INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

CONTRASTING ASSOCIATIVE AND STATISTICAL THEORIES OF CONTINGENCY JUDGMENTS

Rick R. Mehta

Department of Psychology

McGill University, Montréal

May 2000

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements of the degree of Doctorate of Philosophy

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First, I thank my thesis supervisor, Professor Andy Baker, for all of his help and encouragement over the course of this project. I am grateful to him for having shared his knowledge, expertise and insights with me.

Second, I would like to thank Louise LeBrun for her secretarial assistance; Marie-Claire Albanese for her French translation of the abstract; Matthew Holahan for organizing a graduate student forum in which I could present my research to my peers and get valuable feedback; as well as Winnifred Louis and Dr. Geert-Jan Boudewijnse for proofreading and critiquing my thesis.

Third, I want to thank the following individuals for their warmth and inspiration: Professor John Abela, Marie-Claire Albanese, Rhonda Amsel, Kelly Barlow, Sean Barrett, Dr. Deborah Black, Jennifer Bossert, Evelyne Bougie, Kimberly Burton, David D'Alessandro, Sensei Thierry Debeur, Kourosh Farrokhzad, Kristina Gill, Neil Hanley, Sandra Jittler, Dr. Mich Kabay, Anne Lacasse, Dr. William Lariviere, Winnifred Louis, Paul Mackenzie, Dr. Kenneth Mah, Professor Ronald Melzack, Carmelo Milo, Abigail Poore, Caroline Pukall, Michelle Prostak, Kelly Purdy, Tanja Rapus, and Jonathan Webber. All of you made my life at graduate school the wonderful experience that it was and I will always cherish the memories.

Finally, words can not express the gratitude I have for my mother, Kanta Mehta, whose unconditional love and emotional support was never unnoticed and always appreciated.

During my graduate career, I received substantial financial support from the Natural Sciences and Engineering Research Council (NSERC) and Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR). All of the research reported in this thesis was funded by an NSERC operating grant to Professor Andy Baker.

ABSTRACT

"Blocking" refers to judgments of a moderate contingency being lowered when contrasted with a strong contingency. The Rescorla-Wagner model and causal model theory account for blocking through different mechanisms. To examine the predictions from these two models, seven experiments tested the extent to which "causal scenario" and "causal order" would influence whether blocking was observed in human contingency learning tasks. "Causal scenario" was manipulated by contrasting responses to two causes of one effect or to one cause of two effects; "causal order" was defined as causes preceding effects or effects preceding causes. The four conjunctions of these two factors were investigated separately in Experiments 1 to 5. In Experiments 1 and 2, two causes preceded one effect and two effects preceded one cause, respectively. Blocking was observed regardless of whether the predictors were causes or effects. In Experiments 3, 4 and 5, participants were presented with one antecedent cue and made separate predictions about each of the trial's two outcomes. Blocking was not observed, irrespective of whether the antecedent cue was a cause or an effect. These initial results were consistent with the Rescorla-Wagner model. An alternative explanation was that blocking failed to occur in Experiments 3 to 5 because participants were asked questions between the predictor and two outcomes. Predicting the outcomes might have implicitly led participants to monitor them separately and to report on subsets of the data at the time of judgment. To address this issue, the volunteers in Experiment 6 observed the events on each trial but did not make any predictions about the outcomes. Blocking was observed, signifying that the intervening questions between the antecedent and consequent cues constitute an important variable influencing cue competition effects. In Experiment 7, all four conjunctions of causal scenario and causal order were tested simultaneously. Furthermore, participants were not asked questions between the antecedent and consequent events. Blocking was observed in all four conditions, that is, regardless of causal scenario and causal order. This pattern of results was not anticipated by either the Rescorla-Wagner model or causal model theory. The implications for associative and statistical models are discussed.

RÉSUMÉ

Le concept du 'blocking' fait référence à des jugements d'une contingence modérée étant atténués losque comparés à une forte contingence. Le modèle Rescorla-Wagner ainsi que la théorie du modèle de causalité expliquent le concept du 'blocking' à l'aide de mécanismes différents. Afin d'examiner les prédictions de ces deux modèles, sept expériences ont evalué si le scénario de causalité ainsi que l'ordre de causalité influencent l'apparition du 'blocking' au sein de tâches d'apprentissage des contingences chez des sujets humains. Le scénario de causalité a été manipulé en comparant les réponses à deux causes d'un effet ou a une cause de deux effets; l'ordre de causalité a été défini en terme de causes précédant les effets ou en terme d'effets précédant les causes. Les quatre combinaisons de ces deux facteurs ont été étudiées séparément dans les expériences 1 à 5. Dans l'expérience 1, deux causes ont précédé un effet et dans l'expérience 2, deux effets ont précédé une cause. Le 'blocking' a été observé indépendamment de si les facteurs de prédiction étaient des causes ou des effets. Dans les expériences 3,4 et 5, les sujets ont été exposés à un avis antécédant et ont fait des prédictions séparées à propos des deux conclusions de l'essai. Le 'blocking' n'a pas été observé, indépendamment de si l'avis antécédant était une cause ou un effect. Ces résultats initiaux étaient en accord avec le modèle Rescorla-Wagner. Une explication alternative était que le 'blocking' n'a pas été observé dans les expériences 3 et 5 parce que les sujets ont été questionnés entre le facteur de prédiction et les deux conclusions. La prédiction des conclusions a peut-être implicitement induit les sujets à porter attention aux conclusions séparément et à se référer à des portions de données au moment du jugement. Afin d'élucider ce probleme, les participants volontaires dans l'expérience 6 ont observé les événements à chaque essai, mais n'ont fait aucune prédiction à propos des conclusions. Le 'blocking' a été observé, signifiant que les questions posées entre les avis antécédants et conséquents constituent une variable importante influençant les effets de compétition des avis. Dans l'expérience 7, les quatre combinaisons du scénario de causalité et de l'ordre de causalité ont été examinées simultanément. De plus, les sujets n'ont pas été questionnés entre l'événement antécédant et l'événement conséquent. Le

'blocking' a été observé dans toutes les quatre conditions, c'est-à-dire, indépendamment du scénario de causalité ou de l'ordre de causalité. Ces résultats n'étaient prédits ni par le modèle Rescorla-Wagner, ni par le modèle de causalité. L'implication de ces résultats au sein des modèles associatifs ainsi que statistiques est discutée.

INTRODUCTION

It is important for animals and people to accurately process information about covariation and causal relationships. This knowledge allows them to predict and react to relationships in the world. Hence, accurate processing of contingencies is essential for survival, as well as for understanding, predicting, and controlling events that occur in the surrounding environment. For example, rodents process covariation information when determining an optimal foraging strategy. Similarly, people normally cross a street when the light is green based on the belief that they are more likely to cross safely when it is green, as opposed to yellow or red. Furthermore, some people may avoid foods high in fat if they believe this food will increase their risk of being obese, or they may avoid smoking cigarettes if they believe that smoking causes lung cancer. Causal reasoning is also seen in more complex activities such as scientists evaluating which independent variable is responsible for an experimental outcome or economists determining which factors caused a recession. These examples attest to the importance and ubiquity of causal reasoning and demonstrate that contingency plays a prominent role in diverse aspects of animals' and people's lives.

Philosophical Issues

The fundamental question of how people infer causality has been of philosophical interest since at least the time of David Hume (Fales & Wasserman, 1992; Wasserman, 1990b). In *A Treatise of Human Nature*, Hume (1946/1739) postulated that causal relations were deduced solely by sensory experience. For example, he asserts that "...cause and effect are relations, of which we receive information from experience, and not from any abstract reasoning or reflexion" (p. 69) and that "... by experience only ... can [one] infer the existence of one object from that of another" (p. 89). In this book he outlines a series of assumptions by which causal relationships are deduced from the environment, and asserts that these assumptions apply both to humans and to animals.

Since both species process causal information for survival, it seems reasonable to question whether there is a common mechanism.

A central tenet in Hume's view was his emphasis on contiguity for inferring causality. He deemed contiguity "...as essential to that of causation" (p. 75). This view was elaborated in the postulation of spatial and temporal proximity, namely that "the cause and effect must be contiguous in space and time" (p.173). Apart from contiguity, however, Hume and others (e.g., Mackie, 1974; Reichenbach, 1956; von Wright, 1974) have argued that there is a temporal asymmetry between causes and effects. That is, causes precede effects but effects do not precede causes. Furthermore, causes have precedence over effects because it is possible to manipulate the cause to alter the probability of an outcome, but an outcome can not be manipulated to alter the occurrence of the cause. Hume labelled this postulate the "priority of time in the cause before the effect" (Hume, 1946/1739, p. 76).

A third assumption is that "[t]here must be a constant union betwixt the cause and effect. 'Tis chiefly this quality that constitutes the relation" (p. 173). However, other philosophers have considered this notion simplistic. For example, Mackie (1974) has argued that a causal connection between two events is defined not only as their co-occurrence, but also on the fact that when the first event does not occur, neither does the second. This idea has been termed the counterfactual, and experimental evidence suggests that the ability to reason about it is evident in children as young as 3 to 5 years old (Harris, German, & Mills, 1996).

Hume's fourth axiom was that "... the same cause always produces the same effect, and the same effect never arises but from the same cause" (p. 173). However, a moment's reflection reveals that this notion is also simplistic. A cause does not always produce the same effect because there are other intervening factors in the environment that can influence what effect the cause produces. For example, although rain is important in determining whether crops will grow, it will not help the crops if the soil is infertile. Similarly, if there is too much rain, then the crops will also not grow. The second half of this assumption, that a specific effect arises from only one cause, is also incorrect. This is

because the same effect can arise from a variety of causes. For example, a death by internal bleeding can be caused not only by a weak vein or artery bursting, but also by a fatal blow or injury.

The following assumptions are of interest because they complement those traditionally attributed to John Stuart Mill (1949/1843). Hume's fifth axiom was that when several causes produce an effect, the causes must have a common quality that allows them to do so. This notion is similar to Mill's method of agreement. Sixth, if two similar objects produce different effects, the difference in the effects arises due to the property or properties by which the objects differ. This assumption resembles Mill's method of difference, which states that if an event occurs in one instance but not another, the circumstance by which the two instances differ is the cause of the event. Seventh, Hume states that the presence or absence of one part of the cause is proportional to the presence or absence of a proportional part of the outcome. This notion is similar to Mill's method of residues. The method of residues states that two phenomena are causally related to one another when one varies with the other.

Hume last assumption is that "an object, which exists for any time in its full perfection without any effect, is not the sole cause of that effect but requires to be assisted by some other principle, which may forward its influence and operation."(p 174 - 175). This notion implies that although an object might not directly influence an outcome, it still might be important if it is a necessary condition or if it works in conjunction with other causes.

An objection raised by Mill (1949/1843) and others (e.g., Hart and Honoré, 1959; Mackie, 1974; see also Hilton & Slugoski, 1986) against Hume's analysis was that causal events are not evaluated in a vacuum, but instead are evaluated against a point of reference. That is, to be considered causal, a potential event has to be different from events that occur merely as part of the background. Kant (1965/1781) also objected to Hume's reasoning and postulated that people infer causality by intuitively understanding that one event causes the second because of the power or energy that the first has over the second. For example, a pool cue causes a billiard ball to move because the cue hits the

ball, transfers energy to it, which causes the ball to move. This notion has been termed generative transmission (Shultz, 1982).

Sensitivity to Contingency

There are a variety of approaches that can be used to study how people infer causality. For example, one could test how people are influenced by intervening variables (Busemeyer, McDaniel, & Byun, 1996), how people reason about complex biological or engineering systems (White, 1998), or how people assess the relationship between two continuous variables (Jennings, Amabile, & Ross, 1982; Trolier & Hamilton, 1986). For brevity and clarity, this thesis will focus on the simplest causal relationship, involving only one binary cause and one binary effect. This relationship can be summarized in a 2 x 2 contingency table (Figure 1) in which the cause and effect are either present or absent. In this table, cell A refers to the frequency of the co-occurrence of the cause and effect; cell B to the cause and the absence of the effect; cell C to the effect in the absence of the cause; and cell D to the frequency of both the cause and effect being absent. Although there are various methods of assessing contingency (e.g., phi, chi-squared), the accepted normative statistic for the relationship between a cause and effect is Δp (Allan, 1980; see also Inhelder & Piaget, 1958). This statistic is defined as the difference between the probability of the effect occurring in the cause's presence and absence, i.e., Δp = p (Effect | Cause) - p (Effect | No Cause). In terms of the frequencies in Figure 1, $\Delta p = [A / (A+B)] - [C / (C+D)]$. If Δp equals zero, there is no relationship between the cause and effect. If Δp is greater than zero, the relationship is positive. And if Δp is less than zero, the relationship is negative. That is, the cause makes the effect less likely to happen. The advantages of Δp over the other methods of assessing covariation are 1) it uses all four cells in a contingency table and 2) it is unbiased to variations in cell frequency.

