Blocking and reciprocal blocking in predictive and causal learning

by

Duncan Yao Amegbletor

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
August 2, 2014

Keywords: associative learning, causal learning, predictors, blocking, propositional reasoning

Approved by

Francisco Arcediano, Chair, Assistant Professor of Psychology
Martha Escobar, Associate Professor of Psychology
Ana M. Franco-Watkins, Associate Professor of Psychology

Abstract

The blocking effect (Kamin, 1968) demonstrated that responding to a stimulus was attenuated (blocked) if it was presented in compound with a previously trained stimulus. Prior work with non-human animals, which examined the ability of the blocked stimulus to attenuate responding to the blocking stimulus (Arcediano, Escobar, & Miller, 2004), indicated that in some circumstances blocking is a reciprocal effect. That is, even though the blocking stimulus diminished responding to the blocked stimulus, the blocked stimulus in turn reduced responding to the blocking one. We sought to examine this reciprocal blocking effect in human learning. In a predictive learning task (Experiment 1) and in a causal learning task (Experiment 2) both the blocking and the blocked stimuli were found to mutually compete with each other. This observation seems to indicate that the blocking stimulus loses behavioral control because of its additional training with the blocked stimulus. This finding is not only at odds with the assumptions of traditional and modern associative learning models, but also with recent models of causal learning based on inferential reasoning. However, this finding seems to be consistent with the assumptions of the comparator hypothesis (Miller & Matzel, 1988).

Table of Contents

List of Tables

List of Figures

Blocking and reciprocal blocking in predictive and causal learning

In the fourth century B.C.E., the Greek philosopher Aristotle proposed the basic laws of association: rules which dictate how we acquire knowledge (e.g., Barnes, 1991). Arguably, the best known of said rules is the law of contiguity: the closer two events are in space and time, the more strongly associated they become. The idea of spatiotemporal contiguity was considered to be both necessary and sufficient for learning to occur in most known views of learning (Pierce & Epling, 1995). As contiguity was considered to be the key to learning associations between events, it is no wonder that experimental psychologists in the first half of the 20th century incorporated this idea into the theories of the day (e.g., Bush & Mosteller, 1951; Guthrie, 1930). However, this millennia-long dominance of contiguity came to an end in the second half of the 20th century when several independent findings began to reveal that contiguity alone may not be sufficient for learning to occur (e.g., Kamin, 1968, 1969; Rescorla, 1968; Wagner, 1969).

Blocking (Kamin, 1968, 1969) is a learning phenomenon in which responding to a target stimulus is attenuated by prior experience with a companion stimulus (aka the blocking stimulus). Typically, blocking proceeds in two phases. Phase 1 consists of pairings of the blocking stimulus, A, with an outcome, O. In the second phase, Stimulus A is presented in compound with a new stimulus, the target, X, and again followed by the same outcome. In Kamin's landmark experiment (see Table 1), rats in the experimental group were presented with a white noise, A, followed by an electric shock, O. After this initial training, the rats experienced the white noise presented in compound with light onset, X, and again followed by shock. Test trials to the presence of the light alone (X) indicated that the rats displayed little behavioral suppression (an index of conditioned fear) to the light, compared to a control group of rats which never received the white noise alone trials during Phase 1. Kamin interpreted the experimental

group's reduced responding to the target stimulus to mean that the previously trained white noise blocked the learning of the added stimulus, the light.

Table 1

Blocking Design

Group	Phase 1	Phase 2	Test
Experimental	A→ O	A+X→ O	X?
Control		A+X→ O	X?

One of the more important implications of the blocking effect is that it serves as a demonstration that contiguity alone is not sufficient for learning to occur. The contiguity between the light and the shock is the same in both groups of rats, yet one group seemingly shows decremented learning about the association between the light and the shock. As such, Kamin's blocking effect was integral in the development of contemporary associative theories of learning.

Older, contiguity-based models (e.g., Bush & Mosteller, 1951) could not explain blocking and other similar effects which arise from training stimuli in compound. A new class of models (e.g., Mackintosh, 1975; Rescorla & Wagner, 1972; Wagner, 1981) focused on the amount of information provided by a stimulus as the key determinant of learning rather than contiguity. This idea is often called the *information hypothesis*. Learning occurs because a stimulus non-redundantly predicts an outcome; redundant or non-informative stimuli are not learned. In blocking for example, subjects fail to learn about X because it provides no unique

information about outcome occurrence. Stimulus A already serves as an excellent predictor of the outcome, rendering X redundant.

The Rescorla-Wagner model (Rescorla & Wagner, 1972) along with most of the models of the time (and most subsequent models) is based on the information hypothesis. Arguably, the Rescorla-Wagner model has been the most relevant associative model, due to its influence in the field and contribution to the development of new research. The Rescorla-Wagner model can be formalized using the following equations:

$$\Delta V_A^i = \alpha_A \beta_O \left(\lambda - \sum V\right)$$

$$V_A^i = V_A^{i-1} + \Delta V_A^i$$

ΔV^i_A represents the change in associative strength of stimulus A which occurs on a given trial (i). V^i_A represents the current associative strength of A on a given trial; it is the sum of the associative strength of the target CS on all previous trials (V^{i-1}_A), and the change in associative strength which occurred on the current trial (ΔV^i_A). α and β are learning rate parameters for the CS (or stimulus) and US (or outcome), respectively. These are constrained by the salience and associability of the stimuli, and affect the speed and effectiveness of learning. When the CS is absent, α takes a value of zero; this is not the case for β when the US is absent. λ is the total amount of associative strength that the US can support (the asymptote of learning). This constrains how much an organism can learn about a CS. ΣV is the total amount of associative strength currently allocated to all stimuli that have been presented with the US (including the target stimulus). Note that λ minus ΣV is the amount of associative strength which can be allocated in that CS-US training trial. This discrepancy between λ and ΣV is the amount of prediction error. A larger discrepancy indicates there is still much to learn about the US. A US which is surprising -- i.e., not well predicted -- is highly associable, meaning that learning occurs

easily and quickly. In blocking for example, on the very first A-O trial during Phase 1, there is a large discrepancy between λ and ΣV. Subsequent A-O trials increase V_A, while the amount of error decreases; i.e., the value of ΣV (which is currently affected only by V_A) approaches λ. During Phase 2, when A and X are paired with the outcome, O, most of the associative strength has already been allocated to A leaving little for X. As such, learning about X (V_X) remains weak, as compared to a control group in which previous A-O learning did not occur.

While Rescorla-Wagner and other traditional associative models were gaining momentum in the non-human animal learning literature, the predominant view of learning in humans was that contingency judgments closely tracked the strength of a contingency given by statistical rules such as ΔP (Allan, 1980; Tangen & Allan, 2003). ΔP is an index of the strength of contingency between two stimuli, based on the difference between two probabilities. The probability of the outcome given the stimulus or cue, X, and the probability of the outcome given the absence of the cue, ~X. During contingency learning, frequency information about the co-occurrence of cue and outcome is acquired. This information can be represented in a 2x2 contingency matrix (see Table 2).

Table 2

The ΔP 2x2 Contingency Matrix

	Outcome	
Cue	O	~O
X	a	b
~X	c	d

Cell a contains all trials where both the cue and outcome are present, cell b contains all trials where the cue is present and the outcome is absent, cell c contains all trials where the outcome is present but the cue is absent, and cell d contains trials where both the cue and

outcome are absent. Adding all four cells together gives you the total number of "trials" experienced. The following equation is how ΔP is computed from frequency information in the 2x2 matrix:

$$\Delta P = P(O|X) - P(O|{\sim}X) = \frac{a}{a+b} - \frac{c}{c+d}$$

The value of ΔP can signify positive, negative and neutral contingencies. As it is based on the difference between conditional probabilities, the number itself is constrained between positive and negative 1. The strength of the ΔP rule is in its ability to closely track participants' contingency judgments. Despite this major strength, the basic ΔP rule has difficulty accounting for many stimulus competition effects as it lacks a mechanism for taking stimuli paired with the target into account.

Up until the 1980s, the work that had been done with stimulus competition effect in non-human animals had little impact on how human learning was studied. It was then, that the field of associative learning and human contingency learning began examining interactions between stimuli trained together. In this field, the CS and US in non-human animal learning were considered to be roughly analogous to cues and outcomes in human learning. Dickinson, Shanks, and Evenden (1984) were the first to report blocking with human participants in a task measuring contingency judgments. Dickinson et al. used a video game-like task in which participants had to judge the effectiveness of specific cues with respect to a common outcome. In the task, participants attempted to destroy a tank by firing a shell at it. However, the destruction of the tank could also be caused by the minefield the tank drove through on each trial. Participants saw three sets of trials: one blocking set, and two control sets. In Phase 1of the blocking set, the only possible cause of destruction was the minefield. In the second phase, participants were able to fire shells at tanks; thus both mines and participant-fired shells were potential causes of

destruction. In both control sets, there was only one phase, where both shells and mines were equally likely to be potential causes of destruction (i.e., there was no previous experience with the mines as the only possible cause of tank destruction). At test, participants rated the effectiveness of the shells, that is, the contingency between shell firing and tank destruction. Participants rated the effectiveness of the shells in the blocking set (where the mines were initially the sole cause of destruction) lower than the effectiveness of the shells in the control sets, indicating that this manipulation resulted in blocking. Additionally, the contingency judgments of the shells closely tracked the associative strength predicted by the Rescorla-Wagner model. Other studies (e.g., Allan, 1993; Arcediano, Matute, & Miller, 1997; Arcediano, Escobar, & Matute, 2001; Chapman, 1991; Ghirlanda, 2005; López, Shanks, Almaraz, & Fernández, 1998; Wasserman & Neunaber, 1986) demonstrated marked similarities between contingency learning in humans and non-human animal learning. Furthermore, just as contingency judgments seem to be subject to the same manipulations which affect animal learning, phenomena first reported in human contingency learning were later confirmed with animal learning (e.g., Beckers, Miller, De Houwer, & Urushihara, 2006; Miller & Matute, 1996; Wheeler, Beckers, & Miller, 2008).

Despite the massive surge in the use of associative frameworks to examine human learning effects, the normative approach was not abandoned. Spellman (1996; see also Tangen & Allan, 2003) sought to ameliorate one of the major weaknesses of the normative approach by examining how a conditional ΔP rule could explain effects involving interactions between two stimuli, which could not be explained by the standard ΔP (e.g., overshadowing; Spellman, 1996). The conditional ΔP rule allows for stimulus interactions by basing the value of ΔP on other stimuli present during learning (e.g., ΔP of X in the presence of A, or $\Delta P_{X|A}$). The 2x2 matrix is

expanded into a 4x2 contingency matrix (Table 3) to take into account all possible combinations of the two stimuli and the outcome.

Table 3

The Conditional ΔP 4x2 Contingency Matrix

	Outcome	
Cues	O	~O
X A	a	b
X ~A	c	d
~X A	e	f
~X ~A	g	h

The following equations detail how the conditional ΔP of X in the presence and absence of A is determined:

$$\Delta P_{X|A} = P(O|XA) - P(O|{\sim}XA) = \frac{a}{a+b} - \frac{e}{e+f}$$

$$\Delta P_{X|{\sim}A} = P(O|X{\sim}A) - P(O|{\sim}X{\sim}A) = \frac{c}{c+d} - \frac{g}{g+h}$$

In the experimental group of blocking, X has a low conditional ΔP value due to the strong A-outcome contingency. There are many trials (i.e., all Phase 1 trials) where the outcome occurs with A, but not with X, thus reducing X's conditional contingency with the outcome. In the control condition, the outcome never occurs without the compound of A and X, meaning that they both have identical contingencies with the outcome.

