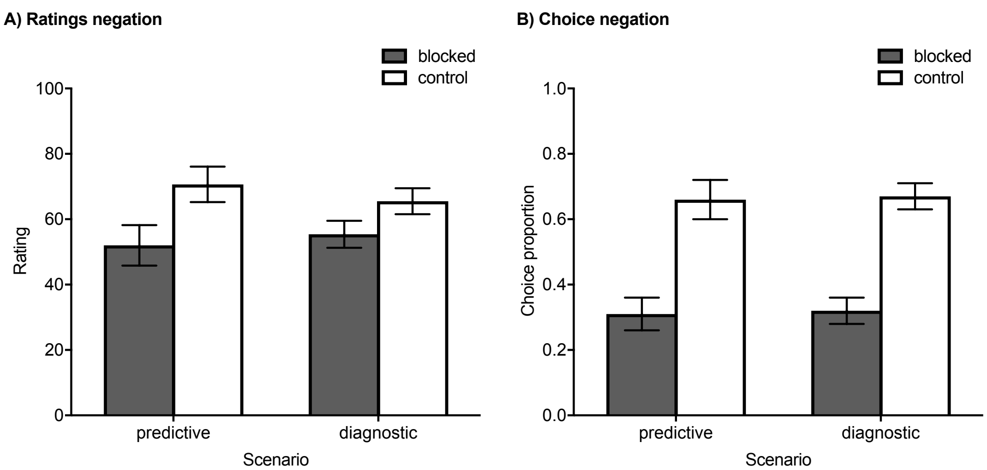

Figure 3: Responding to negation trials at test. Panel A shows ratings for the outcomes paired with the blocked and control cues. Panel B shows the proportion of choice of the outcome paired with the blocked and control cues. Error bars show standard error of the mean.

Discussion

In this experiment, evidence for causal model effects on blocking was dependent on the format of the test measure, and the kind of test trial. When measured by accuracy in continuous ratings using single-cue trials, blocking was evident in the predictive scenario, but not in the diagnostic scenario, replicating previous findings (Waldmann, 2000; Waldmann & Holyoak, 1992). However, when measured by accuracy in discrete choice, the strength of blocking did not differ between the two causal scenarios. Causal model-effects were also less evident when measured by negation test trials compared to single-cue trials.

Sensitivity to causal model manipulations has classically been interpreted as evidence that blocking and other cue competition effects are a result of reasoning at test rather than deficits in learning a connection between the cue and the outcome. For instance, the explanation favored by Matute et al. (1996) is that the blocked and control cues are learned about equally well, and participants use this information flexibly in the test phase, such that different test conditions influence the observation of cue competition effects. For example, participants may remember the outcome that was paired with a particular cue, but may give a lower rating if they infer that the cue does not cause the outcome. If this were the full story then *all* of the diagnostic conditions should, in principle, produce essentially the same result – equal judgements for the blocked and control cues (no blocking). Our results require a different explanation.

One hypothesis is that, regardless of the causal model, there are deficits in learning about the blocked cue, as predicted by associative learning models, and the strength of this learning serves as a basis for judgements at test under some circumstances. However, rational decisions based on the causal model rather than memory strength come to dominate test judgements when the conditions of the test are particularly conducive to reasoning, for instance when using single cues and a continuous rating. This is also consistent with research showing that people who score high on cognitive reflection tests are less likely to show blocking when inferential reasoning would predict there should be none (Livesey, Lee & Shone, 2013). The challenge for this account is to explain why single-cue ratings for the blocked cue under the diagnostic model are high, if learning about the blocked cue is supposedly deficient. One must assume that at least some learning about the blocked cue has occurred but that the strength of that learning is less relevant to strength of the judgement, since ratings for the blocked and control cues may be shifted up or down according to assumed causal status. This interpretation is consistent with Tangen & Allan's (2004) explanation for the presence of causal model effects in ratings at test, but not in trial-by-trial predictions during training.

A variation on this hypothesis is that causal models *do* affect learning about cues during training, and that a relatively weak blocking effect in the diagnostic groups is underestimated by single-cue ratings but overestimated by negation compound and forced choice conditions. Under a predictive model, reasoning about causes may lead to deliberate inattention to the blocked cue (see Mitchell et al., 2006). Thus although there is some learning deficit for the blocked cue in both model conditions, this deficit may be much stronger when reasoning about competing causes of a common effect. Similar to the first hypothesis, the challenge for this account is to explain why the choice and ratings measures are differentially sensitive to differences between the blocked and control cues.