Since the focus of this thesis is on the mechanism by which information about cause-effect relations is acquired, it is important to establish whether people are sensitive to contingencies. This section will briefly review the literature on this issue and will

	Effect	No Effect
Cause	A	B
No Cause	C	D

$$\Delta p = p(\text{effect} \mid \text{cause}) - p(\text{effect} \mid \text{no cause})$$

$$= (A / (A+B)) - (C / (C+D))$$

Figure 1. A 2 x 2 contingency table displaying the four possible conjunctions between a cause and effect. The cell entries represent the frequency of each type of event, and Δp represents the contingency between the event and outcome.

describe the conditions under which people tend to make accurate contingency judgments.

According to Inhelder and Piaget (1958), children as young as 15 are aware of the concept of contingency and that all four cells are needed to assess covariation. In some early studies, however, animals were found to be sensitive to contingencies (e.g., Hammond, 1980; Rescorla, 1968, 1969), whereas humans were not (e.g., Crocker, 1981; Jenkins & Ward, 1965; Nisbett & Ross, 1980; Shaklee, 1983; see also Peterson & Beach, 1967). For example, when introducing the topic of contingency, Nisbett and Ross state that people are not only poor at assessing covariation, but also that

> a priori theories or expectations may be more important to the perception of covariation than are the actually observed data configurations. That is, if the layperson has a plausible theory that predicts covariation between two events, then a substantial degree of covariation will be perceived, even if it is present only to a very slight degree or even if it is totally absent. Conversely, even powerful empirical relationships are apt not to be detected or to be radically underestimated if the layperson is not led to expect such a covariation." (Nisbett & Ross, 1980, pp. 10).

This assertion was derived from early reports demonstrating that people either did not use the information provided or used sub-optimal strategies when performing covariation tasks. For example, L.J. Chapman and J.P. Chapman (1967, 1969; see also Golding & Rorer, 1972) have reported that people are insensitive to valid predictors of events, or report a correlation between events when there is none. The latter finding has been termed the *illusory correlation*. Furthermore, people have been reported to use only cell A (Smedslund, 1963; however, for a critical review of this study, see Vallée-Tourangeau, Hollingsworth, & Murphy, 1998), to compare cells A and B (Shaklee & Mims, 1982; Shaklee & Wasserman, 1986), or to sum cells A and D (Ward & Jenkins, 1965). These strategies tend to be used when the tasks' memory demands are high (Arkes & Harkness, 1983; Shaklee & Mims, 1982). The obvious problem with these tactics is that they use only one or two of the four cells in the table. A more complex strategy involves comparing the sums of the diagonals (Allan & Jenkins, 1983; Arkes & Harkness, 1983; Shaklee & Tucker, 1980), for example, comparing the sum of A and C to that of B and D.

The problem with this technique is that it is biased and hence can lead to errors in estimates of covariation if one of the totals is much larger than the other.

More recent studies have found that people in general are accurate at judging Δp (Allan & Jenkins, 1980; Alloy & Abramson, 1979; Baker, Berbrier, & Vallée-Tournageau, 1989; Dickinson & Shanks, 1985; Dickinson, Shanks & Evendon, 1984; Shanks, 1986; Wasserman, 1990b), although some systematic biases have been noted. Participants sometimes use a positive test strategy in which they examine and remember instances that confirm rather than falsify a relationship between the cause and effect (Klayman & Ha, 1987, 1989; see also Fischoff & Beyth-Marom, 1983). For instance, participants rate the presence of the outcome as more important than its absence (Kao & Wasserman, 1993; Schustack & Sternberg, 1981; Wasserman, Dorner, & Kao, 1990; see also Beyth-Marom, 1982; Crocker, 1982), especially when they expect the contingency to be positive (Levin, Wasserman, & Kao, 1993). Furthermore, small cell frequencies (i.e., less than two) tend to be overestimated (Arkes & Harkness, 1983), presumably because events of low frequency are more available and thus are remembered as occurring more often than they did (Tversky & Kahneman, 1974). Last, people can discriminate moderately positive and negative contingencies (Δp = 0.5 and -0.5 respectively) from zero contingencies (Baker et al., 1989; Dickinson et al., 1984). But for zero contingencies, judgments of Δp are sometimes influenced by outcome density (Alloy & Abramson, 1979; Baker et al., 1989; Jenkins & Ward, 1965; Vallée-Tourangeau, Hollingworth & Murphy, 1998; Wasserman & Shaklee, 1984), although this is not a finding that is consistently replicated (Allan & Jenkins, 1980; Chatlosh, Neunaber & Wasserman, 1985; Shanks & Dickinson, 1991; Wasserman, Chatlosh, & Neunaber, 1983). Outcome density is the mean probability of an outcome occurring regardless of whether the cause is present or absent. For example, Baker et al. (1989) found that a zero contingency was rated as negative when the outcome density was 0.25 [p (outcome | cause) = p (outcome | no cause) = 0.25; low density 0 contingency], but was rated as positive when the outcome density was 0.75 [p (outcome | cause) = p (outcome | no cause) = 0.75; high density 0 contingency]. Furthermore, Wasserman, Elek, Chatlosh, and Baker (1993; see also Baker,

Murphy, & Vallée-Tourangeau, 1996) have found that although participants' contingency judgments were highly correlated with Δp ($r = 0.97$), at any particular contingency these ratings deviated more from Δp as the outcome density decreased. Outcome density influences not only human contingency judgments, but also influences conditioning in animal learning experiments (e.g., Benedict & Ayres, 1972; Kremer, 1971, 1974; Kremer & Kamin, 1971; Quinsey, 1971). For example, levels of conditioning for zero contingencies are faster as the outcome's occurrence increases.

In general, though, the evidence implies that people can be quite good at judging and discriminating contingencies. Different methods have been proposed to help increase this accuracy. For example, judgments tend be more accurate when they are presented with symmetric rather than asymmetric variables (Beyth-Marom, 1982; Trolier & Hamilton, 1986). An example of a symmetric variable would be a tumor defined as malignant or benign, whereas an asymmetric variable would consist of the tumor being either cancerous or not cancerous (Beyth-Marom, 1982). The former variable is symmetric because it can take on two levels of severity, i.e., extremely severe or less severe. The latter variable, on the other hand, is asymmetric because a non-cancerous tumor implies the absence of malignancy. Judgments are more accurate with symmetric variables, because the outcome is present but to different degrees. In this case, participants are likely to consider both levels as equally important.

Apart from the type of variable used during experimentation, another means of increasing accuracy is to ensure that participants understand what the task entails. For example, Crocker (1982) has argued that participants should be informed of all four conjunctions of a binary cause and effect when they are asked for their judgment. She has also asserted that a detailed description of covariation helps clarify what the experimenter wants the participant to assess (Crocker, 1981). Furthermore, explicitly stating that contingencies might be positive or negative, and using a rating scale that includes both positive and negative numbers helps ensure that people will rate positive and negative contingencies symmetrically (Wasserman,1990b). Informing people that there could be

no relationship between the cause and effect helps them recognize non-contingent relationships (Peterson, 1980).

Another factor that influences people's accuracy is the method by which contingencies are presented. That is, the contingencies can be presented over a series of trials or in a summary table. Some have argued that judgments are more accurate when presented in summary table format (e.g., Beyth-Marom, 1982; Ward & Jenkins, 1965; Wasserman & Shaklee, 1984; Shaklee & Mims, 1982). However, Baker et al. (1989) found that participants were less accurate when the individual trials were printed on a questionnaire. More recent reports suggest that acquiring contingencies over trials does not hinder accuracy, although there are systematic errors. For example, Vallée-Tourangeau, Hollingsworth, & Murphy (1998) and Kao and Wasserman (1993) reported that participants are accurate at judging contingencies when they are presented over trials, although judgments were less accurate and were influenced by outcome density when the contingency was zero. Similarly, Catena, Maldonado, & Cándido (1998) reported decreases in accuracy as the frequency of giving ratings increased. In summary, judgments in older studies were worse when contingencies were presented over trials, but more recent studies show that people can be accurate with this presentation format. Thus, the argument that participants are considerably worse at assessing covariation when acquiring contingencies over trials is overstated. Last, it seems reasonable to assume that information about ecologically relevant cues is normally acquired through experience rather than through reference to a summary table. Hence, trial by trial presentation is more likely to emulate real world causal induction mechanisms.

In summary, although there are sometimes systematic biases, participants can accurately judge Δp. Unfortunately, this issue is still oversimplified in textbooks of cognition and decision-making. For example, prominent authors such as Baron (1994) and Plous (1995) continue to advocate the position that people are poor at assessing covariation and rely extensively on cell A, despite the evidence contrary to this assertion.

Mechanisms of Contingency: Associative and Normative Models

Although it is apparent that people can accurately judge and discriminate contingencies, the cognitive mechanism by which they achieve this accuracy is not understood. There are two prominent classes of models that explain how people infer causality: associative theories and normative, or statistical, models. The purpose of this dissertation is to compare them and to further understand how they can account for contingency learning. The models will be described in this section. In the two sections that follow, this thesis will examine conditions under which the models make similar and different predictions for experimental results.

Associative models were initially proposed to explain findings from animal learning. However, it has been observed that people's judgments on causal acquisition tasks mirrored findings from animal learning and it has thus been argued that theoretical accounts used for animal conditioning might be extended to human causal judgments (e.g., Alloy & Tabachnik, 1984; Shanks & Dickinson, 1987; Young, 1995; see also Allan, 1993; Siegel & Allan, 1996; Shanks, Lopez, Darby, & Dickinson, 1996) and category learning (Gluck & Bower, 1988; Shanks, 1990). For example, if the probability of an outcome in the presence of a cause is held constant, people's contingency judgments decrease as the probability of the outcome in the cause's absence increases (Allan & Jenkins, 1980; Alloy & Abramson, 1979; Dickinson et al., 1984). Similarly, animals will display less Pavlovian conditioning and instrumental responding to a conditioned stimulus as the probability of the unconditioned stimulus in the absence of the conditioned stimulus increases (Rescorla, 1968; Dickinson & Charnock, 1985, respectively). Furthermore, signaling the outcomes that occur in the absence of the cause can modestly increase human causal judgments (Shanks, 1986) and levels of both classical and instrumental conditioning in animals (Durlach, 1983; Rescorla, 1984).

Associative models stem from the views of the British Empiricists such as Hume, and include temporal and spatial contiguity as the mechanism by which cues are linked to outcomes. Although there are a variety of associative models (e.g., Pearce & Hall, 1980; Mackintosh, 1975; Miller & Matzel, 1988; Van Hamme & Wasserman, 1994), the one

most cited and used to explain people's contingency judgments is the Rescorla-Wagner model (Rescorla & Wagner, 1972; see also Wagner & Rescorla, 1972; hereafter R-W model). The R-W model is equivalent to the delta rule used to update weights in some connectionist models (e.g., Gluck & Bower, 1988; McClelland & Rumelhart, 1985) and can be summarized by the formula

$$\Delta V_i = \alpha_i \beta_{outcome} (\lambda - \Sigma V)$$

In this formula, ΔV_i is the change in associative strength on a given trial i. α_i is a learning rate parameter that is unique for each cue and represents the salience or associability of that cue. It has a positive value when it is present and is set to zero when the cue is absent[1]. $\beta_{outcome}$ is a learning rate parameter for the outcome. λ is the maximum associative strength that the outcome will support. This parameter is equal to 0 when the outcome is absent, and takes on any value greater than 0 when it is present. And ΣV is the amount of associative strength accrued to all of the cues present on a given trial, including the context. The context is usually defined as the constant features of the testing situation.

An advantage of the R-W model is it includes few assumptions. The R-W model is based on the principles of contiguity and cue competition, and it assumes that there is a monotonic relationship between a cue's associative strength and the causal judgment. Thus, the contiguity or co-occurrence of the cue and effect is sufficient for a cue to gain associative strength. Cue competition arises because an outcome can support only a limited amount of associative strength. Thus, different cues compete to be associated with the outcome.

For a single cue-outcome relationship, the cue competes with the context for associative strength. Depending on the contiguity between the cue and outcome, the associative strength can accrue to either the cue or context. The cue will gain associative strength on trials in which it is followed by an outcome. But the cue will lose associative

---

[1] Although this issue will not be elaborated, it should be noted that Van Hamme and Wasserman (1994) argue that α_i should sometimes take on a negative value, that is, when the cue is expected but does not occur.

strength if it is presented and not followed by the outcome. And if the outcome occurs in the cue's absence, the associative strength between the cue and outcome will remain unchanged, but the associative strength between the context and outcome will increase because it is the context that is reinforced with the outcome.

For example, for a moderately positive contingency, the cue and context will initially have no associative strength. When the cue is followed by the outcome, the associative strength between the cue and outcome will increase. Thus ($\lambda - \Sigma V$) will initially be large and the cue will gain large increments of associative strength. The cue will gain associative strength as it is reinforced, and will lose associative strength when it is not reinforced. The context will gain associative strength when the outcome is presented in the absence of the cue. Over trials, because the cue is reinforced more often than the context, it will accumulate more associative strength. The increments in the cue's associative strength will initially be large but will decrease over trials. Thus, the resulting acquisition function will be a negatively accelerating curve. At asymptote, the cue will have a moderate amount of associative strength, whereas the context will have accumulated less associative strength. The causal relationship will therefore be judged as moderately positive. Thus, the R-W model is sensitive to contingency but does not explicitly calculate it (Baker et al., 1996, see also Van Overwalle, 1996, 1998).