Unlike normative and traditional associative models, one of the key assertions of the comparator hypothesis is that learning is based on contiguity rather than the information provided by a stimulus (Miller & Matzel, 1988; Miller & Schachtman, 1985). The comparator hypothesis separates the process of acquisition from the generation of a response (whether the response in question is a judgment of contingency, causal attribution, or a conditioned response). The expression of an association – i.e., responding– is dependent on other cues or causes (called comparator stimuli) which were present along with the target stimulus. When an opportunity to respond occurs, the strength of the association between the target stimulus and the outcome is compared to the strength of the association between comparator stimuli and the outcome. In blocking (see Table 1), the experimental group's Phase 1 training of A with the outcome results in an association between A and the outcome, and compound training during Phase 2 results in additional learning of the A-outcome association, as well as X-outcome and X-A associations. When the target stimulus X is presented at test, the X-outcome association acquired during Phase 2 is compared with the A-outcome association acquired during Phases 1 and 2; this comparison is facilitated by the X-A association acquired during Phase 2. Responding to X diminishes as a function of the strength of both the X-A and A-outcome associations. In the control group, the A-outcome association should have a lesser impact on responding to X because A was not associated with the outcome during Phase 1.Thus, we should expect weaker responding to X in the experimental group than the control group. Key evidence in support of the comparator hypothesis has come from post-training changes in the value of the comparator cue (e.g. extinction of the comparator cue [Arcediano, Escobar, & Matute, 2001; Blaisdell, Gunther, & Miller, 1999; Cole, Barnet, & Miller, 1995; Miller, Barnet, & Grahame, 1992] and exposure to unpredicted outcome presentations [Balaz, Gutsin, Cacheiro, & Miller, 1982; Miller, Hallam, &

Grahame, 1990]). These manipulations (often called retrospective revaluation effects) result in the previously masked responding to the target becoming evident. For example, recovery from blocking (Arcediano et al., 2001, with humans; Blaisdell et al., 1999, with rats) demonstrates that extinguishing A after both phases of blocking have been completed enhances responding to X.

One of the key weaknesses of traditional associative models (e.g., Mackintosh, 1975; Rescorla & Wagner, 1972; Wagner, 1981) is that they cannot explain retrospective revaluation effects. With retrospective revaluation, the associative value of an absent cue changes, due to a cue which was paired in compound with the absent target being presented either with or without the outcome. Van Hamme and Wasserman (1994), modified the Rescorla-Wagner model (1972), allowing prediction of retrospective revaluation effects, and allowing the associative approach to contend with the comparator hypothesis on this front. In the original Rescorla-Wagner model, an absent cue has an α of zero; this results in no change in associative strength on a trial where the cue is absent. When the outcome is absent, β either maintains its current value or takes a lesser, non-zero value; in either case the associative strength of the cue declines. Van Hamme and Wasserman (1994) proposed that α should take a negative value on trials where the cue in question is absent. Thus, when both the cue and outcome are absent, the negative $\alpha\beta$ interacts with the negative prediction error term (λ-ΣV) and results in a gain in associative strength for the absent cue. In other words, having trials of A- after AX+ trials would decrease the associative strength of A, while increasing the associative strength of X. Similarly, having A+ trials after AX+ trials would increase the associative strength of A, while reducing the associative strength of X (negative $\alpha\beta$, positive prediction error).

Dickinson and Burke (1996) continued in the direction of Van Hamme and Wasserman (1994) with a proposal of how the SOP model (Wagner, 1981) could be modified to explain

retrospective revaluation. Simply put, in the base SOP model, there are two levels of activation A1 and A2. A cue is in A1 activation when it is directly presented. A2 activation of a cue occurs when it is indirectly activated by presentation of a previously paired cue or outcome. Changes in cue-outcome associative strength only occur when both the cue and outcome are presented (i.e., both in A1 activation). Dickinson and Burke (1996) proposed that a) associative strength can increase when both cues are in A1 *or* both in A2, and b) if two cues are in opposing states (i.e., a cue is in A1 and the outcome is in A2), associative strength decreases. Having A+ training after AX+ trials will put X and the outcome in opposing states of activation, lowering the associative strength of X; similarly, having A- training after AX+ trials will put both X and the outcome in an A2 state, increasing the associative strength of X with the outcome.

The use of associative and statistical accounts in human learning has recently been challenged by a newer, more cognitive approach. Although the study of cognitive influences on learning is not a new one, this approach in the learning literature can be traced back to the early 2000s (De Houwer, Beckers, & Glautier, 2002; Lovibond, Been, Mitchell, Bouton, & Frohardt, 2003), and includes both descriptive and quantitative explanations. This view is referred to by many names in the literature: frequently called the propositional or inferential reasoning approach (De Houwer, 2009; Lovibond et al., 2003; Pineño & Miller, 2007; Shanks, 2010) and more recently as the belief distribution approach (Carroll, Cheng, & Lu, 2013) as quantitative explanations have become closely incorporated into this view. In this framework, stimulus competition arises from controlled, effortful decision making about the effectiveness of stimuli in *producing* outcomes. Inference making is supported by expectations that humans have about the nature of causality, and our prior knowledge of causal relationships. Data violating our knowledge of causal structure produce many of the effects seen in human causal learning.

Additivity of causes is one such assumption about causal structure; we have specific expectations of how events interact within the environment. Causes are a unique form of event; they directly impact outcomes. That is, changing the value of the cause changes the value of the outcome. For example, changing the amount of allergen present affects the intensity of the allergic reaction. Causes which occur together should also increase the strength or the probability of the occurrence of the outcome more so than when occurring alone. If adding additional causes does not impact the outcome, then one can infer that the added cause is ineffective (i.e., it is not a cause).

De Houwer et al. (2002), and Lovibond et al. (2003) nearly simultaneously published findings which indicated that blocking relies on inference making, these and similar descriptive accounts of stimulus competition are alternately referred to as the inferential or propositional reasoning account in the literature (e.g., Pineño & Miller, 2007). In this view, blocking only occurs when one can infer that the added cause is ineffective. Rather than a deficit of learning about the added cause due to its lack of informative value, blocking arises from the knowledge that the added cue is not an effective cause. In contrast, blocking should not occur with predictors: stimuli which serve as indicators, but not causes, of the outcome; predictors do not impact the outcome magnitude. For example, a weather forecast is a predictor of future weather patterns, but increasing the urgency of the forecast, or adding additional weather forecasts will not impact the magnitude of the meteorological event. Pineño and Miller (2007) describe the chain of reasoning used in blocking as akin to a modus tollens logical argument:

If stimulus X modifies the outcome, then it is an effective cause.

X does not modify the outcome.

Therefore, it is not an effective cause.

Associative models based on the information hypothesis use informative value to provide increments in the associative strength between a cue and outcome, by means of error-reduction. In the propositional approach, a stimulus provides novel information through the ability to change outcome magnitude; stimuli which leave outcome magnitude unchanged fail to provide novel information about the outcome and are not considered causes.

In recent years, more and more investigation has been directed towards models of learning based on Bayes rule (henceforth Bayesian models). Although they share many of the same assumptions, it is only recently that Bayesian models have been discussed as quantitative propositional models (Carroll, Cheng, & Lu, 2013; De Houwer, 2009). Just like their descriptive counterparts based on propositional logic, Bayesian models incorporate prior beliefs about causality and previously learned causal relationships. In many models attempting to represent causality, learners entertain multiple hypotheses (e.g. Fernbach, Darlow, & Sloman, 2011); that is, multiple potential causal models which could produce the observed outcome. Bayesian models also incorporate alternative hypotheses, each hypothesis has a degree of uncertainty as well as a probabilistic likelihood (Holyoak & Cheng, 2011). For example if a patient develops hives, their doctor will have a number of causes or combinations of causes in mind; each individual cause or combination constitutes a hypothesis.

Another constant among the numerous Bayesian models is the use of Bayes rule to update beliefs after observing data. The probability, p(h|d), of a hypothesis (specific combination of cue weights and values) given specific data (observations) is calculated as follows.

$$p(h|d) = \frac{p(d|h)p(h)}{p(d)}$$

P(h) is the probability of a hypothesis before observing any data, it is often called the prior distribution or prior belief. P(d|h) is the likelihood function, this is the probability of the data

occurring if the hypothesis is true. P(d) is the likelihood of the observed data, this probability gives the amount of evidence for the data.

Despite the commonality of Bayes rule as the foundation of these models, the sets of assumptions can differ wildly. Take, for example, the Kalman Filter; this specific model is a Bayesian adaptation of the assumptions of associative models (Kruschke, 2008; see Rescorla & Wagner, 1972). One of the very early Bayesian formulations which was applied to learning used the idea of causal support (Griffiths & Tenenbaum, 2005). The Noisy-Logic Gate (e.g. Carroll et al., 2013) incorporates assumptions from theories of causal learning, such as Power-PC (Cheng, 1997), and causal support (Griffiths & Tenenbaum, 2005). Additionally, recent work has attempted to incorporate ideas such as trial order (Courville, Daw, & Touretzky, 2006; Daw, Courville, & Dayan, 2007) and sample/set size (Lu, Yuille, Liljeholm, Cheng, & Holyoak, 2008), which have long been problematic for Bayesian and other normative models.

The flexibility of Bayesian models allow them to explain a very large number of effects, including retrospective revaluation (Carroll et al., 2013; Kruschke, 2008; Lu, Rojas, Beckers, & Yuille, 2008), overshadowing / highlighting (Daw et al., 2007), and of particular relevance for our purposes, blocking (Lu et al., 2008). At the beginning of blocking training, Cue A could be a cause of the outcome or have no effect on the outcome (disregarding preventative causes). Phase 1 training centers the model around A being an effective cause of the outcome. In Phase 2, X paired with A does nothing to shift the belief distribution. Blocking is simple to explain in the Bayesian approach as there is no ambiguity about the role of each cause in producing the outcome. In contrast, consider backwards blocking (e.g. Carroll et al., 2013; Kruschke, 2008) where the order of Phases 1 and 2 is reversed (AX-O trials, followed by A-O). Phase 1 of backward blocking is ambiguous, as both A and X are equally likely to be causes of the outcome.

The presentation of a compound (rather than a singular cue) during Phase 1 increases the number of potentially valid beliefs: A could be the only valid cause, A and X could both be valid causes, or only X could be a valid cause. Phase 2 training of A alone shifts the belief distributions for both A and X, and adds support for the belief that A is the cause of the outcome.

All of the previously mentioned views of learning are capable of explaining learning effects (such as blocking) equally well. To continue expanding the work done in the field it becomes necessary to focus on the dissimilarities between the various models as a means to discriminate between models as well as improve existing ones. One of these unique predictions is that of reciprocal stimulus competition.

In addition to blocking and other basic learning effects, the comparator hypothesis predicts, unlike the previously mentioned views of learning, that stimulus competition should be reciprocal. For example, in the experimental group, the blocking stimulus, A, should elicit lower responding than a stimulus which has similar Phase 1 training, but lacks Phase 2 compound training with the blocked stimulus (X). If A is evaluated at test, its responding should depend not only on the strength of the A-outcome association, but also on the strength of the X-A and X-outcome associations. Thus, the process of competition during blocking should be reciprocal: In the same way that A blocks responding to X, X should diminish responding to A. In other words, X affects responding to A, just as A affects responding to X. There is already some support in the literature for such an effect. There is evidence that stimuli with low validity (weakly correlated with reinforcement) can acquire control over responding, but will also lose said control if a better predictor of reinforcement is available (e.g., relative stimulus validity; Wagner, Logan, & Haberlandt, 1968). Hall, Mackintosh, Goodall, and Dal Martello (1977) sought to extend the loss

of stimulus control found with stimulus validity to stimulus salience. In their Experiment 3, they compared four groups of rats (see Table 4) using a blocking design.