The conditions that appear to be less susceptible to causal model effects each involve some form of response competition. In discrete choice, participants are required to choose between alternative options, and so choice of one outcome necessarily prevents choice of another. Negation trials, which pair two cues associated with conflicting outcomes, tacitly encourage an appraisal of which outcome is more likely, which may enhance subtle differences in prediction strength. In contrast, ratings are non-competitive, as each outcome can be rated independently, and multiple outcomes can be given high ratings. Each cue was only ever paired with one outcome, such that, provided at least some learning occurred, single-cue test trials would elicit relatively little response competition.

Perhaps then there is an influence of response competition on the way in which participants respond to test trials. Conditions that require participants to choose between options may be more sensitive to detecting small differences in associative strength, because the nature of the choice emphasizes this type of decision evidence. In contrast, single-cue ratings trials emphasize causal reasoning because competition among choices is not very salient. This would suggest a difference in the way the same evidence is translated into a decision in different test measures.

Previous research indicates that the sensitivity of cue competition effects to causal model depends on what is asked on test. This study shows that it also matters *how* you ask it. Causal model effects are more likely to be observed in continuous ratings for single-cue trials. Further research should assess whether causal models influence learning or performance to tease apart potential explanations for these results. Nevertheless, it is possible that both test sensitivity and differences in decision processes contribute. The results also suggest that researchers should consider the sensitivity of different test measures when interpreting the strength of learning and causal reasoning effects, and to consider what processes different test measures may reflect.

References

Arcediano, F., Matute, H., Escobar, M., & Miller, R. R. (2005). Competition between antecedent and between subsequent stimuli in causal judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 228.

Blanco, F., Baeyens, F., & Beckers, T. (2014). Blocking in human causal learning is affected by outcome assumptions manipulated through causal structure. *Learning & behavior*, *42*, 185-199.

Don, H. J., & Livesey, E. J. (2017). Effects of outcome and trial frequency on the inverse base-rate effect. *Memory & cognition*, *45*(3), 493-507.

Kamin, L.J. (1969). Selective association and conditioning. In N.J. Mackintosh & W.K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 42–64). Halifax: Dalhousie University Press.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*, 1

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

Livesey, E., Lee, J., & Shone, L. (2013, January). The relationship between blocking and inference in causal learning. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 35, No. 35).

López, F. J., Cobos, P. L., & Caño, A. (2005). Associative and causal reasoning accounts of causal induction: Symmetries and asymmetries in predictive and diagnostic inferences. *Memory & Cognition*, *33*, 1388-1398.

Luque, D., Cobos, P. L., & López, F. J. (2008). Interference between cues requires a causal scenario: Favorable evidence for causal reasoning models in learning processes. *Learning and Motivation*, *39*(3), 196-208.

Matute, H., Arcediano, F., & Miller, R. R. (1996). Test question modulates cue competition between causes and between effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22*, 182–196.

Mitchell, C. J., Lovibond, P. F., Minard, E., & Lavis, Y. (2006). Forward blocking in human learning sometimes reflects the failure to encode a cue–outcome relationship. *The Quarterly Journal of Experimental Psychology*, *59*, 830-844.

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioural and Brain Sciences, 32*, 183–98.

Rehder, B. (2003). Categorization as causal reasoning. *Cognitive Science*, *27*(5), 709-748.

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

Shanks, D. R., & López, F. J. (1996). Causal order does not affect cue selection in human associative learning. *Memory & Cognition*, 24, 511–522.

Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology*, *60*(3), 291-309.

Shone, L. T., Harris, I. M., & Livesey, E. J. (2015). Automaticity and cognitive control in the learned predictiveness effect. *Journal of Experimental Psychology: Animal Learning and Cognition*, *41*, 18.

Tangen, J. M., &Allan, L. G. (2004). Cue interaction and judgments of causality: Contributions of causal and associative processes. *Memory & Cognition, 32*, 107–124.

Thorwart, A., & Livesey, E. J. (2016). Three ways that non-associative knowledge may affect associative learning processes. *Frontiers in psychology*, *7*, 2024.

Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 53–76.

Waldmann, M. R. (2001). Predictive versus diagnostic causal learning: Evidence from an overshadowing paradigm. *Psychonomic Bulletin & Review, 8*, 600–608.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121*, 222–236.