Although the R-W model is most frequently used to account for judgments about covariation, two other associative models have recently gained popularity: Pearce's (1994, 1987) stimulus generalization model and, to a lesser extent, Miller's comparator hypothesis (Miller & Matzel, 1988). Pearce's model was originally designed to explain how animals performed discrimination tasks, but has recently been successful at explaining human contingency judgments (López, Shanks, Almaraz, & Fernández, 1998; Vallée-Tourangeau, Hollingsworth, & Murphy, 1998; Vallée-Tourangeau, Murphy, & Baker, 1998; Vallée-Tourangeau, Murphy, Drew & Baker, 1998). Although Pearce's model will not be discussed in detail, it should be noted that Pearce's model also assumes that organisms have a limited capacity to process information to which they are exposed. A key difference between Pearce's model and the R-W model is that Pearce (1994, 1987)

postulates that the constellation of stimuli presented on a trial are learned about as a whole, whereas the R-W model is an elemental theory that treats stimuli as individual elements that are learned about independently of one another. Thus, in Pearce's (1994, 1987) model, the internal representation corresponds to the overall pattern of events and outcomes, rather than to individual cues and outcomes.

Miller's comparator hypothesis (Miller & Matzel, 1988) differs from both the R-W and Pearce's models. During acquisition, organisms form associations between cues and outcomes. After acquisition, when an animal is presented with the test stimulus or when a person is asked for a contingency judgment, the organism will compare the target cue's associative strength with the associative strength accumulated by all the other cues present during the acquisition phase. The strength of the response or contingency judgment will then depend on the target cue's associative strength relative to the cue most strongly associated with the outcome.

In contrast to associative models, normative models compute contingencies across trials. These models assume that humans act like naive scientists who infer causes by considering whether outcome events take place more often in the presence or absence of alternative factors (e.g., Kelley, 1973; Spellman, 1996). Two influential models are the Probabilistic Contrast model (hereafter PC model; Cheng & Novick, 1990; 1992; see also Cheng & Holyoak, 1995) and Waldmann and Holyoak's (1992) causal model theory. The PC model has recently been modified and is known as the Power Theory of the Probabilistic Contrast Model (hereafter Power PC model; Cheng, Park, Yarlas, & Holyoak, 1996; Cheng, 1997; Wu & Cheng, 1999). The PC model includes two important theoretical tools: Δp and the focal set. Whereas the R-W model states that participants' judgments are based on a cause's accumulated associative strength, both the PC model and causal model theory postulate that participants implicitly calculate Δp.

Furthermore, according to the PC model, participants make contingency judgments based on "conditional contingencies computed over a contextually constrained set of events termed the *focal set*" (Cheng et al., 1996: p. 315; italics in original). In other words, Δp is not computed unconditionally, but instead is computed over contextually

determined events in which the cause and effect occur. Thus, within an appropriate focal set, a potential cause is evaluated by computing the probability of the outcome occurring in both the presence and absence of the cause. When there are multiple causes, participants evaluate each potential cause in the presence and absence of alternative causes. The notion of focal sets stems from the view that causal events are evaluated against a point of reference, and that to be considered causal, a potential event has to be different from events that occur merely as part of the background (Einhorn & Hogarth. 1986; Mackie, 1974). For example, causal events are usually unexpected and can be evaluated by comparison with alternative explanations for the effect.

While Δp and the focal set are powerful theoretical tools that allow people to judge whether two variables covary, statistical regularity is insufficient for making judgments about causality. To address this problem, the Power PC model includes a third theoretical construct: "causal power". Causal power, or generative transmission (Shultz, 1982), incorporates Kant's (1965/ 1781) notion that one event causes another because of power or energy transferred from the first to the second.

Waldmann and Holyoak's causal model theory (1992; see also Waldmann, 1996) is similar to the Power PC model in that it includes the notion of conditionally calculating Δp. However, the two models differ in terms of which theoretical constructs they emphasize. The Power PC model focuses on computational processes underlying Δp calculations, whereas causal model theory focuses on how mental models of cause influence causality judgments. This model's assumptions will be described and discussed in detail later.

Opposing Explanations for Empirical Phenomena

Normative and associative mechanisms differ from one another, yet sometimes yield similar predictions for empirical findings. I will discuss three examples: temporal delays, blocking; and relative validity. The first example is the finding that judgments of Δp are lowered if there is a delay between an action and an outcome, and that the effect of the delay is diminished if a signal is interposed during the delay. For example, Shanks

and Dickinson (1987, 1991) have found that a delay of two or four seconds between pressing a computer's space-bar and producing an outcome decreased causal judgments (see also Gruber, Fink, & Damm, 1957; Leslie, 1982; Shanks, Pearson & Dickinson, 1989; Siegler & Liebert, 1974; Wasserman & Neunaber, 1986; however, see Wasserman, Chatlosh, & Neunaber, 1983). Schlottmann and Shanks (1992) obtained similar results using Michotte's (1963) launch task. In this task, two objects are seen on a computer screen; one is on the left end and the other is in the middle of the screen. The first object moves across the screen until it reaches the border of the second object. At this point, the first object stops moving and the second one begins to move. As delays increased from 17 to 289 ms, the authors found that causal ratings decreased. Furthermore, similar results are seen in animal learning. For example, pigeons display less responding for food if there is a 3 second delay between responding for and receiving food (Richards, 1981; Williams, 1976). However, if a tone is present during the delay between the cause and outcome, judgments of the causal relationship are elevated (Reed, 1992; Shanks, 1989; see also Dickinson & Charnock, 1985; Rescorla, 1984; and Williams & Heyneman, 1982 for similar findings from animal learning). According to an associative approach, the tone could serve as a secondary reinforcer that would allow the cause and outcome to be associated with one another (Reed, 1999). On the other hand, according to a normative approach, the tone would make it more likely that the cause and effect would be linked in a causal chain (Einhorn & Hogarth, 1986).

The second example has to do with the "blocking" phenomenon. In the first phase of a blocking experiment (Kamin, 1969; see also Arcediano, Matute, & Miller, 1997; Chapman, 1991; Chapman & Robbins, 1990; Dickinson et al., 1984), an experimental group is exposed to a stimulus that is paired with an outcome (i.e., A => outcome). In the second phase, the stimulus is presented in compound with another stimulus, and this compound stimulus is paired with an outcome (i.e., AB => outcome). The control group, on the other hand, does not experience trials in which A is first paired with the outcome. In the final or test phase, the response to B is measured. The common finding is that there is little responding to B in the experimental group relative to the control group. Similarly,

in human causal inference experiments, judgments of B are lowered. According to the R-W model, the associative strength of A would approach λ during the first phase. At the beginning of the second phase, since ΣV is equal to λ, the linear operator ($\lambda - \Sigma V$) would equal zero. Because B is introduced in this phase, no associative strength would accrue to it, and little responding would be expected.

A normative account makes a similar prediction, albeit through a different mechanism (Cheng & Holyoak, 1995; Waldmann & Holyoak, 1992). In the first phase, A is established as a strong predictor of the outcome since the outcome always occurs in A's presence but never in its absence. In the second phase, B is also a clear predictor of the outcome since it is always presented in compound with A. Thus, it will also have a large unconditional Δp. However, participants will not be able to establish whether B is an independent cause of the effect since B is always presented in compound with A and there are no trials in which B is presented in A's absence. Since the participants have incomplete information about the conditional contingency between B and the outcome, they will be uncertain about B's predictability and will give it a low rating. In summary, both the normative and the R-W accounts predict that prior training with A will reduce judgments of B. The normative model explains blocking in terms of the undefined Δp and the resulting uncertainty about the relationship between B and the outcome, whereas the R-W model explains blocking solely in terms of competition for associative strength.

A more complex finding that is predicted by both models but explained differently by each is the "relative validity" phenomenon observed in animals by Wagner, Logan, Haberlandt, & Price (1968; see also Wasserman, 1974). This has also been demonstrated in humans by Wasserman (1990a; Van Hamme & Wasserman, 1993) and Baker (Baker, Mercier, Vallée-Tourangeau, Frank, & Pan, 1993; Baker et al., 1996; Vallée-Tourangeau, Baker, & Mercier, 1994). In the original experiment by Wagner et al. (1968), a light (L) was paired with a shock on 50% of trials. In both the experimental ("True Discrimination") and control ("Pseudo-Discrimination") conditions, the light was presented in a two-element compound with one of two tones (T_1 and T_2). In the True

Discrimination, the outcome was present only when L was presented in compound with T_1 (i.e., T_1L^+) but was absent when L was presented in compound with T_2 (i.e., T_2L^-). In the Pseudo-Discrimination, the light was paired with the outcome on half of the trials in which it was presented in compound with T_1 and half of the trials in which it was presented in compound with T_2 (i.e., $T_1L^{+/-}$, $T_2L^{+/-}$). In the test phase, the animals responded more to the light following the Pseudo-Discrimination than following the True Discrimination, even though the light had been paired with the shock on 50% of the trials in both conditions.

According to the R-W model, the cues in the True Discrimination condition (i.e., T_1L, T_2L and the context) would initially gain associative strength. Over trials, T_1L would continue to gain associative strength because it is consistently reinforced. The context and T_2L, on the other hand, would lose associative strength because they are never paired with the outcome. Thus, responding at asymptote would be high for T_1L since it accumulates lots of associative strength, whereas there would be little responding to T_2L and the context since they acquire little associative strength. In terms of the individual cues, T_1 would accumulate more associative during training because it is consistently paired with the outcome and T_2 would lose associative strength because it is never paired with the outcome. L would gain associative strength when it is paired with the shock, that is, on the T_1L trials. However, it would lose associative strength on the T_2L trials when the shock is not presented. L would thus experience both increases and decreases in associative strength over trials, and by the test phase would have little associative strength. In the Pseudo-Discrimination, the context would initially gain some associative strength when either of the tone-light compounds is reinforced. However, this associative strength would decrease over trials because there are no trials in which the outcome occurs in the presence of the context by itself. The context would thus accumulate the least amount of associative strength. Both T_1 and T_2 would acquire little associative strength since they are both reinforced on half the trials in which they are presented. L, on the other hand, would gain more associative strength since it is presented in compound with T_1 and T_2 and thus would be paired with the outcome more than either of the two

tones. According to the R-W model, L would accumulate more associative strength in the control than in the experimental condition, and this difference would account for the greater response to the light following the Pseudo-Discrimination.

According to the PC model, by contrast, participants would first decide on an appropriate focal set. For the tones, the appropriate focal set is the presence of the light. In the Pseudo-Discrimination, each of the tones would have a Δp of zero because the outcome occurs on half of the trials in which they are present, as well as on half of the trials in which they are absent. In the True Discrimination, T_1 would have a Δp of 1 because the outcome always occurs in its presence but never in its absence; T_2 would have a Δp of -1 because the outcome always occurs in its absence but never in its presence.

For the light, the appropriate focal set is the time spent in the experimental context in which the tones and light occur. Without the context, the light's Δp is undefined because there are no trials in which the light is absent; in other words, [C/(C+D)] is indeterminate. But by expanding the focal set to include the context, the PC model can predict the relative validity effect by calculating Δp for the light in the absence of T_1 (Baker et al., 1996). For the Pseudo-Discrimination, T_2 has a lower contingency than the light and thus does not act as a conditional cue. The light's unconditional Δp is therefore calculated as 0.5. In the True Discrimination, the light's Δp is zero because the outcome does not occur in either the light's presence or absence. Since the light's Δp is greater in the Pseudo-Discrimination condition, greater responding to the light would be expected in this group than in the True Discrimination condition. Thus, both the R-W and PC models can account for the relative validity effect, albeit through different mechanisms.

Opposing Predictions: Effects of Mental Models on Contingency Judgments

These examples demonstrate how associative and normative models often make similar predictions about empirical phenomena. However, these models sometimes make different predictions about people's judgments of Δp. One illustration of this occurs when

determining whether blocking is influenced by "causal scenario" and "causal order". "Causal scenario" refers to whether participants are presented with two causes of one effect or one cause of two effects. "Causal order" refers to whether information about causes is presented before effects or effects before causes. Waldmann and Holyoak (1992) have revived the debate on this issue and have proposed a causal model theory to explain how mental models would influence blocking. After describing causal model theory, this section will review the literature comparing the predictions of Waldmann and Holyoak to those of the R-W model. The analysis will be followed by a series of experiments to further examine this issue.

According to Waldmann and Holyoak's (1992) causal model theory, participants use prior knowledge to guide their learning in a causal induction task. The authors make three important assumptions about how people reason in these tasks. The first assumption is that "people have a strong disposition to learn directed links from causes to their effects, rather than vice versa" (p. 224). This assumption arises from the empirical notion of a temporal asymmetry between causes and effects. That is, causes precede effects but the reverse does not occur. Experimental evidence suggests that children as young as 3 - 5 years old are aware of this asymmetry (Bullock & Gelman, 1979). Furthermore, it has been argued that participants can easily reason from cause to effect but often have trouble reasoning from effect to cause (Einhorn & Hogarth, 1986). Waldmann and Holyoak (1992) cite Tversky and Kahneman (1980) and Eddy (1982) to support this argument. Listed on the next page are two of the questions Tversky and Kahneman asked their participants and the number of participants that chose each answer.