Table 4

Groups in Hall et al. (1977)

Group	Phase 1	Phase 2
t-tL	tone→Shock	Tone+LIGHT→Shock
t-t	tone→Shock	tone→Shock
t-L	tone→Shock	LIGHT→Shock
tL	___	Tone+LIGHT→Shock

Note. *lowercase* indicates low salience. *ALLCAPS* indicates high salience.

In Group t-tL, a faint tone was followed by an US (electric foot shock) in Phase 1; then it was presented in compound with a salient light, and again followed by the US in Phase 2. Group t-t was trained on the faint tone followed by the US in both phases. Group t-L was trained on faint tone – US in Phase 1 and the salient light alone – US in Phase 2. Finally Group tL was trained on the faint tone-salient light compound followed by the US in Phase 2. All groups were tested on both the light and the tone. Results indicated weaker responding to (i.e., fear of) the light in Group t-tL as compared to Groups t-L and tL. This was interpreted as evidence that the light was blocked by the less salient tone in Group t-tL. Additionally, a comparison of Group t-tL versus Groups t-L and t-t revealed weakened responding to the tone in Group t-tL as a consequence of the tone being paired with the more intense light; fear of the tone was greater in

Groups t-L and t-t than in Group t-tL. In other words, although the tone was a strong enough predictor to reduce responding to the more intense light (blocking) in Group t-tL, compound pairing with the light also attenuated responding to the tone; responding to the blocking stimulus, the tone, was weakened by the blocked stimulus, the light. Thus, Hall et al. (1977) were able to replicate the loss of control to a low validity stimulus, first shown by Wagner et al. (1968), by manipulating stimulus intensity or saliency instead of cue-outcome contingency, and point out the inability of the Rescorla-Wagner model to account for this result. Hall et al. (1977) explained this effect within an attentional version of the information hypothesis. A low saliency stimulus can acquire behavioral control if it is the sole predictor of reinforcement, but being trained in compound with a more salient stimulus causes the low salience stimulus to lose its acquired behavioral control. Arcediano, Escobar, and Miller (2004) further examined this loss of control due to compound training, but controlled for salience. In addition to the high-salience stimulus condition from Hall et al. (1977), Arcediano et al. (2004) also included a condition where all relevant stimuli were of similar saliency. The results indicated that responding to the blocking stimulus (A) was reduced in both the high and similar salience groups, compared to control groups where A was never presented in a stimulus compound. Being presented in compound with X lowered responding to A regardless of whether X was more salient (although a more salient X resulted in a larger response decrement to A, as should be expected by a comparator mechanism). In contrast, the account proposed by Hall et al. (1977) only predicts a response decrement in the blocking stimulus when it is less salient than the blocked stimulus. They interpreted this finding as being consistent solely with the comparator hypothesis, and incompatible with models based on the information hypothesis, and with normative views of learning. Although the revisions proposed by Van Hamme and Wasserman (1994), and by

Dickinson and Burke (1996) allow the traditional associative approach to extend to retrospective revaluation, they do not provide an explanation for this reciprocal blocking effect. With the revision to SOP (Dickinson & Burke, 1996), A and the outcome are always in A1 together, as is the control cue (which is never paired in compound, i.e., no Phase 2 training) and its outcome. As such, the associative strength of the two cues should be identical. AX pairing does not affect the associative strength of A, as this manner of pairing does not put A and the outcome in opposing states of activation. Likewise, with Van Hamme and Wasserman's (1994) modified Rescorla-Wagner rule, there is nothing that occurs during reciprocal blocking training that results in reduction in the strength of A (A continues to be presented with the outcome). Again, A and the control cue should have near identical associative strength.

The purpose of the current experiments is twofold. Firstly, we seek to replicate the effect of reciprocal competition in blocking (i.e., lowered responding to the blocking stimulus, A, due to compound pairing with the added, blocked stimulus, X), previously observed in non-human animals (rats), using a predictive scenario with human participants (Experiment 1). Secondly, we seek to examine the potential of cues in a causal scenario (Experiment 2) to elicit reciprocal blocking. Reciprocal blocking is seemingly incompatible with the explanations of blocking by most views of learning, excluding the comparator hypothesis. Likewise, blocking in a predictive task is incompatible with the inferential reasoning view of learning.

Experiment 1

The purpose of Experiment 1 was to investigate the hypothesis that reciprocal blocking occurs during traditional blocking. That is, lowered responding to the blocking stimulus, A, due to pairing A with what is typically known as the blocked stimulus, X. Although A blocks responding to X (traditional blocking), X should reciprocally lower responding to A (reciprocal

blocking). This occurs despite A continuing to serve as a reliable predictor of (and as the cue most informative of) outcome occurrence. The comparator hypothesis uniquely predicts such an effect, while other models of learning would have difficulty accounting for it. To investigate this hypothesis we employed a standard within-subjects blocking design within a predictive learning task. Responding to both X (to assess blocking) and A (to assess reciprocal blocking) was measured.

Method

Participants. The participants were 25 undergraduate students enrolled in psychology courses at Auburn University. All participants were recruited through the online service SONA, and received course credit in exchange for participation. All participants were at least 19 years of age (or 18 with parental consent).

Materials and apparatus. The experiment was completed in a room containing six Dell XPS computers; the computers were set up in two rows of three, on long tables on opposite walls of the room. Each computer was connected to a 20-inch flat panel Dell monitor; a large tower case was placed next to each monitor to serve as a divider between participants. This occluded the screen of each participant from others in the room.

To assess blocking and reciprocal blocking we used a computer task in which participants were required to learn the relationships between cues and outcomes. In the task, fictional preferences for specific foods (cues) were framed as *predictors* of different personality traits (outcomes) in a series of fictional subjects. After a food or a pair of foods was displayed on screen, participants selected the personality trait they thought was most likely predicted by that food preference. Participants received feedback as to the correct outcome in all training trials; at test participants did not receive feedback.

Design. Experiment 1 employed a within-subjects blocking design, meaning that all participants saw both the experimental and control cues. The task was presented in three uninterrupted phases: two training phases, and one test phase (see Table 5). Phase 1 of training consisted of ten presentations of two individual foods, referred to here as Predictor 1 (P1) and Predictor 2 (P2). Both Phase 1 foods were followed by specific personality traits: P1 by Outcome 1 (O1) and P2 by Outcome 2 (O2). In Phase 2 of training, P1 was paired with a new predictor, P3, and followed, once again, by O1. In addition, a novel pair of foods, P4 and P5, was followed by a new outcome, O3. There were four presentations of each compound during Phase 2. Phase 3 consisted of two individual test trials: The compound of P3 and P5 to assess blocking and the compound of P1 and P2 to assess reciprocal blocking. The cue-outcome relationships were automatically counterbalanced by the program, thus the food preferences and associated personality traits used to stand in for Predictors 1-5 and Outcomes 1-3 differed across participants. Similarly, the order of trial types during Phases 1 and 2, and the order of test trials were counterbalanced.

Table 5

Design of Experiment 1 (Predictors)

Phase 1	Phase 2	Test 1: Blocking	Test 2: Reciprocal blocking
10 P1→O1	4 P1P3→O1		
		P3P5?	P1P2?
10 P2→O2	4 P4P5→O3		

Note: The foods used to stand in for P1-P5 were: eggs, mustard, oatmeal, raisins, and wine. The personality traits used to represent O1-O3 were: sophisticated, perfectionist, and spontaneous.

Procedure. After entering demographic information (age and gender) participants went through two instruction screens (see Appendix A for task instructions) which gave them information on the cover story of the task, and how to enter responses. At the beginning of each trial the cue(s) were presented and remained on screen until the end of the trial. One second after the food preference(s) were presented, participants used the mouse to make a selection of one of three possible outcomes (personality traits). After the selection was made, feedback as to the correct outcome was presented and remained onscreen until the participant selected a button reading "Press here to see the next subject.", ending the trial. Participants completed 30 trials, including test trials. At the end of each test trial, a message reading "NO FEEDBACK!" was displayed on screen (see Appendix B for task images). There were no transitions or additional instruction screens between Phases 1 and 2 of training, or between training and test.

Results and Discussion

A repeated measures ANOVA was used to test for equivalence between cues during Phase 1 training. We found a main effect of trial $F(9,71)=26.4$, $p<.001$. There was no main effect of cue $F(1,24)=1.86$, $p=.185$. Additionally, there was no interaction between cue and trial, $F(9,216)=1.11$, $p=.361$. In Phase 2, responses and progression of learning were equivalent for both compounds: We found a main effect of trial $F(3,72)=8.34$, $p<.001$. There was no main effect of cue $F(1,24)=.05$, $p=.824$, and no cue/trial interaction, $F(3,72)=.26$, $p=.854$. In summary, participants' accuracy for P1 and P2 were equivalent, at the end of Phase 1. Likewise, accuracy was equivalent for both compounds at the end of Phase 2 (see also, Figure C1).

Next, we used a χ^2 test to examine independence of test trial order (whether a participant sees the test for blocking or reciprocal blocking first) and outcome selection. The outcomes selected during the tests for blocking and reciprocal blocking are independent of the order in

which the participant sees the tests: $\chi 2(1, N = 20) = 1.5$, $p = .219$, $\chi 2(1, N = 21) = .01$, $p = .920$, for blocking and reciprocal blocking respectively.

In order to test for blocking, we examined which outcome was selected by the participants when P3 and P5 were presented together at test. During training, P3 was followed by O1, and P5 was followed by O3 an identical number of times. If no blocking occurred, the probability of participants selecting either O1 or O3 should be the same (50%). However, if P1→O1 training during Phase 1 blocked responding to P3 during Phase 2 (P1P3→O1), then blocking should be observed, and participants should show greater preference for O3 than O1 (i.e., P5 would be a more reliable predictor of O3 than P3 was of O1). To analyze outcome selection data for blocking, we employed Pearson's χ^2 test to test for equal selection of both O1 and O3 ($\alpha = .05$). A null result indicates that there are no statistical differences in selection of each outcome (either O1 or O3); this would indicate that blocking did not occur. Effect size for the χ^2 test was calculated using the phi-coefficient (φ). Phi is interpreted in the same manner as Cohen's *w* statistic: a value of .1 indicates a small effect, .3 indicates a moderate effect size, and .5 or above is indicative of a large effect size (Cohen, 1992).

Five participants who did not select either O1 or O3 were excluded from the analysis. Seventeen out of the remaining 20 participants selected O3 at test (see Figure C2), indicating substantial blocking of P3 by P1 ($M = .85$, $\chi^2(1, N = 20) = 9.8$, $p < .01$, $\varphi = .7$).