Table 1

Experimental Designs and Results of Waldmann and Holyoak (1992)

Experiment 1

Group	Phase 1	Phase 2	*Test Question*	Results
2C-1E	P	PR	*?CE*	Blocking
2E-1C	P	PR	*?CE*	No Blocking

Experiment 2

Group	Phase 1	Phase 2	*Test Question*	Results
2C-1E	P	PR	*?CE*	Blocking
2E-1C	P	PR	*?EC*	Blocking

Experiment 3

Group	Phase 1	Phase 2	*Test Question*	Results
2C-1E	P	PR	*?CE*	Blocking
2E-1C	P	PR	*?EC*	No Blocking

The first column indicates whether two causes preceded one effect (2C-1E) or two effects preceded one cause (2E-1C). The letters in the Phase 1 and Phase 2 columns refer to the cues presented during these phases (P: Perfectly correlated cue; R: Redundant cue). The test question columns state whether participants were asked to reason from cause to effect (?CE) or from effect to cause (?EC) at the time of judgment. This variable is italicized to emphasize that it is the key difference in the design of the three experiments.

Problem 1: Which of the following is more probable?
(a) That a girl has blue eyes if her mother has blue eyes. (N = 69)
(b) That the mother has blue eyes, if her daughter has blue eyes. (N = 21)
(-) The two events are equally probable. (N = 75)

Problem 2: In a survey of high-school seniors in a city, the height of boys was compared to the height of their fathers. In which prediction would you have greater confidence?
(a) The prediction of the father's height from the son's height. (N = 23)
(b) The prediction of the son's height from the father's height. (N = 68)
(-) Equal confidence. (N = 76). (Tversky & Kahneman, 1980, pp. 51).

In each problem, participants rated the probability of predicting information about an offspring's characteristics based on a parent's as being more probable and reliable than predicting a parent's characteristics based on the offspring's, even though the probabilities of the two events are the same. Similarly, Eddy (1982) reports how doctors sometimes misinterpret the probability of having cancer given a positive test result [i.e., p (cancer | test)] with the probability of attaining a positive test result given that the patient has cancer [i.e., p (test | cancer)]. These findings imply that participants treat probabilities differently when doing problems involving knowledge retrieval. Furthermore, causal mechanisms were not assessed in Tversky and Kahneman's (1980) study. Hence, the implications of these results to how participants process information in a causal induction task are not clear. A causal induction task measures acquisition of new information, whereas problems such as the ones used in Tversky and Kahneman (1980) and Eddy (1982) measure how information is processed and retrieved after it has been acquired.

Waldmann and Holyoak's second assumption is that "the perceived strength of a causal connection is related to the *contingency* between the possible cause and the effect" (1992: p. 224; italics in original). This assumption is similar to the notion of covariation described in the PC model. Third, although the links in a causal model are directed from cause to effect, people can make both predictive and diagnostic judgments in which they have to reason from cause to effect and from effect to cause.

Waldmann and Holyoak (1992) report three experiments which they claim support this position. The experimental design and results are shown in Table 1. In each

experiment a predictive learning scenario was compared to a diagnostic scenario. In the predictive scenario, there were two causes and one effect, and participants were first informed about the causes and then about the effect (hereafter 2C-1E), while in the diagnostic scenario there were two effects and one cause, with the effects being presented before the cause (hereafter 2E-1C). This information was acquired over trials in a computer game.

The first experiment was an attempt to demonstrate that blocking would occur in 2C-1E but not in 2E-1C. In this experiment and the second experiment, the authors used a cover story in which the names of the antecedent cues were the same in both the 2C-1E and 2E-1C scenarios, but differed in whether they were defined as causes or effects. In 2C-1E, participants were told that a person's appearance could evoke an emotional response in an observer. Thus, they were informed about a person's appearance and had to guess about the observer's emotional response. In other words, they were asked how the cue "appearance" caused an "emotional reaction." In 2E-1C, the participants were told that a virus could affect a person's appearance. During the task, they were presented with patient files containing information about the patient's appearance and had to guess about the presence or absence of the virus. Here the cues were symptoms, or the effects of a disease. Hence, the antecedent cues in 2E-1C were identical to 2C-1E, except that they were labeled as outcomes rather than causes. The subsequent cues, however, were different in the two conditions.

There were three phases in each scenario. In the first phase, skin quality was a perfect predictor (P) of an outcome. Pale skin was always followed by the outcome whereas normal skin was never followed by the outcome. Two other irrelevant causes were also presented, but their role was less important and thus will not be discussed. The second phase was similar to the first phase, except an additional redundant cue (R) was added. This cue was weight; hence, people who had pale skin were underweight and people with normal skin had normal weight. In the test phase, participants in 2C-1E were asked how strongly each feature was a cause of the emotional response and those in 2E-1C were asked how strongly each cue was an effect of the virus (see p. 229; ?CE, as

shown in the last column of Table 1). By the authors' terminology, their participants were asked to perform a predictive inference in both the 2C -1E: ?CE and 2E - 1C: ?CE tasks, although it is not clear how the test question in 2E-1C asks people to reason from cause to effect.

According to Waldmann and Holyoak (1992), the R-W model would make the same prediction in both the 2C-1E: ?CE and 2E-1C: ?CE preparations. Namely, P would be rated highly and R would have a low rating since R would have no associative strength by the end of the trials. In other words, learning about P in phase one should block learning about R. The authors argue that this prediction arises because the two scenarios are logically symmetrical, but that associative models cannot encode the semantic distinction between causes and effects. Accordingly, the cues presented first, regardless of whether they are causes or effects, will be encoded as inputs that compete for associative strength with the output. The R-W model thus predicts that cue competition should arise in both scenarios. On the other hand, causal model theory predicts that there will be cue competition between P and R in the 2C-1E: ?CE condition but not in 2E-1C: ?CE because people learn the link from causes to effects but not from effects to causes, and because causes compete for control of effects but effects do not compete for control of causes. As shown in Table 1, P blocked ratings of R in 2C-1E: ?CE but not in 2E-1C: ?CE, which was interpreted as being consistent with causal model theory.

In the second experiment, the investigators attempted to replicate the first experiment with two changes to the experimental protocol. First, they changed the scale used during the test phase. More important, participants in the 2C-1E group were asked to rate how strongly each cue was a cause of the effect (i.e., predictive inference: ?CE) whereas those in the 2E-1C were asked "...questions intended to elicit diagnostic inferences (from effects to a hypothetical cause)" (Waldmann & Holyoak, 1992, pp. 230), although the exact wording of these questions is not stated in their paper. Nevertheless, the authors argue that participants in 2E-1C rated how strongly each cue was an effect of the cause (i.e., diagnostic inference: ?EC). The two groups were thus 2C-1E: ?CE and 2E-1C: ?EC. The cover story and the events that occurred on each trial were the same as their

first experiment. As shown in Table 1, participants gave similar ratings to P and to R in both the 2C-1E: ?CE and 2E-1C: ?EC tasks. That is, blocking was observed in both groups. However, the authors argued that the participants in the 2E-1C: ?EC had come into the experiment with existing causal models and that these models competed with the information presented in the experiment to produce blocking. Using Cheng's terminology, participants can be said to be using a focal set larger than that dictated by the experiment. Furthermore, the authors argued that these existing causal models influenced participants' ratings in the second experiment, but not in the first, because participants had to perform a diagnostic inference in the second experiment whereas they had to perform a predictive inference in the first one. However, since the authors did not independently assess the participants' existing causal models or prior knowledge, it is difficult to assess what causal models the participants might have had or how they might have influenced the judgments. Moreover, it is not clear whether the questions asked at the time of judgment elicited the type of inference the authors intended.

To circumvent the problem of prior experience influencing participants' performance on the experimental task, the authors attempted to create a neutral causal scenario in which the cues were whether certain buttons or lights were on or off, and whether a security alarm system was operating. In both the 2C-1E: ?CE and 2E-1C: ?EC conditions, participants were first told which buttons were pressed (or which lights were on) and were then informed as to whether the alarm was operating. In 2C-1E: ?EC condition, the buttons were the causes and the alarm was the effect, whereas in 2E-1C: ?EC the lights were the outcomes and the alarm was the cause.

It was hypothesized that participants would have few prior notions regarding the cues and outcomes used in the experiment and thus the participants' focal set would be restricted to the cues presented in the experiment. Without putative alternative causal models competing with the information displayed, Waldmann and Holyoak (1992) predicted that there would be blocking in the predictive but not the diagnostic condition. This result was obtained and was interpreted to support the claim that participants used wider focal sets in the second experiment than in the third experiment. However, this

argument is circular since observing blocking confirms the presence of alternative causal models and failing to observe blocking confirms their absence. Because of its circularity, the argument may be unfalsifiable (Popper, 1965).

In summary (Table 1), in the 2C-1E: ?CE condition, blocking was observed in all three experiments. However, in the 2E-1C condition, the authors reported reliable blocking in the second experiment, but not in the first and third. Unless one accepts their untested assumptions about the role of alternative causal models, the evidence is at best only somewhat consistent with causal model theory. However, the authors conclude that their data "clearly refute connectionist learning theories that subscribe to an associationist representation of events as cues and responses" (p. 233) and speculate that "lower-order associative learning should be reduced to higher order causal induction" (p.235).

A number of methodological and theoretical issues need to be addressed before this conclusion can be accepted. The first issue is that the causal scenarios were confounded with causal order. That is, in the predictive (2C-1E) condition, there was always two causes and one effect, with the causes being presented first, while in the diagnostic (2E-1C) condition, there was one cause and two effects, with the effects being presented first. In none of the experiments were there conditions in which there were two causes and one effect with the effect being presented first (i.e., 1E-2C) or one cause and two effects with the cause presented first (1C-2E).[2] Second, all three experiments lacked an appropriate control group. Traditionally, in experiments investigating blocking (Kamin, 1969), the learning in the blocking group is compared to a group that has not undergone the blocking procedure. This control procedure thus confirms that blocking has occurred. However, instead of using a control group or control condition, Waldmann and Holyoak (1992) assessed blocking by comparing ratings of P with R. Since participants experienced more trials with P, and therefore more pairings of P with the

[2] Blocking in these two conditions, however, has been investigated. For brevity and clarity, this section will focus on 2C-1E and 2E-1C, the conditions studied most thoroughly. Discussion of 1C-2E and 1E-2C will be withheld until the introduction to Experiments 3 and 4.

outcome, any differences between P and R could be attributed to differential experience with the two cues (Shanks & Lopez, 1996). Third, the cues in their experiments were not counterbalanced. Fourth, the authors argue that existing causal models accounted for the opposing results in their first two experiments. But without an independent assessment of these causal models, any arguments about how they would influence their participants can only be speculation. Fifth, the authors stress the importance of contingency in causal inference tasks. But in their experiments, the outcome always occurred in the presence of P and R but never in their absence. Participants also did not experience trials in which outcomes occurred in the absence of the cues. Thus, Waldmann and Holyoak (1992) did not use all of the cells in a contingency table. Last, a critical assumption in causal model theory is that people learn from causes to their effect and not vice versa. But this assumption was never tested directly in any of the three experiments.

Since Waldmann and Holyoak's (1992) findings were claimed to be at variance with associative accounts, a number of experiments have been conducted to test this model. In 1993, Van Hamme, Kao, and Wasserman tested whether competitive asymmetry would be observed between 2C-1E and 1E-2C. This experiment differed from Waldmann and Holyoak's (1992) in three important ways; these differences and a critique of this experiment will be presented before discussing the results and their implications.

First, Van Hamme et al. (1993) used the relative validity paradigm of Wagner et al. (1968) instead of Kamin's blocking (1969) model. Similar to Wagner et al.'s (1968) original study, the relative predictiveness of a stimulus, X, was determined by the relative predictability of two other stimuli (A and B) presented in compound with X. Thus, as opposed to Waldmann and Holyoak's study that employed two cues that could be defined as causes or effects, this study used three cues. Their two conditions were 3C-1E and 3E-1C. However, unlike Wagner et al.'s (1968) original experiment, this experiment used a range of contingencies. These conditions are summarized in Table 2. At one extreme, the outcome was always present when X was presented with A [i.e., p (outcome | AX) = 1] but was absent when X was presented with B [i.e., p (outcome | BX) = 0]. Thus the difference in the probability of an outcome after combining the two probabilities was 1

Table 2

Design and Results of Van Hamme et al. (1993)

p (outcome | AX - BX)	Cover Story	Test Question	Results (Ratings of X)
0	3C-1E	?CE	Ratings for X decrease as p (outcome | AX - BX) increases.
0.25	3C-1E	?CE	
0.5	3C-1E	?CE	
0.75	3C-1E	?CE	
1	3C-1E	?CE	
0	3E-1C	?EC	Ratings of X do not change as p (outcome | AX - BX) increases.
0.25	3E-1C	?EC	
0.5	3E-1C	?EC	
0.75	3E-1C	?EC	
1	3E-1C	?EC	

Cover story
3C -1E: 3 causes - 1 effect
3E -1C: 3 effects - 1 cause

Test Question
?CE: cause to effect
?EC: effect to cause

[i.e., p (outcome | AX-BX) = 1]. At the other extreme, the outcome was present on half of the AX trials and half of the BX trials [i.e., p (outcome | AX) = 0.5, p (outcome | BX) = 0.5, p (outcome | AX-BX) = 0]. Between these two extremes, the AX-BX relationship varied from 0 to 1 in steps of 0.25. That is, the probabilities for an outcome given AX and BX, respectively, were 0.625 and 0.375 (AX - BX = 0.25), 0.75 and 0.25 (AX - BX = 0.50), and 0.875 and 0.125 (AX - BX = 0.75). Thus there was a total of five possible contingencies.