The same analyses were used for the reciprocal blocking test (between P1 and P2). After Phase 1 training, both P1 and P2 have similar roles as predictors of O1 and O2, respectively. During Phase 2 training, there are additional presentations of P1 followed by O1. All learning theories we know, with exception of those supporting a comparator mechanism, would predict equal or greater O1 selection than O2 selection at test. Specifically, they predict

P1 serving as a similar (assuming asymptotic learning of P1 and P2 at the end of Phase 1) or stronger (due to additional P1 – O1 training during Phase 2) indicator of the occurrence of O1 than P2 does for O2. Four participants who did not select either O1 or O2 were excluded from the analysis. Sixteen out of the remaining 21 participants selected O2 at test (see Figure C3), indicating that pairing P1 with P3 reciprocally diminished (blocked) selections of O1 (M = .76, χ^2 (1, N = 21) = 5.76, p < .05, φ=.52).

We observed greater selection of O2, despite the additional experience with P1 and its respective outcome. These data are consistent with the findings of Arcediano et al. (2004), that is, they seem to support the view that blocking, at least under certain circumstances, is reciprocal.

Experiment 2

The purpose of Experiment 2 was to investigate reciprocal blocking with causes, in order to extend the effect of reciprocal blocking to a causal learning task. Reciprocal blocking (and traditional blocking) with predictors is incompatible with current inferential analyses of learning. The finding of reciprocal blocking in causes also has important implications for the inferential reasoning view as it violates the causal structure required for blocking. In this view, judgments of a strong cause should not be affected by compound training with a potential cause which is unable to change the outcome magnitude.

Method

Participants. We recruited 26 Auburn University undergraduate students. Recruiting procedures remained unchanged from Experiment 1.

Materials and apparatus. Participants completed the experiment in the same room outlined in Experiment 1. The task used was similar to Experiment 1; however, rather than

learning the relationship between food preferences and personality traits, participants were required to learn the relationship between the consumption of certain foods (causes) and the adverse reactions they cause (effects).

Design. The design of the task (including trial types and counterbalancing) was identical to that of Experiment 1. The sole difference is the use of causes rather than predictors. See Table 6 for a depiction of the design.

Procedure. The procedure in Experiment 2 was similar to the procedure in Experiment 1 with the corresponding changes from a predictive to a causal scenario. Participants entered demographic information and read through the instructions and cover story (see Appendix A for task instructions). At the beginning of each trial the food(s) eaten by the patient were presented and remained on screen until the end of the trial. One second after the food(s) were presented, participants used the mouse to make a selection of one of three possible adverse reactions. After the selection was made, feedback as to the correct outcome was presented and remained onscreen until the participant selected a button reading "Press here to see the next subject.". For test trials a message reading "NO FEEDBACK!" was displayed on screen. There were no transitions or additional instruction screens between Phases 1 and 2 of training, or between training and test.

Table 6

Design of Experiment 2 (Causes)

Phase 1	Phase 2	Test 1: Blocking	Test 2: Reciprocal blocking
10 C1→O1	4 C1C3→O1		
		C3C5?	C1C2?
10 C2→O2	4 C4C5→O3		

Note: The foods used to stand in for C1-C5 were: beans, cocoa, sesame, soy, and tuna. The adverse reactions used to represent O1-O3 were: asthma, diarrhea, and dizziness.

Results and Discussion

A repeated measures ANOVA was used to test for equivalence between cues across all Phase 1 trials. We found a main effect of trial $F(9,225)=18.95$, $p<.001$. There was no main effect of cue $F(1,25)=.298$, $p=.596$. Additionally, there was no interaction between cue and trial, $F(9,225)=1.05$, $p=.398$. In Phase 2, responses and progression of learning were equivalent for both compounds: We found a main effect of trial $F(3,75)=11.23$, $p<.001$. There was no main effect of cue $F(1,25)=.3.68$, p=.067; this trend toward significance is likely due to the novelty of the C4C5 compound: accuracy towards this novel compound was close to chance in the first trial of Phase 2, while the higher accuracy for the C1C3 compound is driven by prior experience with C1. As in Phase 1, there was no cue/trial interaction, $F(3,75)=.1.38$, $p=.255$. As in the previous experiment, participants' accuracy for Cause 1 (C1) and Cause 2 (C2) were equivalent at the end of Phase 1. Likewise, accuracy was equivalent for both compounds at the end of Phase 2 (see also, Figure C4).

We once again used a $\chi 2$ test to examine independence of test trial order and outcome selection. The outcomes selected during the tests for blocking and reciprocal blocking are independent of the order in which the participant sees the tests, $\chi 2(1, N = 24) = 1.61$, $p = .204$., $\chi 2(1, N = 20) = .85$, $p = .357$, for blocking and reciprocal blocking respectively.

As in Experiment 1, to analyze data for blocking, we employed Pearson's χ^2 to test for equal selection of Outcomes 1 and 3 (O1 and O3). Effect size for the χ^2 test was calculated using the phi-coefficient (φ). If no blocking occurred, the probability of participants selecting either O1 or O3 should be similar (roughly 50%). However, if C1→O1 training during Phase 1 blocked responding to C3 during Phase 2, participants should show greater preference for O3 than O1. Two participants who did not select either O1 or O3 were excluded from the analysis. Seventeen out of the remaining 24 participants selected O3 at test (see Figure C5), indicating substantial blocking of C3 by C1 ($M = .71$, $\chi^2 (1, N = 24) = 4.167$, $p < .05$, φ=.42).

The same analyses were used for the reciprocal blocking test (between C1 and C2). Six participants who did not select either O1 or O2 were excluded from the analysis. Fourteen out of the remaining 20 participants selected O2 at test (see Figure C6), indicating a strong trend towards C3 reciprocally diminishing (blocking) selections of C1 ($M = .70$, $\chi^2 (1, N = 20) = 3.2$, $p = .074$, φ=.4). Although reciprocal blocking is not statistically significant in the causal task, the effect size of .4 (φ-coefficient) suggests that this is a substantial effect, and that an issue of low sample size likely drove this result.

General Discussion

The current study supports the view that blocking may be reciprocal in humans, as it is in non-human animals (specifically rats; Arcediano et al., 2004); that is, the same way that the blocking cue reduces responding to the blocked cue, the blocked cue reduces responding to the

blocking one. Reciprocal blocking is pertinent because it challenges most prevalent traditional theories of learning, as well as the more recent view on learning based on inference making.

We observed strong evidence of the blocking and the blocked cues mutually competing with each other. Pairing a cue in compound with a previously trained cue results in blocking (diminished responding to the added cue). However, responding to the previously trained cue, despite the stronger relationship it should have acquired with the outcome, is also diminished. This reciprocal blocking effect occurred in predictive learning, and approached significance (with a moderate effect size) in causal learning. The data from the current study are consistent with the findings of Arcediano et al. (2004).

The traditional associative approach (Mackintosh, 1975; Rescorla & Wagner, 1972; Wagner, 1981) and the revised associative approach (e.g. Dickinson & Burke, 1996; Van Hamme & Wasserman, 1994) predict responding to C1 (the blocking cue) to be equal to or greater than responding to C2 (the control cue for reciprocal blocking). If both C1 and C2 reach asymptotic learning in Phase 1, responding to these cues should be equal. Otherwise, training in Phases 1 and 2 should lead to greater associative strength for C1, when it is compared to C2, which was only trained in Phase 1. In either case, C2 should not elicit greater responding than C1. As such, reciprocal blocking is problematic for these classes of models (as well as other models which are based on error prediction rules).

A normative analysis of the current results based on the traditional Δ-P (e.g., Allan, 1980, 1993) or the conditional Δ-P (e.g., Tangen & Allan, 2003) also fails to explain reciprocal blocking. The contingencies between C1, C2, and their respective outcomes, O1 and O2, are identical. As such, no differences in responding to these cues are to be expected. This pattern is

expected in both the "basic" and conditional Δ-P as responses track the cue-outcome contingency calculated by these rules.

The present results are also difficult to explain in the view of inferential and propositional models. Firstly, it is important to note that, in the inferential view, blocking with predictors (Experiment 1) is difficult to explain, as predictors are not privy to the same logical propositions which lead to blocking in causality. Specifically blocking with predictors violates the modus tollens argument (e.g., Pineño & Miller, 2007) which allows the blocked cue to be disregarded as an effective cause. However, blocking seemed to occur equally well regardless of the type of scenario (see also Arcediano et al., 2001, 1997; Chapman & Robbins, 1990; Chapman, 1991 for additional evidence of blocking in predictive situations with human participants). Secondly, even when blocking is predicted to occur (i.e., in the causal task) it is unknown how the propositional approach could explain reciprocal blocking: C1 was always followed by O1, while C3 was only followed by O1 in the presence of C1 (whenever O1 is present, C1 is as well). A cue which should not be viewed as a cause by participants should not have the ability to negatively affect responses to what should be seen as a very strong cause of the outcome.

The effect of reciprocal blocking seems to be uniquely predicted by the comparator hypothesis (Miller & Matzel, 1988; Stout & Miller, 2007). In this view, learning occurs through contiguity, and blocking is due to a response deficit. Responding is dependent on both what you have learned (by contiguity) about the target cue and the outcome, and what you have learned about the relation of other cues paired with the target cue and that same outcome. The stronger the relation between these other cues (called comparators in this framework) and the outcome, the more responding in the presence of the target cue is diminished. The aforementioned comparator mechanism differentiates this view of learning from those of its counterparts, and

allows for the prediction of both blocking and reciprocal blocking. The diminished responding to the blocked cue is driven by the strong association between the blocking cue and the outcome established in Phase 1. Likewise, this same mechanism allows the blocked cue to serve as a comparator stimulus and diminish responding to the blocking stimulus due to its pairings with the outcome in Phase 2.

References

Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, *15*(3), 147–149.

Allan, L. G. (1993). Human contingency judgments: Rule based or associative? *Psychological Bulletin; Psychological Bulletin*, *114*(3), 435.

Arcediano, F., Escobar, M., & Matute, H. (2001). Reversal from blocking in humans as a result of posttraining extinction of the blocking stimulus. *Animal Learning & Behavior*, *29*(4), 354–366. doi:10.3758/BF03192901

Arcediano, F., Escobar, M., & Matute, H. (2001). Reversal from blocking in humans as a result of posttraining extinction of the blocking stimulus. *Animal Learning & Behavior*, *29*(4), 354–366.

Arcediano, F., Escobar, M., & Miller, R. R. (2004). Is stimulus competition an acquisition deficit or a performance deficit? *Psychonomic Bulletin & Review*, *11*(6), 1105–1110.

Arcediano, F., Matute, H., & Miller, R. R. (1997). Blocking of Pavlovian conditioning in humans. *Learning and Motivation*, *28*(2), 188–199.

Balaz, M. A., Gutsin, P., Cacheiro, H., & Miller, R. R. (1982). Blocking as a retrieval failure: Reactivation of associations to a blocked stimulus. *The Quarterly Journal of Experimental Psychology Section B*, *34*(2), 99–113. doi:10.1080/14640748208400879

Balaz, M. A., Gutsin, P., Cacheiro, H., & Miller, R. R. (1982). Blocking as a retrieval failure: Reactivation of associations to a blocked stimulus. *The Quarterly Journal of Experimental Psychology*, *34*(2), 99–113.

Barnes, J. (1991). *The Complete Works of Aristotle, The Revised Oxford Translation, ed. Jonathan Barnes*. New Jersey: Princeton University Press.

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*(1), 92–102. doi:10.1037/0096-3445.135.1.92

Blaisdell, A. P., Gunther, L. M., & Miller, R. R. (1999). Recovery from blocking achieved by extinguishing the blocking CS. *Animal Learning & Behavior*, *27*(1), 63–76.

Bush, R. R., & Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, *58*(5), 313–323. doi:10.1037/h0054388

Carroll, C. D., Cheng, P. W., & Lu, H. (2013). Inferential dependencies in causal inference: A comparison of belief-distribution and associative approaches. *Journal of Experimental Psychology: General*, *142*(3), 845.

Chapman, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(5), 837.

Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition*, *18*(5), 537–545.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*(2), 367.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155.

Cole, R. P., Barnet, R. C., & Miller, R. R. (1995). Effect of relative stimulus validity: Learning or performance deficit? *Journal of Experimental Psychology: Animal Behavior Processes*, *21*(4), 293.

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, *10*(7), 294–300.

Daw, N. D., Courville, A. C., & Dayan, P. (2007). *Semi-rational Models of Conditioning: The Case of Trial Order*.

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*(1), 1–20. doi:10.3758/LB.37.1.1

De Houwer, J., Beckers, T., & Glautier, S. (2002). Outcome and cue properties modulate blocking. *The Quarterly Journal of Experimental Psychology Section A*, *55*(3), 965–985. doi:10.1080/02724980143000578

Dickinson, A., & Burke, J. (1996). Within compound associations mediate the retrospective revaluation of causality judgements. *The Quarterly Journal of Experimental Psychology: Section B*, *49*(1), 60–80.

Dickinson, A., Shanks, D., & Evenden, J. (1984). Judgement of act-outcome contingency: The role of selective attribution. *The Quarterly Journal of Experimental Psychology*, *36*(1), 29–50.

Fernbach, P. M., Darlow, A., & Sloman, S. A. (2011). Asymmetries in predictive and diagnostic reasoning. *Journal of Experimental Psychology. General*, *140*(2), 168–185. doi:10.1037/a0022100

Ghirlanda, S. (2005). Retrospective revaluation as simple associative learning. *Journal of Experimental Psychology: Animal Behavior Processes; Journal of Experimental Psychology: Animal Behavior Processes*, *31*(1), 107.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*(4), 334–384.

Guthrie, E. R. (1930). Conditioning as a principle of learning. *Psychological Review*, *37*(5), 412–428.

Hall, G., Mackintosh, N. J., Goodall, G., & Martello, M. D. (1977). Loss of control by a less valid or by a less salient stimulus compounded with a better predictor of reinforcement. *Learning and Motivation*, *8*(2), 145–158.

Holyoak, K. J., & Cheng, P. W. (2011). Causal learning and inference as a rational process: The new synthesis. *Annual Review of Psychology*, *62*, 135–163.

Kamin, L. J. (1968). Attention-like" Processes In Classical Condiioning. In *Miami Symposium On The Production Of Behavior Aversive Stimulation* (pp. 9–33).

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. (2008). Bayesian approaches to associative learning: From passive to active learning. *Learning & Behavior*, *36*(3), 210–226.

López, F. J., Shanks, D. R., Almaraz, J., & Fernández, P. (1998). Effects of trial order on contingency judgments: A comparison of associative and probabilistic contrast accounts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 672–694. doi:10.1037/0278-7393.24.3.672

Lovibond, P. F., Been, S. L., Mitchell, C. J., Bouton, M. E., & Frohardt, R. (2003). Forward and backward blocking of causal judgment is enhanced by additivity of effect magnitude. *Memory & Cognition*, *31*(1), 133–142.

Lu, H., Rojas, R. R., Beckers, T., & Yuille, A. (2008). Sequential causal learning in humans and rats. Cognitive Science Society. Retrieved from http://dare.uva.nl/record/292508

Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review*, *115*(4), 955–984. doi:10.1037/a0013256

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

Matzel, L. D., Schachtman, T. R., & Miller, R. R. (1985). Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. *Learning and Motivation*, *16*(4), 398–412. doi:10.1016/0023-9690(85)90023-2

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1992). Responding to a conditioned stimulus depends on the current associative status of other cues present during training of that specific stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, *18*(3), 251.

Miller, R. R., Hallam, S. C., & Grahame, N. J. (1990). Inflation of comparator stimuli following CS training. *Animal Learning & Behavior*, *18*(4), 434–443.

Miller, R. R., & Matute, H. (1996). Biological significance in forward and backward blocking: resolution of a discrepancy between animal conditioning and human causal judgment. *Journal of Experimental Psychology. General*, *125*(4), 370–386.

Miller, R. R., & Matzel, L. D. (1988). *The comparator hypothesis: A response rule for the expression of associations. ln GH Bower (Ed), The psychology of learning and motivation (Vol. 22, pp. 51-92)*. San Diego, CA: Academic Press.

Miller, R. R., & Schachtman, T. R. (1985). Conditioning context as an associative baseline: Implications for response generation and the nature of conditioned inhibition. *Information Processing in Animals: Conditioned Inhibition*, 51–88.

Pierce, W. D., & Epling, F. (1995). *Behavior analysis and learning*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Pineño, O., & Miller, R. R. (2007). Comparing associative, statistical, and inferential reasoning accounts of human contingency learning. *The Quarterly Journal of Experimental Psychology*, *60*(3), 310–329.

Pineño, O., Urushihara, K., & Miller, R. R. (2005). Spontaneous Recovery From Forward and Backward Blocking. *Journal of Experimental Psychology: Animal Behavior Processes*, *31*(2), 172–183. doi:10.1037/0097-7403.31.2.172

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, *66*(1), 1.

Rescorla, R. A., & Wagner, A. R. (1972). A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. *Cognitive Science*, 147.

Shanks, D. R. (2010). Learning: From Association to Cognition. *Annual Review of Psychology*, *61*(1), 273–301. doi:10.1146/annurev.psych.093008.100519

Spellman, B. A. (1996). Conditionalizing Causality. In D. L. M. and K. J. H. David R. Shanks (Ed.), *Psychology of Learning and Motivation* (Vol. Volume 34, pp. 167–206). Academic Press.

Stout, S. C., & Miller, R. R. (2007). Sometimes-competing retrieval (SOCR): A formalization of the comparator hypothesis. *Psychological Review*, *114*(3), 759.

Tangen, J. M., & Allan, L. G. (2003). The relative effect of cue interaction. *The Quarterly Journal of Experimental Psychology: Section B*, *56*(3), 279–300.

Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and Motivation*, *25*(2), 127–151.

Wagner, A. R. (1969). Stimulus selection and a "modified continuity theory." *The Psychology of Learning and Motivation*, *3*, 1–41.

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In *Information processing in animals: Memory mechanisms* (Vol. 85, pp. 5–47).

Wagner, A. R., Logan, F., & Haberlandt, K. (1968). Stimulus selection in Animal discrimination learning. *Journal of Experimental Psychology*, *76*(2, Pt.1), 171–180. doi:10.1037/h0025414

Wasserman, E. A., & Neunaber, D. J. (1986). College students' responding to and rating of contingency relations: The role of temporal contiguity. *Journal of the Experimental Analysis of Behavior*, *46*(1), 15.

Wheeler, D. S., Beckers, T., & Miller, R. R. (2008). The effect of subadditive pretraining on blocking: Limits on generalization. *Learning & Behavior*, *36*(4), 341–351. doi:10.3758/LB.36.4.341

Appendix A.
Task Instructions

The introductory screen contained the same information in both experiments. The text read as follows:

> Please carefully read the following instructions and try to do your best.
>
> You will be using the computer's mouse for this task. Feel free to change the mouse and mouse pad to your left if you prefer to do so.
>
> You know you have finished with the experiment when you see a screen with a message thanking you for participating in the study.
>
> --Press the right button on the mouse to start--

Next, participants saw the instructions specific to the task (causal or predictive) they were undergoing. The instructions for Experiment 1 read as follows:

> Imagine you are part of a research group that wants to analyze if some food preferences may predict certain personality traits. Over the next few minutes, you are going to see the records collected from a number of subjects. For each record you will first see which food(s) the subject preferred. Then, you will have to correctly select the personality trait predicted by that food(s) preference.
> --Press the right button on the mouse when you are ready to start--

The instructions for Experiment 2 read as follows:

> Imagine you are part of a research group that has discovered that eating some foods may cause certain adverse reactions. Over the next few minutes, you are going to see the records collected from a number of subjects. For each record you will first see which food(s) the subject ate. Then, you will have to correctly select what adverse reaction was caused by eating that food(s).
> --Press the right button on the mouse when you are ready to start--

Finally, after completing the task participants saw the following exit screen, regardless of the task they underwent:

> THANK YOU FOR PARTICIPATING IN THIS STUDY!

Appendix B.
Task Images

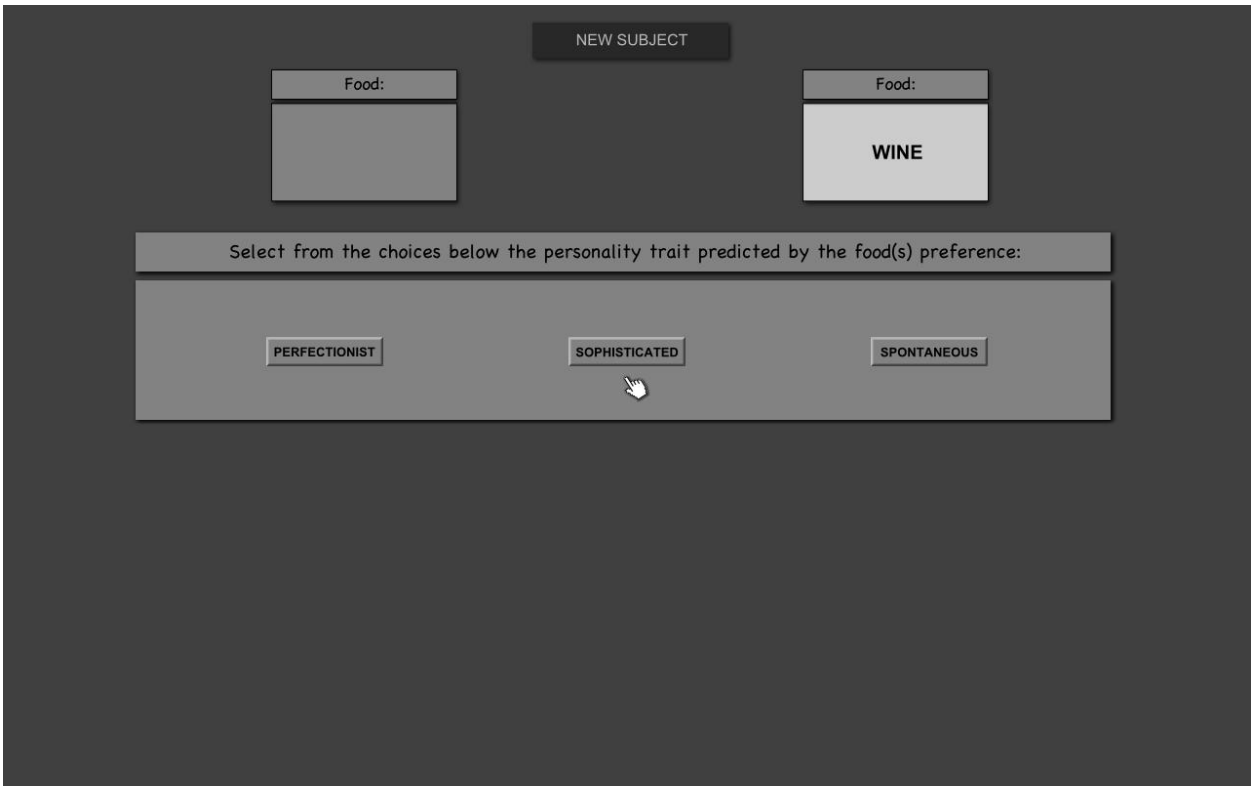

Figure B1. Example of cue presentation in a Phase 1 trial. (Experiment 1)