Second, this experiment differed from Waldmann and Holyoak's (1992) in terms of the cover stories and cues employed. Whereas Waldmann and Holyoak (1992) used the same cues in both the 2C-1E: ?CE and 2C-1E: ?EC conditions but changed only whether the cues were defined as causes or effects, Van Hamme et al. (1993) used different cues and cover stories in their 3C-1E and 3E-1C conditions. In the 3C-1E condition, the three cues defined as causes were foods patients had eaten (X was shrimp, A was strawberries, and B was peanuts) and the outcome was whether an allergic reaction occurred. On the other hand, in the 3E-1C condition, the participants were told a patient had either eaten on not eaten shrimp and that the symptoms that occurred were recorded (X was headache, A was fever, and B was rash). For this preparation, shrimp consumption was defined as the cause and the symptoms were defined as the effects. At the test phase, participants in the 3C-1E: ?CE condition were asked to rate the degree to which each food was a cause of the allergic reaction (i.e., ?CE question) whereas participants in the 3E-1C: ?EC condition were asked to rate how strongly each symptom was an effect of eating shrimp (?EC question).

Third, this experiment differed from Waldmann and Holyoak's (1992) in terms of how the contingencies were presented. In Waldmann and Holyoak's (1992) study, participants acquired the data over trials, whereas participants in Van Hamme et al.'s (1993) study were provided with a sheet of paper containing all of the results. That is, in the 3C-1E: ?CE condition, the list contained 16 rows, each stating which foods were consumed and whether the allergic reaction had occurred. And in the 3E-1C: ?EC

condition, each row stated whether shrimp had been consumed and which symptoms had occurred.

Before discussing the results and their implications, the problems with this experiment will be addressed. First, inspection of Van Hamme et al.'s (1993) Table 1 reveals that the participants were not provided with all of the information in a contingency table. Namely, in the 3C-1E: ?CE condition, there were no "trials" in which all causes were absent. Conversely, in the 3E-1C: ?EC condition, there were no "trials" in which no symptoms occurred. Without using all of the cells in a contingency table, it is difficult to make any firm conclusions about how participants reason about contingencies.

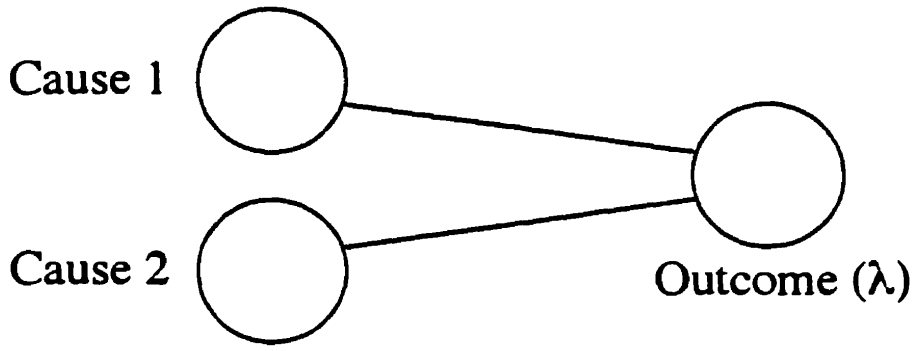
Second, the methodology in this experiment was dramatically different from that of Waldmann and Holyoak's (1992) in that different cues and cover stories were used in the 3C-1E and 3E-1C conditions. More important, however, is that participants were given the information on a summary list. With this method of presentation, it seems reasonable to assume that participants based their response on calculations performed on the information in the list rather than building associative strength to the cues. An important assumption of the R-W model is that associative strength is incremented over trials. Since the participants did not receive the contingencies over trials, this experiment did not evaluate the R-W model per se.

Furthermore, with the information presented in a list, it is difficult to state the causal order in this experiment or to even speculate on how this information could be represented. For example, in Van Hamme et al.'s (1993) 3E-1C: ?EC condition, participants were told that patients had either eaten or not eaten shrimp and had then recorded which symptoms occurred. It could be argued that participants had represented the scenario as 1C-3E instead of 3E-1C since the events in the cover story occurred in the order cause-effect. Finally, this experiment shares two problems with Waldmann and Holyoak's (1992). Causal order was confounded with whether the scenarios had multiple causes and one effect or one cause and multiple effects, and the names of the cues were not counterbalanced .

Despite the methodological differences between these two studies, Van Hamme et al. (1993) obtained results consistent with Waldmann and Holyoak (1992), but interpreted them differently. In both the 3C-1E: ?CE and 3E-1C: ?EC scenarios of Van Hamme et al. (1993), ratings of A increased as A's contingency increased. At the same time, ratings for B decreased. Ratings for X, however, decreased as the correlation difference between AX and BX increased in the 3C-1E condition, but not in the 3E-1C condition (see Table 2). The authors concluded that this finding was consistent with Waldmann and Holyoak's (1992) suggestion that causal cues compete for associative strength while effect cues do not. However, they reasoned that this asymmetry would not pose a problem for associative accounts. In particular, they argue that, from an evolutionary perspective, it would be unadaptive for a reliable predictor to compete with less reliable predictors for associative strength. On the other hand, if a cause is predictive of more than one effect, it would make little sense for the effects to compete with one another; instead, an adaptive response would be for each effect to be associated with its most reliable predictor.

Similarly, Baker et al. (1996) have argued that the difference in results can be accounted for by the Rescorla - Wagner model if one compares two one-layer connectionist networks (Figure 2). In the first network, which is analogous to Waldmann and Holyoak's (1992) predictive scenarios, two input nodes are linked to one output node. Cue competition would be expected in this network since there are two associative links to the output. Since the output can support only a limited amount of associative strength (λ), the two inputs would come to share the maximal associative strength that the outcome can support. On the other hand, in the second network, there is one input node and two output nodes (λ_1 and λ_2). This network is analogous to Waldmann and Holyoak's (1992) diagnostic scenarios. The input is linked to each output by a single associative link. Since each output node is able to support different amounts of associative strength, cue competition would not be expected with this network. Thus, according to Van Hamme et al. (1993) and Baker et al. (1996), the R-W model can accommodate the results of Waldmann and Holyoak (1992).

2 Causes Predict One Outcome



$$\lambda = V_{(\text{Cause 1})} + V_{(\text{Cause 2})}$$

One Cause Predicts Two Outcomes

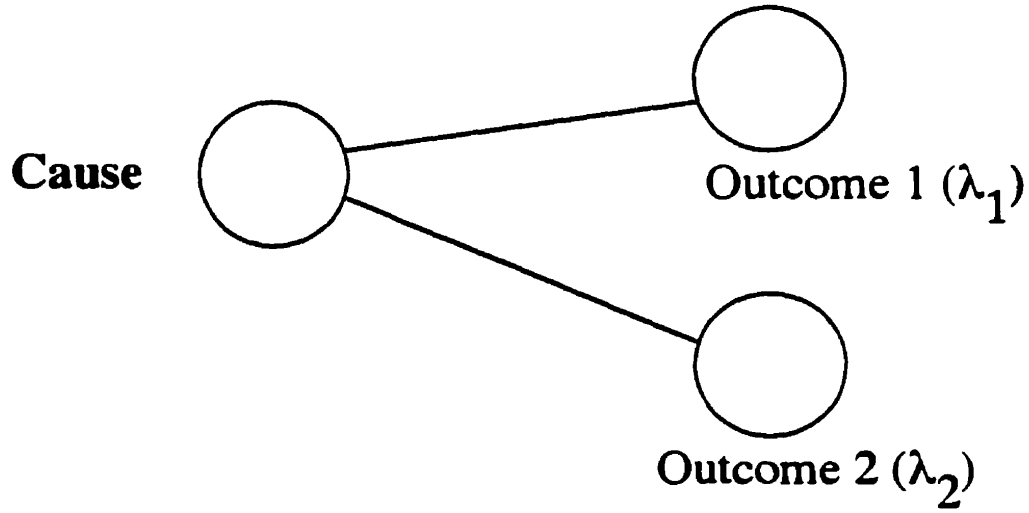


Figure 2. The model of Baker et al. (1996) displaying the assumed associations formed when two causes predict one outcome and when one cause predicts two outcomes.

Further support for the R-W model comes from three experiments reported by Shanks and Lopez (1996). In Experiment 1 (see Table 3A and 3B), three factors were manipulated: the first two were between-subjects factors and the third was a repeated measures variable. First, multiple causes preceded a single effect (MC-SE) or whether multiple effects preceded a single cause (ME-SC). In MC-SE, participants evaluated the extent to which certain foods were associated with an allergic reaction, and in ME-SC they judged how strongly different symptoms were caused by a disease.

The second factor was whether the cues were concrete or abstract. This manipulation was performed in an attempt to manipulate focal sets and the degree to which prior knowledge would influence whether blocking was observed. In the concrete version of MC-SE, the causes and effects were named after foods and symptoms that occur in the real world (e.g., avocados or eggs were causes; puffy eyes or upset stomach could be effects). Similarly, in the concrete version of ME-SC, the effects were symptoms such as an upset stomach, and the causes were fictitious diseases (e.g., Marshall-Isaacs Disease, Phipp's Syndrome). In the abstract version of MC-SE and ME-SC, the cues were labeled with letters and numbers (e.g., food A, allergy reaction 1 in MC-SE; symptom A, Disease 1 in ME-SC).

The third variable was whether the relationships among the cues and outcomes were defined as contingent or non-contingent. As shown in Table 3B, this experiment adapted Kamin's (1969) blocking paradigm. In MC-SE, Cues A and C were paired with an outcome the same number of times. Cue A was defined as non-contingent with 1 because it was a poor predictor of 1 and had an unconditional Δp of zero. That is, outcome 1 occurred regardless of whether A was present or absent. Cue A's contingency was calculated by

$$\Delta p\,(A) = p\,(1 \mid A) - p\,(1 \mid \text{no } A) = 1.0 - 1.0 = 0.$$

Cue C, on the other hand, was defined as contingent because it was a perfect predictor of 2. That is, 2 occurred only in C's presence but never in its absence:

$$\Delta p\,(C) = p\,(2 \mid C) - p\,(2 \mid \text{no } C) = 1.0 - 0.0 = 1.0.$$

Table 3

Design of Shanks and Lopez (1996)

A. Causal Scenario and Causal Order (Experiment 1)

Condition	Test Question
MC-SE: Concrete	"On a scale of 0 to 100, how strongly is [cue] associated with [outcome]?" (p. 516)
MC-SE: Abstract	
ME-SC: Concrete	
ME-SC: Abstract	

B. Contingent and Non-Contingent Trial Types (Experiment 1)

Conditions	Events	Δp (in MC-SE)
Non-Contingent	**AB** => 1, B => 1	0.0
Contingent	**CD** => 2, D => 0	1.0

A to D represent causes in MC-SE and outcomes in ME-SC. A and C are presented in boldface to emphasize that judgments of these two cues were compared. The numbers 1 and 2 represent outcomes in MC-SE and causes in ME-SC. 0 represents a no-outcome trial in MC-SE and a no-cause trial in ME-SC.

C. Contingent and Non-Contingent Trial Types (Experiment 3)

Conditions	Cues => Outcomes	Δp
Non-Contingent	**AB** => 1, B => 1,C => 0	0.5
Contingent	**DE** => 2, E => 0, F => 2,	0.5

Experiment 3 used ME-SC: Abstract. Cues A to D represent effects and the numbers 1 and 2 represent causes. 0 represent a no-cause trial. The cues in boldface represent the cues that were compared.

In this design, the non-contingent condition corresponded to the experimental condition of a blocking experiment and the contingent condition served as the control. Blocking was observed if judgments of C were higher than those of A.

The procedure of Shanks and Lopez (1996) was similar to Waldmann and Holyoak (1992) in that the contingencies were presented over a series of trials in a computer-based game. At the end of the game, participants rated how strongly each cue was associated with each outcome. Shanks and Lopez (1996), however, used a larger sample size than Waldmann and Holyoak (1992) and intermixed the contingent and non-contingent trials rather than presenting them in separate blocks. With this preparation, blocking was observed in all four conditions. That is, the contingent cues were rated higher than the non-contingent cues regardless of whether causes were presented before effects or effects before causes, and regardless of whether the cues were defined as concrete or abstract. This finding was interpreted as supporting the R-W model.