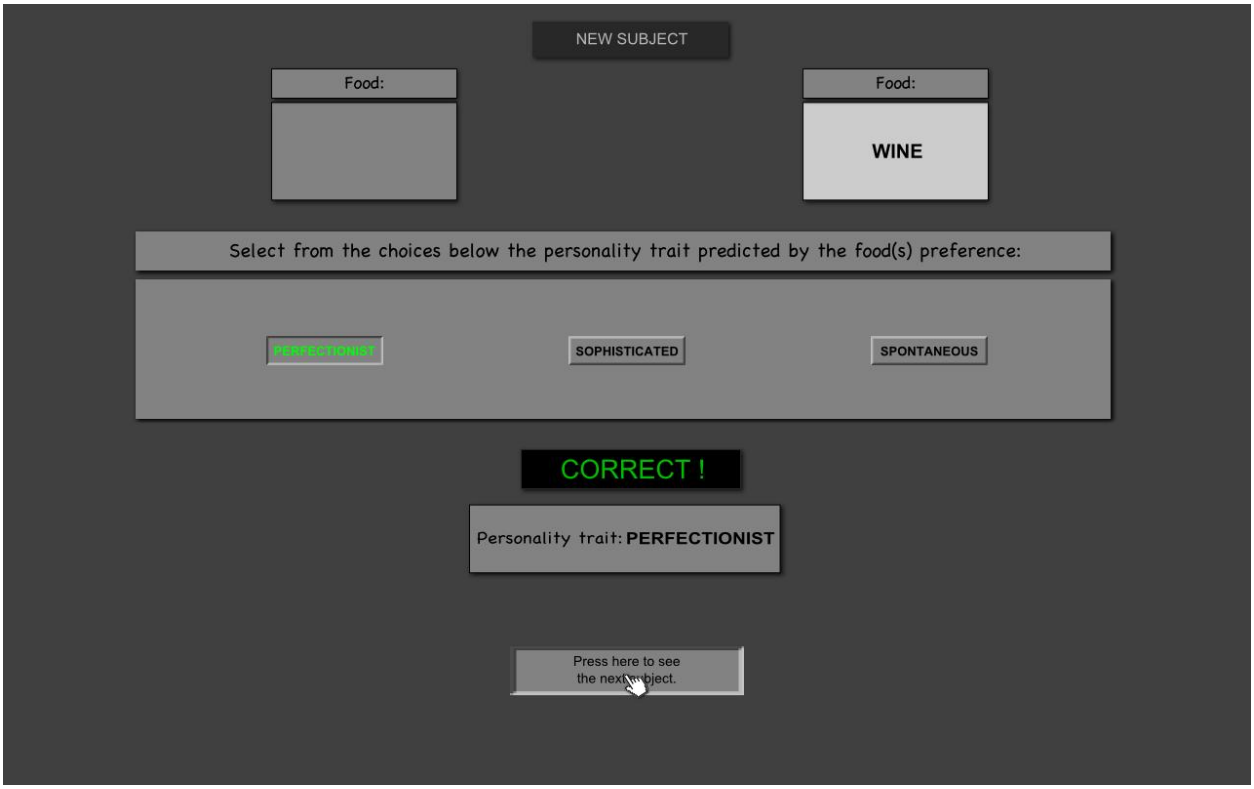

Figure B2. Example of feedback for a correct answer during Phase 1. (Experiment 1)

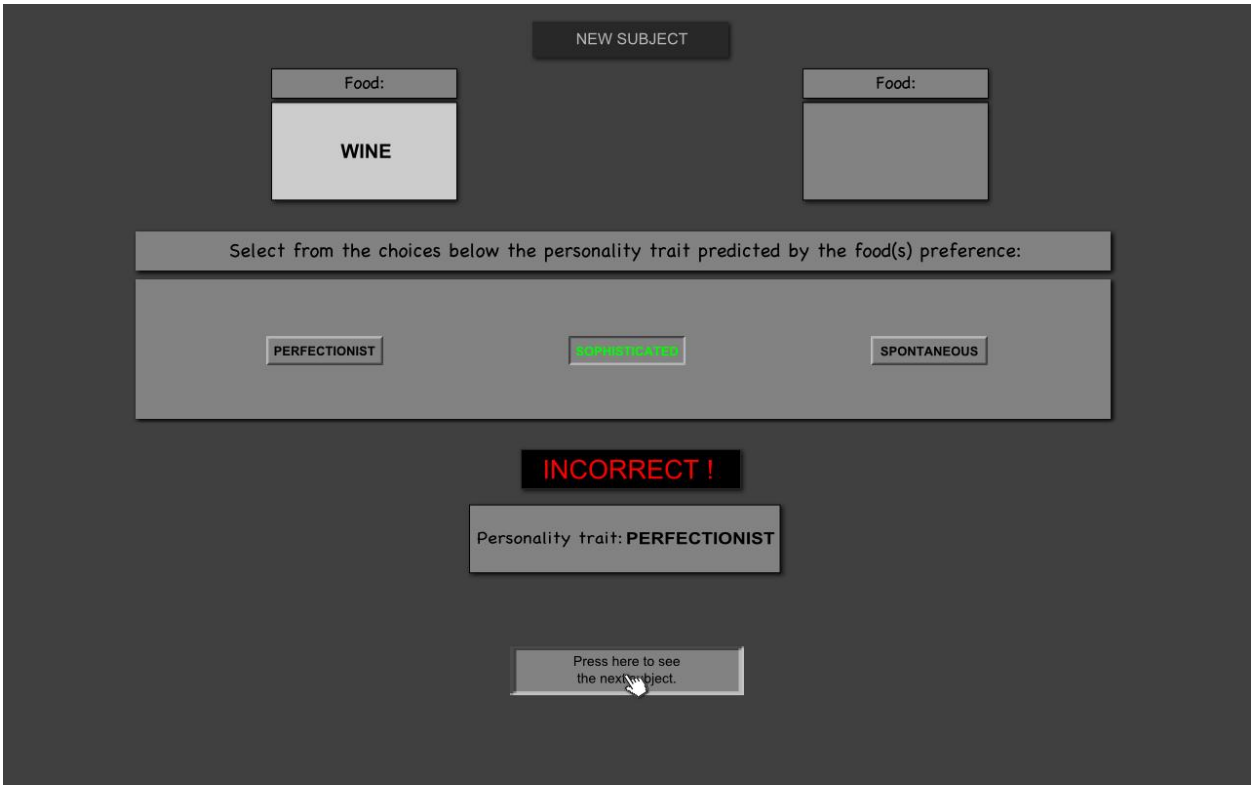

Figure B3. Example of feedback for an incorrect answer during Phase 1. (Experiment 1)

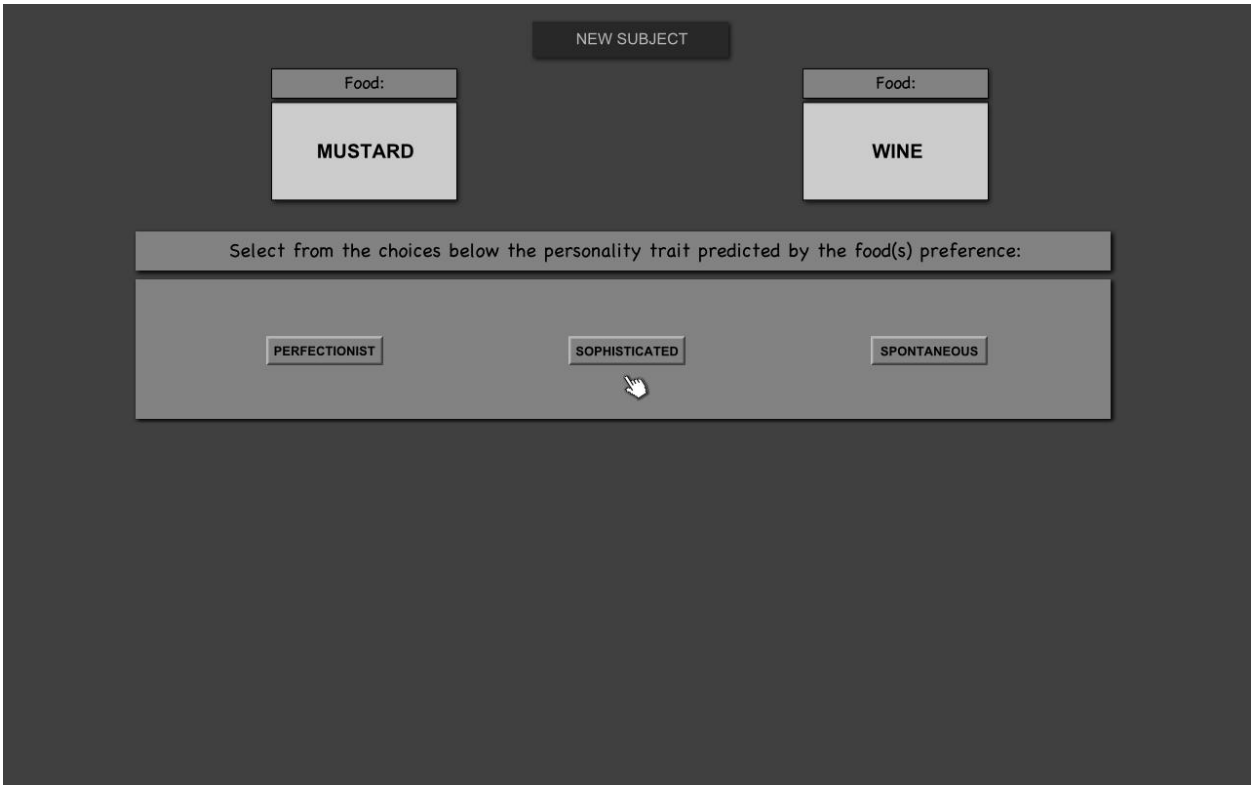

Figure B4. Example of cue presentation in a Phase 2 trial. (Experiment 1)