Experiment 2 was conducted in an attempt to replicate the blocking effect in ME-SC using a different cover story. In this experiment, only the ME-SC: Abstract condition was tested. The effects were whether indicator lights were on or off and the cause was the occurrence or non-occurrence of a problem in a chemical plant. Similar to Experiment 1, the contingent cues were rated higher than the non-contingent cues. This blocking effect was consistent with the predictions from the R-W model.

In Experiment 1, the Δp for the non-contingent and contingent cues were 0 and 1 in MC-SE, and it was assumed that these contingencies were the same in ME-SC. However, an analysis of the contingencies revealed that the Δp for the non-contingent and contingent cues in ME-SC were 0.5 and 1, rather than 0 and 1. This difference arose because, according to causal model theory, participants in ME-SC might have interpreted the AB => 1 and B => 1 trials as AB <= 1 and B <= 1. In other words, (effect 1, effect 2) <= cause 1, and (effect 2) <= cause 1. In this case, the Δp for the non-contingent cue is 0.5 rather than 0. That is,

$$\Delta p \text{ (cause 1)} = p(A \mid 1) - p(A \mid \text{no } 1) = 0.5 - 0.0 = 0.5.$$

On the other hand, the Δp for the contingent relationship was maintained at 1. If one assumes that participants interpreted CD => 2 and D => 0 as CD <= cause 2 and D <= no cause, then

$$\Delta p \text{ (cause 2)} = p \text{ (C | 2)} - p \text{ (C | no 2)} = 1.0 - 0.0 = 1.0 .$$

Experiment 3 was conducted in an attempt to control for this confound. This experiment used a ME-SC: Abstract preparation and the same cover story used in Experiment 1. The Δp in the non-contingent and contingent conditions, were equated by using additional cues and changing the frequency of the events associated with each cue; at the same time, A remained an unreliable predictor of 1, and D was a better predictor of 2 than A (see Table 3C).

In the non-contingent condition, the events were AB => 1, B => 1, C => 0. Under the assumption that A, B, and C were effects, causal model theory would predict that the events were interpreted as AB <= cause 1, B <= cause 1, and C <= no cause. In this case, the Δp between 1 and A would be 0.5 because A occurs half of the time when 1 is present and never occurs when 1 is absent. That is,

$$\Delta p \text{ (cause 1)} = p \text{ (A | 1)} - p \text{ (A | no 1)} = 0.5 - 0.0 = 0.5$$

In the contingent condition, the events were DE => 2, E => 0, and F => 2. Assuming that the events were interpreted as DE <= cause 2, E <= no cause, and F <= cause 2, the contingency between 2 and D would be 0.5 because D occurs half of the time when 2 is present and never occurs when 2 is absent. In other words,

$$\Delta p \text{ (cause 2)} = p \text{ (D | 2)} - p \text{ (D | no 2)} = 0.5 - 0.0 = 0.5$$

The Δp between the cue and outcome was thus the same in the non-contingent and contingent conditions.

However, causal model theory and the R-W models make opposite predictions about the experimental outcome. According to causal model theory, blocking would not be expected because the contingencies are the same in both conditions. On the other hand, the R-W model would predict that blocking should occur because cue D remains a better predictor of the outcome than A. In the non-contingent condition, A would accumulate

little associative strength. And in the contingent condition, cue D would gain more associative strength than A despite the presence of cue F competing for associative strength with D (see also Shanks, 1991).

In Experiment 3, the contingent cues were rated higher than the non-contingent cues. Cue competition was therefore observed in all three experiments. To summarize, blocking was attained in Experiment 1 regardless of whether the task was MC-SE or ME-SC and regardless of whether the cover stories used abstract or concrete cues. In Experiment 2, blocking occurred in ME-SC: Abstract with a cover story different from that used in Experiment 1. And in Experiment 3, blocking was observed in ME-SC: Abstract even though the contingencies for the contingent and non-contingent cues were the same. These findings were anticipated by the R-W model but not causal model theory; the authors thus concluded that their results did not rule out an associative account.

However, Waldmann and Holyoak (1997) criticized Shanks and Lopez's (1996) experiments as being inadequate for testing causal model theory. The first main objection was that Shanks and Lopez (1996) did not clearly state which variables were causes and which ones were effects. For example, they argue that the symptoms in ME-SC may have been interpreted as causes rather than effects ("e.g., puncture wounds may be the cause of blood poisoning, rather than the reverse", p. 128; however, these cues were not used by Shanks and Lopez, 1996), and that a disease is not always a cause but instead "may simply name a collection of symptoms that constitute a syndrome." (p. 128). Waldmann and Holyoak (1997) furthermore argue that this ambiguity was pronounced in the ME-SC: Abstract condition of Experiments 1 and 3, where the cues were letters and numbers. In this condition, participants were not provided with an appropriate context in which they could evaluate causal relationships, and thus had no way of knowing which cues were causes and which ones were effects. Without an appropriate causal context, Waldmann and Holyoak (1997) suggest that participants in ME-SC: Abstract treated the task as a simple associative task in which they learned contingencies between arbitrary stimuli. This argument, however, is based on speculation. The only way to assess whether

cues are interpreted as effects and causes is to directly measure whether the cues are interpreted correctly.

Another criticism of Shanks and Lopez (1996) is that they failed to manipulate causal order. In particular, Waldmann and Holyoak (1997) argue that to properly assess cue competition between causes and effects, the experiment must be designed such that

> (1) participants who are required to give diagnostic judgments do not believe there are multiple alternative causes of the diagnostic signs and (2) all participants in each condition consistently interpret the causal relation in a single direction (either cause-effect or effect-cause). (Waldmann & Holyoak, 1997, p. 130)

Similar to the first criticism, this argument is speculative. Furthermore, it also applies to Waldmann and Holyoak (1997) because the authors failed to measure their participants' beliefs about alternative causes in their 2E-1C, and at no point did they demonstrate that participants consistently interpreted the causal model as either cause-effect or effect-cause.

Waldmann and Holyoak's (1997) third objection to the methodology of Shanks and Lopez (1996) has to do with the wording of their rating scale. Shanks and Lopez (1996) asked participants to rate how strongly a cue was associated with an outcome, whereas their figures were labeled as being ratings of predictiveness. Waldmann and Holyoak (1997) argued that the two measures are not semantically equivalent. In other words, there was a discrepancy between what the participants were asked and how their ratings were interpreted. Unfortunately, the strength of this argument can not be assessed because the authors neither state the exact wording of the questions used in their original experiments nor provide any data to demonstrate how participants interpret the two types of questions.

To summarize, Waldmann and Holyoak (1997) have argued that Shanks and Lopez (1996) have failed to 1) unambiguously state which variables were causes and effects, 2) demonstrate that the causal model was interpreted as cause-effect or effect-cause, 3) control for alternative causes in ME-SC, and 4) use an appropriate rating scale. In each case, the criticism is based on speculation and conjecture because Waldmann and

Holyoak (1997) failed to directly measure the participants' interpretation of the cover story or of the rating scale. Their review thus does not detract from Shanks and Lopez's (1996) findings and interpretation, which were consistent an associative account.

Matute, Arcediano, and Miller (1996), on the other hand, have obtained results that support neither causal model theory nor the R-W model. Based on animal research done in Miller's laboratory (Esmoris-Arranz, Miller, & Matute, 1997), Matute et al. (1996) hypothesized that it does not matter whether causes are presented before their effects or whether effects are presented before their causes during acquisition. However, the authors argued that when participants are asked about whether one variable is the cause of the other or when they are asked about whether one variable is the effect of the other, they interpret the two questions as being the same. Furthermore, the authors hypothesized that both these questions asked about the probability of the effect given the cause implicitly compared with the effect given an alternative cause. Thus, if there are multiple causes in the experiment, cue competition would be expected using both of these questions; on the other hand, if there is only one cause present, then neither question should yield cue competition. Furthermore, Matute et al. (1996) hypothesized that participants should be able to accurately report the co-occurrence of the cause and effect, regardless of the wording. In other words, when asked about the contiguity of events, participants would not expected to exhibit cue competition.

To test these hypotheses, Matute et al. (1996) conducted three experiments. All three used the relative validity paradigm of Wagner et al. (1968). In the first experiment, participants were presented with three causes and one effect (3C-1E); in the second experiment, they were presented with three effects and one cause (3E-1C); and in the third experiment both of these conditions were included. The experimental designs are summarized in Table 4. These experiments were an adaptation of the experimental preparation used by Van Hamme et al. (1993), except that different names were assigned to the cues and different test questions were asked. Since the experimental preparation was similar, this study is subject to the same criticisms as Van Hamme et al., which will be summarized after describing the results and conclusions.

Table 4

Design and Results of Matute et al. (1996)

Experimental Condition: AX^{+}, BX^{-}
Control Condition: $AX^{+/-}$, $BX^{+/-}$

Condition	Test Question	Results
Experiment 1		
3C-1E	1. ?CE Causal (i.e., is C the cause of E)	Cue Competition
	2. ?EC Causal (i.e., is E the effect of C)	Cue Competition
	3. ?CE Contiguity (i.e., when C is present, does E cooccur)	No Cue Competition
	4. ?EC Contiguity (i.e., when E is present, does C cooccur)	No Cue Competition
Experiment 2		
3E-1C	1. ?CE Causal	No Cue Competition
	2. ?EC Causal	No Cue Competition
	3. ?CE Contiguity	No Cue Competition
	4. ?EC Contiguity	No Cue Competition
Experiment 3		
3C-1E	1. Indicator	Cue Competition
	2. Contiguity	No Cue Competition
3E-1C	1. Indicator	Cue Competition
	2. Contiguity	No Cue Competition

?CE: cause to effect test question
?EC: effect to cause test question

In the first experiment (3C-1E), Matute et al. (1996) hypothesized that there would be cue competition when the questions implicitly encouraged participants to compare causes (see Table 4). Cue competition would not be observed in the second experiment (3E-1C), though, since there was only one possible cause. In both experiments, it was expected that questions about contiguity would not yield cue competition. In the first experiment, cue competition was observed for X when participants were asked causality test questions but not contiguity questions. In the second experiment, no cue competition was observed in any of the conditions. Thus the authors concluded that the ?CE and ?EC causality questions were interpreted as the same question.

In the third experiment, cue competition was observed in 3E-1C if the question was worded so that, according to the authors, it implicitly asked participants to compare the probability of a cause given an effect to the probability of the cause given alternative effects. For example, one such question asked "Is taking [medicine's name] *indicative* that the allergic reaction is going to appear?" (Matute et al., 1996: p. 189; italics added). When the question was worded this way, cue competition was found between effects. However, cue competition was not observed with contiguity questions. Hence, the authors concluded that the specific type of question was crucial in determining whether cue competition is observed.

However, the authors' basis for their assertions about the interpretation of the test questions is not clear. For example, it is not apparent how Matute et al. (1996) know that causality questions implicitly encourage participants to compare causes, whereas using the word "indicative" implicitly asks participants to evaluate a cause given alternative effects. The authors neither provide any references from linguistics to support their hypothesis, nor provide any evidence that participants interpreted the questions in the manner suggested. Matute et al. (1996) also contradict themselves about the meaning of contiguity test questions. For example, they argue that contiguity questions ask merely about the co-occurrence of events yet concurrently claim that these questions are synonymous with conditional probabilities of p (E | C) and p (C | E). It is difficult to

imagine how the contiguity questions can simultaneously ask about merely the number of pairings and about conditional probabilities. These are two different issues.

Apart from these difficulties, it is not clear how their arguments about test questions would apply to other experiments investigating causal reasoning. This notion will be clarified with two examples. First, as mentioned, Shanks and Lopez (1996) asked participants how strongly each cue was associated with each outcome and were criticized for arguing that this question was synonymous with questions concerning causal relations (Waldmann & Holyoak, 1997). But according to Matute et al. (1996) , Shanks and Lopez (1996) asked causality test questions. Since both Waldmann and Holyoak's (1997) and Matute et al.'s (1996) arguments rely on conjecture, it cannot be stated which party has a stronger position. The second example concerns other experiments in which participants are not asked ?CE questions corresponding to "Is A the cause of B?". For example, Baker asked participants to "estimate the effectiveness" (Baker et al., 1993, p. 94) of the cue on the outcome. It seems reasonable to question how Matute et al. (1996) would classify this type of test question. Would it be a causality ?CE question? And does this type of test question also ask participants the probability of the outcome given the cause, implicitly compared with alternative causes?