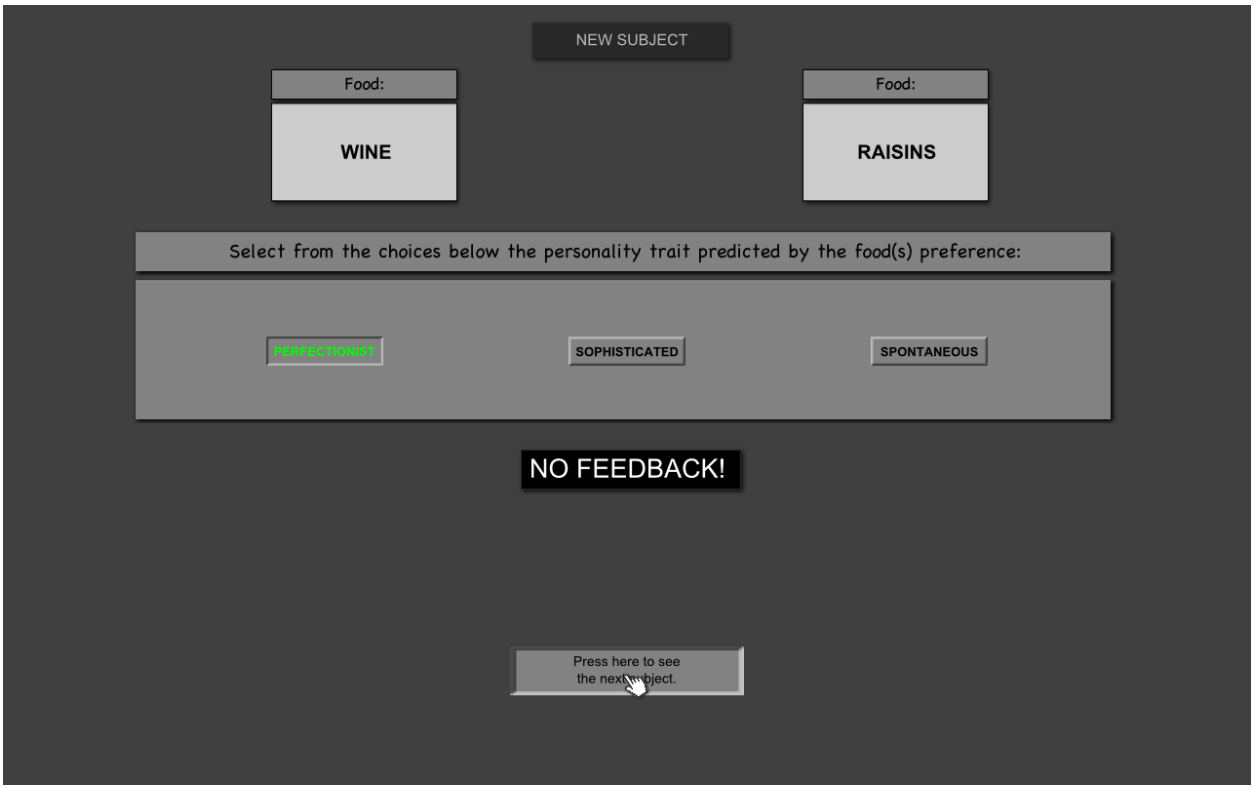

Figure B5. Example of outcome selection in a test trial. No feedback given. (Experiment 1)

Figure B6. Example of cue presentation in a Phase 1 trial. (Experiment 2)

Appendix C.
Results and Graphs

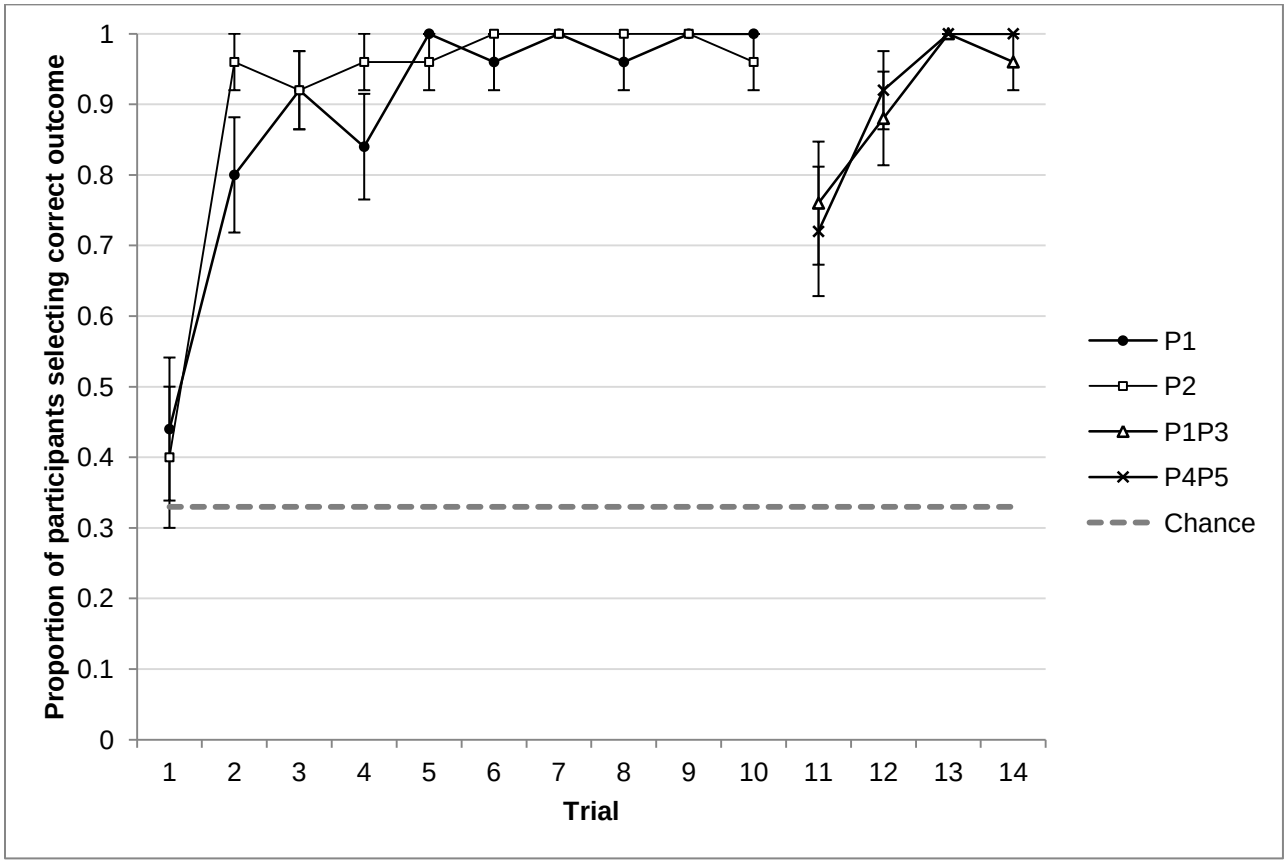

Figure C1. Accuracy across trials for all cues and compounds during Phases 1 and 2. Accuracy is presented as the proportion of participants selecting the correct outcome in each trial. These data are from Experiment 1.

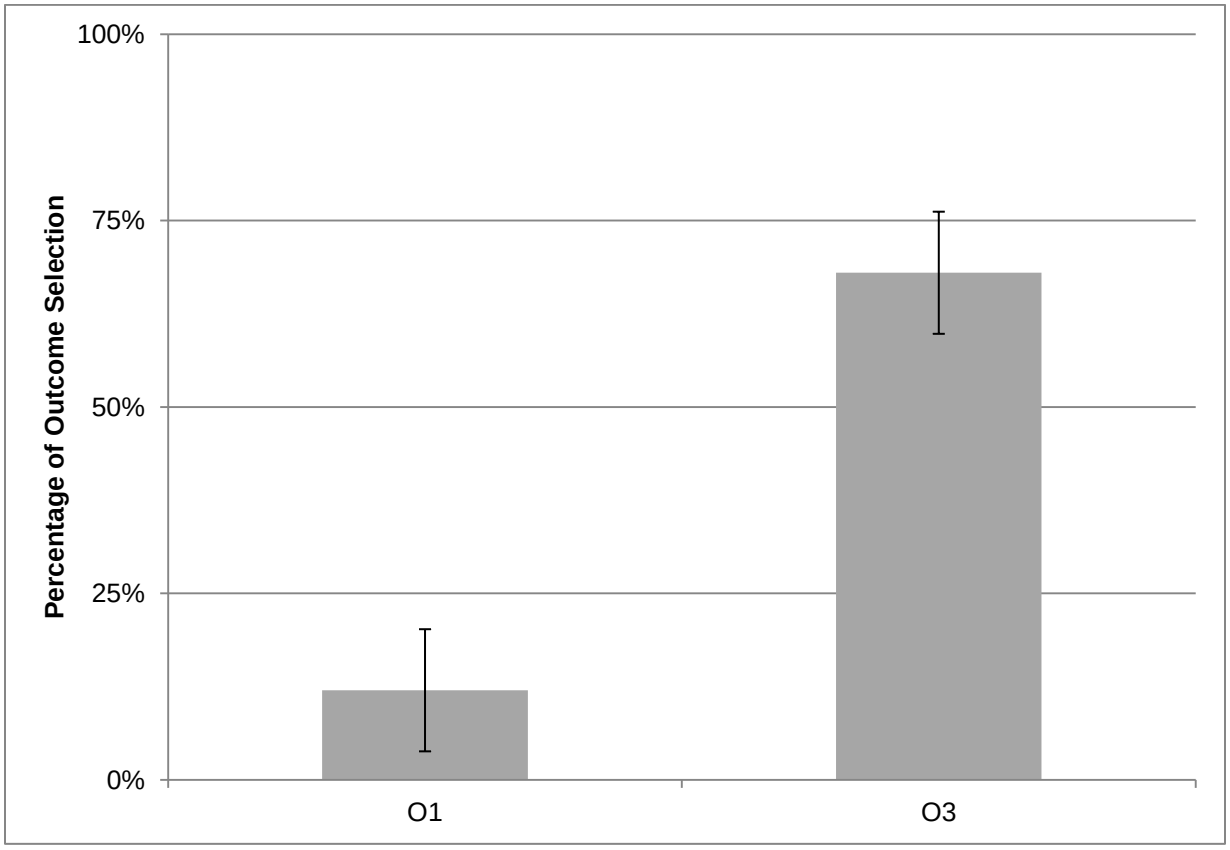

Figure C2. Test for blocking. Depicts the percentage of participants selecting each outcome at test. These results are from Experiment 1.

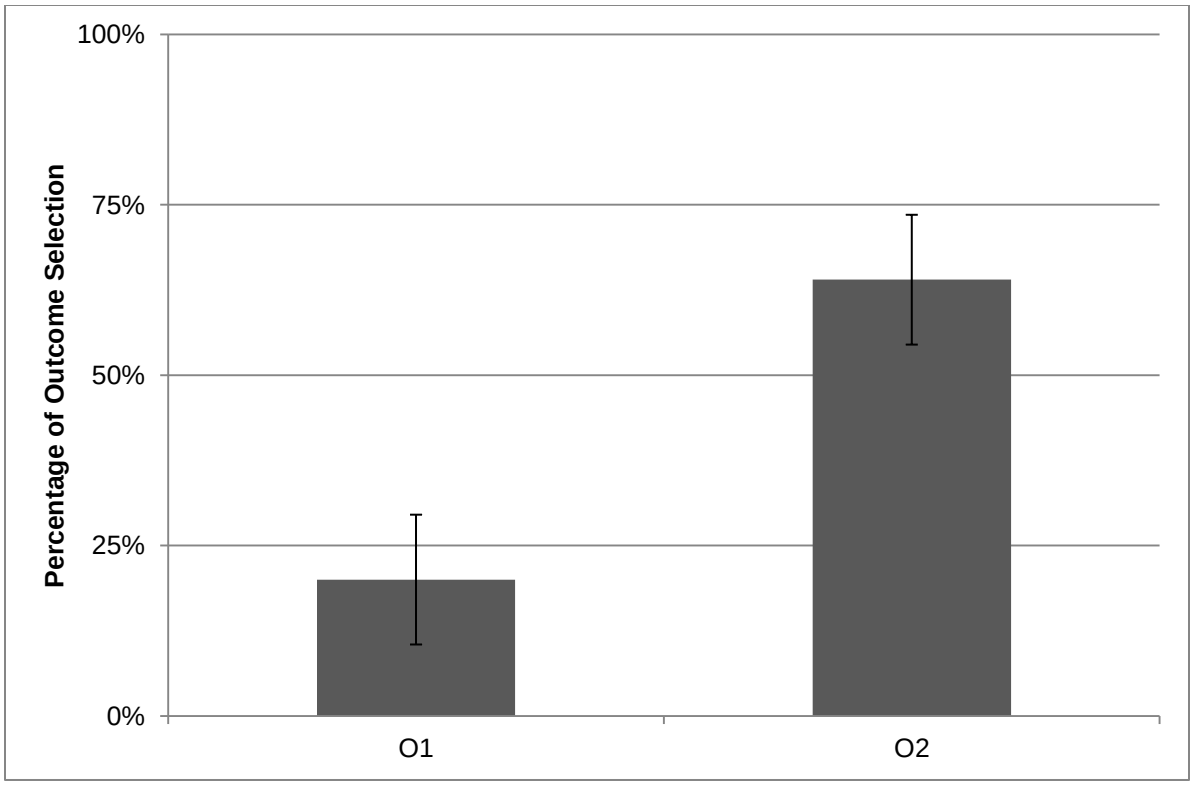

Figure C3. Test for reciprocal blocking. Depicts the percentage of participants selecting each outcome at test. These results are from Experiment 1.