Finally, as mentioned, this study is also open to criticism for the same reasons as Van Hamme et al (1993). First, causal order is confounded with causal scenario. Second, since different cues were used in the 3C-1E and 3E-1C conditions, these experiments do not achieve Waldmann and Holyoak's (1992) ideal design in which the same cues are used in both conditions but differ only in whether they are defined as causes or effects. Third, since the data were presented on a list, this preparation does not evaluate the R-W model in which associative strength is gained or lost over trials. It is also difficult to specify whether participants acquire the information in a cause-effect or effect-cause causal order. Last, the participants were not exposed to all of the information in a contingency table since there were no "trials" in which the cues were absent. Moreover, one way to control for alternative causal models would be to specify what happens in the absence of the trials. Aside from these problems, however, the interesting finding from

these experiments was that participants' ratings were not influenced by whether the questions were worded ?CE or ?EC when they were asked questions about causality.

To summarize, a number of researchers have attempted to test whether participants reason differently when there are multiple causes of a single effect as opposed to when there are multiple effects of a single cause. In all of the experiments reviewed, cue competition was observed when there were multiple causes of a single effect, and the causes were presented before the effects. Waldmann and Holyoak (1992) have suggested that cue competition should not be observed when there are multiple effects of a single cause, and information about the effects are presented before the cause, whereas the R-W model would predict that cue competition should be observed under these conditions. However, in only one of their experiments did the authors clearly demonstrate that cue competition could not occur under these conditions. Follow-up studies have been done, but none adequately compare the R-W model to Waldmann and Holyoak's (1992) causal model theory. For example, none of these experiments used cover stories in which the cues were the same but differed only with regard to whether they were causes or effects. Furthermore, Van Hamme et al. (1993) and Matute et al. (1996) did not manipulate the presentation of the causal order (since the participants received the data on a list), nor did they present all of the information in a contingency table (i.e., the participants never observed what happened in the absence of the cues). Few of the experiments have manipulated causal order, and when it was manipulated, this factor was confounded with whether there were multiple causes of a single effect or multiple effects of a single cause. Last, Shanks and Lopez (1996) attempted to manipulate the participants' causal models and focal sets, but found that these manipulations did not influence cue competition.

Taken together, none of these experiments yields results that would rule out an associative model of causal learning. Nor do they strongly support Waldmann and Holyoak's (1992) causal model theory. To further evaluate causal model theory and to better understand the role of cue competition in predictive and diagnostic reasoning, causal scenarios (multiple causes- single effect; multiple effects-single cause) must be

manipulated independently of causal order. Furthermore, participants must be exposed to all of the information in a contingency table. That is, participants should observe what happens in both the presence and absence of the cues.

An Empirical Project

The following experiments were conducted to accomplish this goal. That is, cue competition was assessed in causal scenarios having either two causes of one effect or one cause of two effects. In addition, causal order was investigated independently in each of the two causal scenarios. Thus, there were four experimental conditions: two causes of one effect with the causes presented before the effect (2C-1 E), two causes of one effect with the effect presented before the causes (1E-2C), one cause of two effects with the cause presented before the effects (1C-2E) and one cause of two effects with the effects presented before the cause (2E-1C). Furthermore, the cues were two chemicals and one strain of bacteria and each could be defined as either causes or effects. In the 2C-1E and 1E-2C conditions, the chemicals were the causes and the bacterial strain was the outcome. On the other hand, in 1C-2E and 2E-1C conditions, the bacterial strain was the cause and the chemicals were the outcome. Scenarios in which the chemicals acted on the bacteria were used in Experiments 1, 4, 5 and 6, while scenarios in which the bacteria produced the chemicals were used in Experiments 2 and 3. In Experiments 1 and 3 participants were first informed about the cause(s) and then about the effect(s), whereas in Experiments 2, 4, 5 and 6, this information was received in the reverse order. In Experiment 7, all four conditions were tested in the same experiment. The experimental designs are summarized in Table 5. The rest of this section provides a brief overview of the studies in this project.

The first experiment was an attempt to replicate cue competition in 2C-1E using a cover story in which the chemicals were the two causes and the survival of the bacteria was the outcome. Experiment 2 (2E - 1C) was an attempt to test whether cue competition would be observed when the causal scenario was changed and the causal order was reversed. Thus, the chemicals were the effects and the bacteria was the cause, and effects

Table 5

Design of Experiments 1 to 7

	Cover Story	Test Question
Experiment 1	2 C - 1 E	?CE
Experiment 2	2 E - 1 C	?EC
Experiment 3	1 C - 2 E	?CE
Experiment 4	1 E - 2 C	?EC
Experiment 5	1 E - 2 C	?CE
Experiment 6	1E - 2C	?CE
Experiment 7	Group CE: 1 C - 2 E, 2 C - 1 E Group EC: 1E - 2C, 2E - 1C	?CE ?CE

were presented before the cause. According to the R-W model, cue competition should occur in both Experiments 1 and 2 since the model can not encode causes or effects, but merely the order in which information is presented. Since two cues are presented first, they should compete to be linked with the consequent; hence cue competition should be observed. On the other hand, according to causal model theory (Waldmann & Holyoak, 1992), cue competition should be observed in Experiment 1 but not in Experiment 2. The authors reason that causes should compete to be linked with an effect, but that effects should not compete to be linked with a cause. However, there is a caveat. Cue competition should be observed in Experiment 2 if alternative causal models are not controlled for. Since we could not state a priori what these models might be or devise a way of independently assessing them, we diminished their influence by stating in the instructions that steps were taken to minimize alternative causal factors. If cue competition was not observed, then the results would be consistent with Waldmann and Holyoak's (1992) model. On the other hand, if cue competition was observed, it would be consistent with the R-W model. When taken in conjunction with the results of Shanks and Lopez (1997), such an observation would provide converging evidence against the notion that cue competition occurs in a 2E-1C condition merely because of alternative causal models or not comprehending the cover story.

To compare the predictions of simple associative networks to causal model theory, Experiments 3 and 4 used treatments similar to those of Experiments 1 and 2. In Experiment 3 (1C-2E), the bacteria was the cause and the chemicals were the two effects, with the cause presented before the effects. On the other hand, in Experiment 4 (1E-2C), the causal order was reversed. Thus, the effect was presented before the causes. According to the model of Baker et al. (1996), the bacteria would be the input to the associative network while each of the two chemicals would be the outputs. Since each output can support its own amount of associative strength, cue competition would not be expected in either experiment. On the other hand, according to Waldmann and Holyoak (1992), cue competition should not be observed in Experiment 3 but should be observed in Experiment 4. In Experiment 3, there is only one cause and that cause is presented

before the effects. Since effects do not compete to be linked with a cause, cue competition would not be expected. On the other hand, in Experiment 4, there are two causes of one effect. However, even though the information is presented as effect-causes, participants should naturally reason from cause to effect. Consequently, since causes compete, causal model theory predicts that cue competition should be observed.

Three follow-up experiments were conducted to test the generality of the findings of the previous experiments and to rule out some methodological issues. In Experiment 1 and 3, participants were asked a ?CE question when they had to give their ratings and in Experiment 2 and 4, participants were asked an ?EC question. These questions corresponded to the causal order in which the contingencies were presented. To test Matute et al.'s (1997) assumption that causality ?CE and ?EC questions are interpreted as being the same question, in Experiments 5 and 6 the contingencies were presented as 1E-2C and participants were asked a ?CE question at test. Another motivation for this experiment was to independently assess a) whether participants encoded the causes and effect appropriately when they were presented in the reverse order, and b) to collect a list of potential alternative causes. In Experiment 6, we attempted to replicate the results of Experiments 4 and 5 using a modified task structure. Again, participants were asked about alternative causes. In Experiment 7, all four experimental conditions (2C-1E, 2E-1C, 1C-2E, and 1E-2C) were used in the same experiment, and participants were informed that various alternative causes (i.e., those listed in Experiments 5 and 6) were accounted for and could not explain the events on each trial. This step was taken in an attempt to rule out the possibility that blocking in 2E-1C could be explained by alternative competing causal models. Causal order was the between-subjects variable and causal scenario was the within-subjects variable.

OVERVIEW OF EXPERIMENTS

Experimental Conditions

The experimental paradigm used in the present series of studies was adapted from Experiment 1 of Baker et al. (1993). In Baker et al. (1993), participants played a video game in which a tank went into a minefield and could either traverse it safely or explode. On each trial, the volunteers could either camouflage or not camouflage the tank. The camouflage could increase, decrease, or have no effect on whether the tank passed through the minefield safely. In addition, a spotter plane appeared on some trials, which could increase, decrease or have no effect the tank's safety. For the experiments in this dissertation, the cues were two chemicals and one strain of bacteria instead of a tank, plane and minefield. In 2C-1E and 1E-2C, the chemicals were the causes and the bacterial strain was defined as the effect, whereas in 1C-2E and 2E-1C the bacterial strain was the cause and the two chemicals were the outcomes. The details of the cover story used in each condition will be explained in the individual experiments (see *Cover Story* and *Procedure* sections).

The design and frequency of events presented to participants in Experiments 1 and 2 are shown in Table 6; the bottom row will be discussed in the Introduction to Experiment 1. One chemical, which will be denoted as the "blocked" chemical (A in Table 6), could make it more likely, less likely, or could have no effect on whether the bacterial strain was observed. The blocked chemical was analogous to the camouflage in Baker et al. (1993). In two experimental conditions, the blocked chemical had a moderately positive contingency ($\Delta p = 0.5$; first two treatments in Table 6) with the bacterial strain's presence. That is, the probability of the strain being present was 0.75 in the chemical's presence and 0.25 in its absence. The difference between these two probabilities was 0.5. In the other two conditions, the blocked chemical had no relationship to the bacterial strain's presence ($\Delta p = 0$; last two treatments in Table 6). The probability of the bacterial sample being present was 0.5 regardless of whether the blocked chemical was present or absent. Thus, the difference between these two probabilities was 0.

Table 6

Frequency of Events in Experiments 1 and 2. A refers to the blocked chemical, B refers to the blocking chemical, and X refers to the bacteria. The number before the slash refers to the blocked chemical-bacteria contingency and the number after the slash refers to the blocking chemical-bacteria contingency.

	Experimental Treatment			
Event	**0.5/0**	**0.5/1**	**0/0**	**0/1**
AB => X	9	18	6	12
A => X	9	0	6	0
B => X	3	6	6	12
no A & no B => X	3	0	6	0
AB => no X	3	0	6	0
A => no X	3	6	6	12
B => no X	9	0	6	0
no A & no B => no X	9	18	6	12
Total Number of Trials	48	48	48	48
Δp (A)	0.5	0.5	0	0
Δp (B)	0	1	0	1
Δp (A) | no B	0.5	0	0	0

The blocked chemical was presented with a second chemical that could also appear on some trials. This chemical, which will be denoted as the "blocking" chemical (B in Table 6), was either a perfect predictor of or was uncorrelated with whether the bacterial sample was observed, i.e., $\Delta p = 1$ or 0, respectively. The blocking chemical was analogous to the spotter plane of Baker et al. (1993). Each blocked contingency was paired with each blocking contingency, resulting in four experimental conditions labelled 0.5/0, 0.5/1, 0/0 and 0/1. The number before the slash refers to the blocked chemical-bacteria contingency and the number after the slash refers to the blocking contingency. This design was used in all experiments except Experiment 7. Unless otherwise stated, these four contingencies were administered to all participants in each experiment. Within each contingency, the names of the blocked and blocking chemicals were counterbalanced across participants.

Table 6 describes the frequency of events for each of the four conditions. For example, when a moderately positive blocked chemical was paired with a non-contingent blocking chemical (0.5/0: see the first treatment in Table 6), the bacteria sample was present on 18 out of 24 trials in the blocked chemical's presence and on 6 out of 24 trials in its absence. Since the bacterial strain was observed on half the trials in which the blocking chemical was present, there were 9 trials in which the outcome occurred in the presence of both the blocked and blocking chemicals (i.e., 9 AB => X trials) and 9 trials in the presence of the blocked chemical by itself (i.e, 9 A => X trials). Of the remaining 6 trials in which the blocked chemical was present, the outcome failed to occur on 3 trials in the presence of the blocked and blocking chemicals together (i.e., 3 AB => no X trials) and 3 trials in the presence of the blocked chemical by itself (3 A => no X trials). Similarly, when the blocked chemical was absent, the bacterial strain was observed on 3 trials in which the blocking chemical was present and 3 trials in which neither cue was present (i.e, 3 B => X and 3 no A & no B => no X trials). Finally, on the remaining 18 trials, the bacterial strain was not observed on 9 trials in which the blocking chemical was present and 9 trials in which neither the blocked nor blocking chemicals was present (i.e., 9 B => X and 9 no A & no B => X trials).

However, when the blocking chemical was perfectly contingent (0.5/1: see the second treatment of Table 6), the bacterial strain was observed every time the blocking chemical was present and was never observed in its absence. Of the 18 trials in which the bacterial strain was observed in the blocked chemical's presence, all included the blocking chemical's presence (AB => X) . And of the remaining 6 trials, the bacterial strain was never observed if the blocking chemical was absent (A => no X). Similarly, the blocking chemical was present on all 6 trials in which the bacterial sample was observed in the blocked chemical's absence (B => X), and was absent on all 18 trials in which the sample was not observed (no A & no B => no X). The logic in the explanation outlined above can be extended to the third and fourth conditions of Table 6, i.e., 0/0 and 0/1.