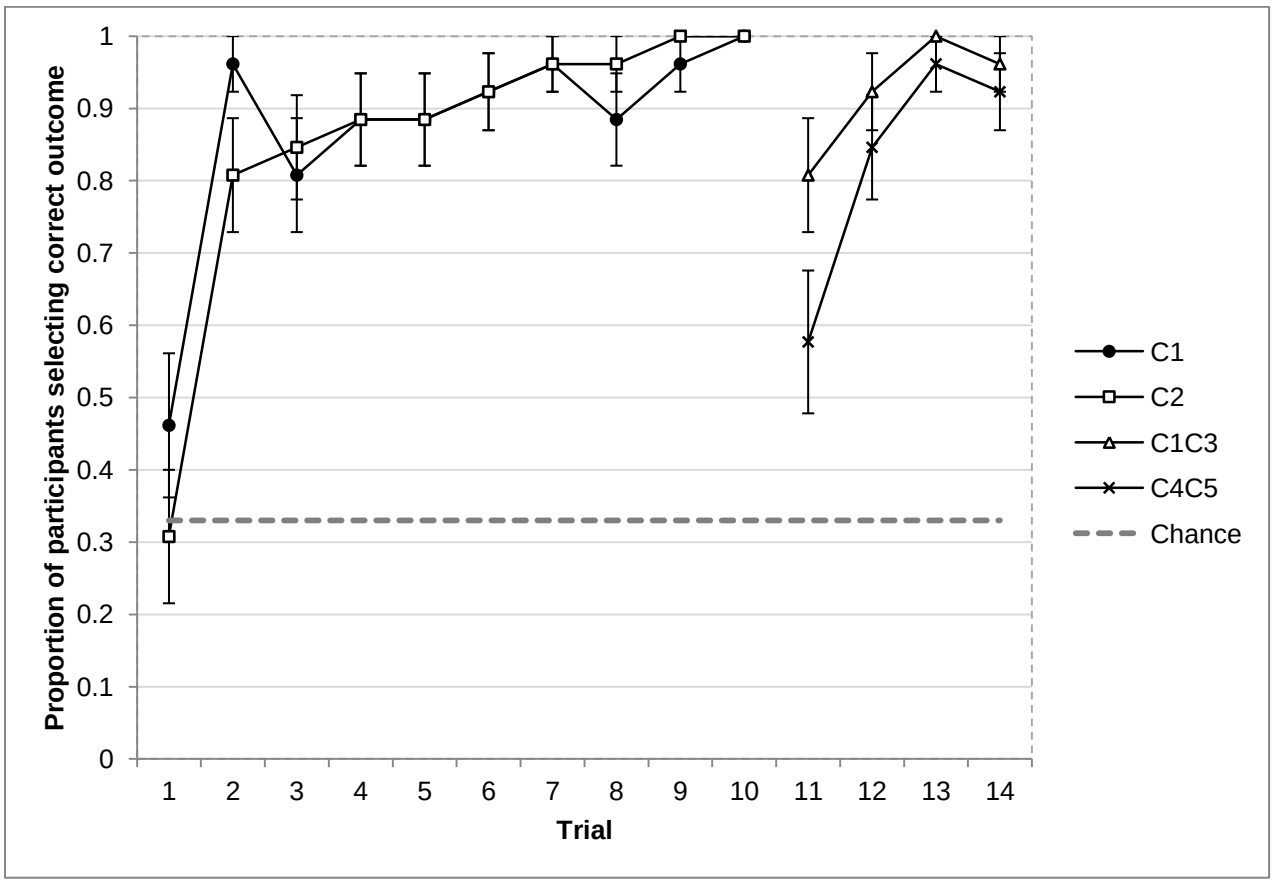

Figure C4. Accuracy across trials for all cues and compounds during Phases 1 and 2. Accuracy is presented as the proportion of participants selecting the correct outcome in each trial. These data are from Experiment 2.

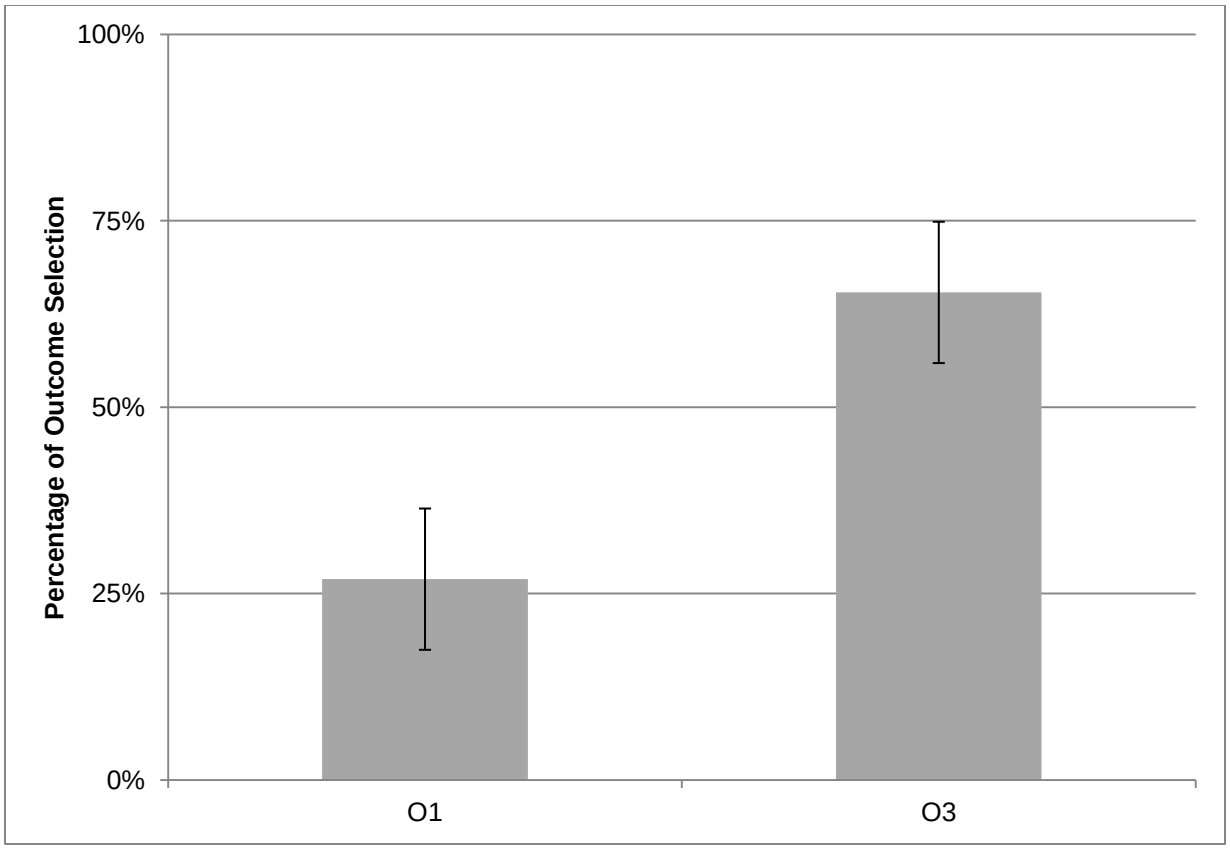

Figure C5. Test for blocking. Depicts the percentage of participants selecting each outcome at test. These results are from Experiment 2.

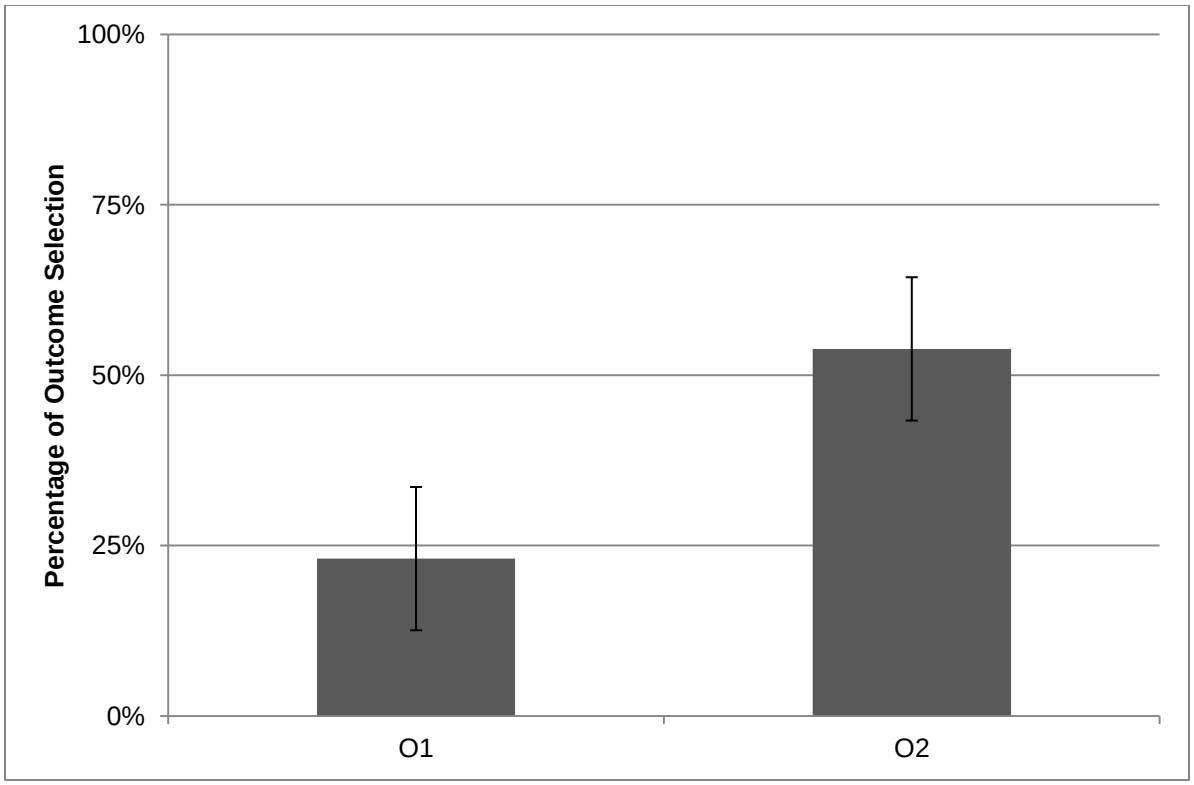

Figure C6. Test for reciprocal blocking. Depicts the percentage of participants selecting each outcome at test. These results are from Experiment 2.