With this paradigm, Baker et al. (1993) have found that people can discriminate moderate and weak blocked contingencies when they are paired with a second weak blocking contingency. That is, judgments of the blocked cue were higher in 0.5/0 than in 0/0. In addition, ratings of the blocked cue were lower when the blocking contingency was perfect than when it was weak. In other words, judgments of the moderate cue were lower in 0.5/1 than in 0.5/0; similarly, judgments of the zero blocked contingency were lower in 0/1 than in 0/0. However, people discriminated the 0.5/1 and 0/1 contingencies. Thus, judgments of the blocked contingency in 0.5/1 were higher than 0/1.

The experiments reported here used the same design as Baker et al. (1993), but differed in one important way. In Baker's experiments, participants decided whether to camouflage the tank. Their decision, in turn, influenced the trial's outcome. Since the task was stochastic, there was some variation in the contingencies. To remove this variability, these experiments were set up so that the frequency of each event was programmed before the experiment had commenced. The 48 trials were divided into three 16-trial blocks, and the program was designed so that the contingency in each game was attained by the end of each block. Estimates of the contingencies were requested twice for each contingency: at the end of the second block (i.e., after 32 trials) and at the end of each game (i.e., after 48 trials). Testing for differences between the two blocks would allow for detecting whether participants' judgments had reached asymptote in the allotted number

of trials. The first estimate was requested at the end of the second block rather than after twenty-four trials so that the attained contingency would be the same as for the final estimates.

Apparatus

A PC computer with a VGA colour monitor was used for presenting the contingencies and collecting data. All experiments were programmed in QuickBASIC.

Data Analysis

Unless otherwise stated, the ratings for the blocked and blocking contingencies were analysed separately with two 3-way repeated-measures analyses of variance (ANOVAs). The three factors were the BLOCKED contingencies ($\Delta 0 = 0.5$ or 0), the BLOCKING contingency ($\Delta p = 1.0$ or 0), and TIME of evaluation (first and second set of ratings). The rejection criterion, α, was 0.05 for the overall ANOVAs. For tests of simple main effects and simple interaction effects, the experiment-wise error rate was corrected using the Bonferroni method.

EXPERIMENT 1

The goal of the first experiment was to replicate Experiment 1 of Baker et al. (1993) using chemicals as the two causes and the survival of the bacteria as the outcome. This experiment used a 2C-1E scenario. That is, there were two causes of one effect, and the causes and were presented before the effect. This order of presentation is consistent with Waldmann and Holyoak's (1992) notion of predictive reasoning.

The R-W model and causal model theory make similar predictions about the experimental outcome. Both models predict that participants will discriminate the 0.5 and 0 contingencies when the blocking contingency is zero. Judgments of the 0.5, but not the 0, contingency should be lowered when the blocking contingency is perfect. Hence, the order of means for the blocked contingency should be $0.5/0 > 0.5/1 = 0/0 = 0/1$. Although

the two sets of models make the same predictions, the outcomes arise from different mechanisms.

According to the R-W model, the cues presented first would be analogous to conditioned stimuli (CS) or inputs of a simple network (Figure 3). The important assumption is that input cues compete to be linked with an outcome. Hence, in the 0.5/0 condition, the 0.5 contingency should gain a moderate amount of associative strength whereas the 0 contingency should accumulate little associative strength. This occurs because the 0.5 contingency is most often reinforced with the outcome, whereas the 0 contingency is both reinforced and not reinforced an equal number of times. The remaining associative strength should accrue to the context. In the 0.5/1 condition, the perfect contingency should gain all of the associative strength since it is always reinforced, whereas the 0.5 contingency and context should accumulate little associative strength. When the blocked contingency is zero, it should gain little associative strength because there are an equal number of trials in which the cue is reinforced and not reinforced. Hence, ratings of the zero contingency should not be lowered when it is paired with a perfect contingency because, at asymptote, it has little associative strength to lose.

Causal model theory's predictions are based on three critical assumptions. The first two are that people reason from causes to effects and that causes compete to be linked with effects (Figure 3). The third is that when there are multiple causes, each cause is evaluated in the presence and/or absence of alternative factors. That is, rather than evaluating unconditional contingencies, people evaluate conditional contingencies. For positive contingencies, candidate causes are evaluated conditional on the absence of other causes (Waldmann & Holyoak, 1992). The formula for calculating the conditional Δp is

$$\Delta p\ (A \mid \text{no } B) = p\ (\text{effect} \mid A \ \&\ \text{no } B) - p\ (\text{effect} \mid \text{no } A \ \&\ \text{no } B)$$

The conditional Δp for each condition is shown in Table 6 (bottom row). For 0.5/0, the conditional Δp for A is 0.5 because the outcome occurs on 9 out of 12 trials in which A is present by itself (i.e., second and sixth rows in Table 6), and on 3 out of 12 trials in which neither A nor B is present (i.e., fourth and eighth rows in Table 6). The difference between these two probabilities is 0.5. Using the above formula,

$$\Delta p\,(A \mid \text{no } B) = ((9 / (9 + 3)) - ((3 / (3 + 9)) = 0.75 - 0.25 = 0.5$$

Similarly, applying the equation to the other conditions yields conditional Δps of 0 in the 0.5/1, 0/0 and 0/1 conditions. Thus, like the R-W model, causal model theory predicts that people will discriminate the 0.5 and 0 contingencies when they are paired with a zero contingency, and that blocking should be observed for the 0.5, but not the 0, blocked contingency.

Method

Participants

Twenty-four McGill University students participated in this experiment. Five participated for course credit and nineteen were paid $5. Eight were male and sixteen were female.

Cover Story

Participants were told that four new strains of bacteria were discovered and that certain pairs of chemicals could affect the bacteria's survival. The chemicals were the causes and the bacterial strain's survival was the effect. To test whether the chemicals affected the strain's survival, the bacteria were placed in culture and one chemical, the other chemical, both chemicals, or neither chemical was added. The culture's survival was later verified.

To help ensure that participants treat each contingency differently, they were told that the scientists were testing four different strains of bacteria and four different pairs of chemicals. For the 0.5/1 contingency, the two chemicals were called ubitone and dichloroparylate and the strain of bacteria was called E. Chronismus. In the 0.5/0 contingency, the two chemicals were chorbine and phylate, and the strain of bacteria was E. D471. In the 0/1 contingency, the two chemicals were blornazene and cyclozentate, and the strain of bacteria was E. Tremus. And in the 0/0 contingency, the two chemicals were pentomorine and isopanone, and the strain of bacteria was E. Conti.

The name of the bacterial strain being tested and the trial number were printed in a box that took up the top half of the screen. To further ensure that the four conditions were treated differently, the colour of the box was different for each game. In the 0.5/1 contingency, the box was cyan; in the 0.5/0 contingency, the box was magenta; in the 0/1 contingency, the box was red; and in the 0/0 contingency the box was yellow.

Procedure

After reading and signing the consent form, the participants were seated in front of the computer. They were told that the instructions would appear on the computer screen and to hit the space-bar to advance from one screen of instructions to the next. As well, they were told that they could ask the experimenter questions at any time during the instructions phase. There were five screens of instructions, which are shown in the Appendix (section I). The first screen explained that there were four strains of bacteria and that for each strain scientists would do a series of experiments to evaluate whether certain pairs of chemicals affected its survival. Each experiment involved adding one chemical, the other chemical, both chemicals, or neither chemical to a bacterial culture and later testing whether it survived. The second and third screens explained that a chemical might make the culture more likely to survive, less likely to survive, or might have no effect on the culture's survival. The fourth screen explained that on each trial participants would first be informed about the presence or absence of each of the two chemicals. After entering whether they thought the bacterial culture would survive, the participants would be informed whether their decision was correct or incorrect. To encourage people to attend to all four cells of a contingency table, the participants were advised to keep track of what happened both in the presence and absence of the chemicals. The fifth screen stated that participants would rate how strongly the chemicals affected the culture's survival on two occasions in each game. The scale ranged from -100 for a perfectly negative contingency to +100 for a perfectly positive contingency, with 0 meaning that a chemical did not affect the bacterial culture's survival.

Before commencing the experiment, the experimenter asked the participants if they had any questions. After answering these questions, the experimenter left the room and allowed the participants to proceed with the task. Once the space-bar was pressed, the first of the four games began. During each trial, participants were informed about the presence or absence of each of the two chemicals. They then entered whether they thought the bacterial strain would survive, after which they were informed whether they were correct or incorrect and were informed of the trial's outcome. Participants then pressed the space-bar to commence the next trial.

Results

Ratings of the Blocked Contingencies

Table 7 contains the first and final estimates of both the blocked and blocking contingencies, and Figure 4 displays the final estimates of the blocked contingencies. The results imply that participants were able to discriminate the blocked contingencies when the blocking relationship was 0, and that judgments of the blocked contingencies were reduced when the blocking relationship was strong. However, it appears that this reduction was larger when the blocked contingency was 0.5 than when it was 0.

Statistical analyses confirm these impressions. The main effects for the blocked and blocking contingencies were significant, and so was the blocked by blocking interaction [F (1, 23) = 5.66, F (1, 23) = 135.05, and F (1, 23) = 5.58, respectively]. A Bonferroni correction was applied for the following two tests of simple main effects; thus, the rejection criterion was 0.025. These tests demonstrated that blocking was observed for both the 0.5 and 0 contingencies, but that the effect was larger for the moderate blocked contingency, F (1, 23) = 53.47, than for the zero blocked contingency, F (1, 23) = 30.08. In summary, participants discriminated the two blocked contingencies but were also influenced by the perfect blocking contingency.

Table 7

Experiment 1: 2C-1E: ?CE

Mean Estimates of the Blocked and Blocking Contingencies after 32 (first rating) and 48 trials (final rating). The number before the slash refers to the blocked chemical's contingency and the number after the slash to the blocking chemical's contingency.

BLOCKED CONTINGENCIES				
	First Ratings		*Final Ratings*	
Condition	**Mean**	**SEM**	**Mean**	**SEM**
0.5/0	36.9	6.8	34.8	8.0
0.5/1	- 37.7	8.7	- 34.6	8.5
0/0	- 2.7	6.9	2.7	4.9
0/1	- 33.3	11.3	- 43.5	7.4
BLOCKING CONTINGENCIES				
0.5/0	- 22.1	9.6	-12.3	9.6
0.5/1	91.7	4.7	94.8	3.7
0/0	-1.9	8.7	2.1	7.1
0/1	91.0	7.3	99.2	0.8

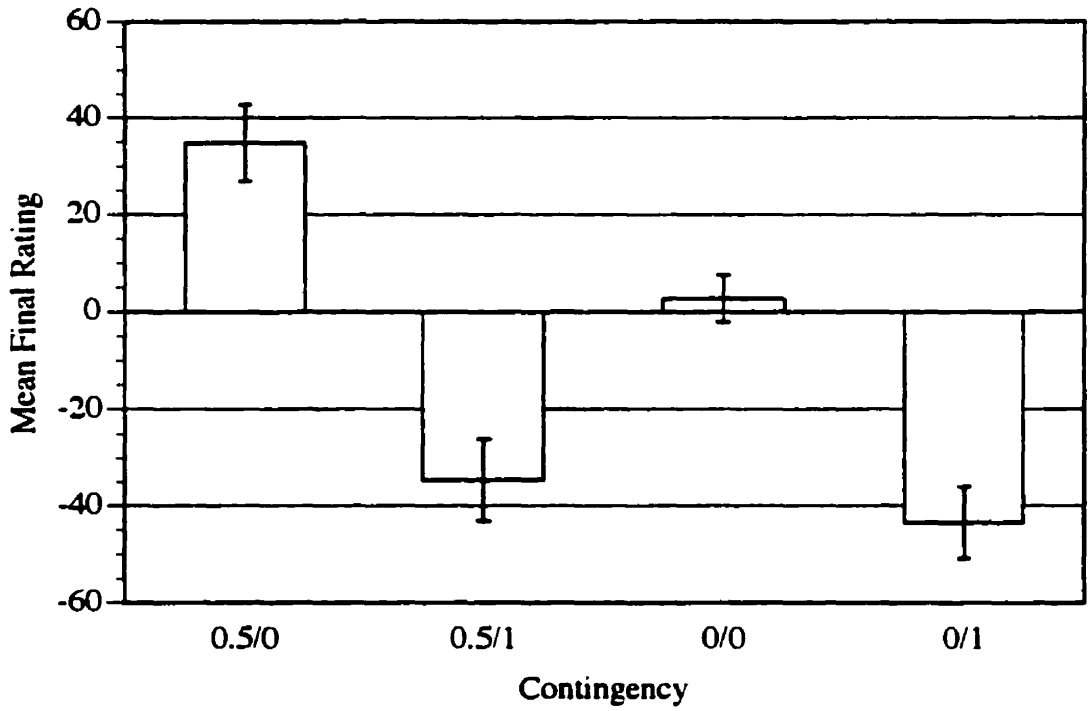


Figure 4. Final judgments and standard errors of Blocked contingency in Experiment 1 (2C - 1E).

Ratings of Blocking Contingencies

Inspection of Table 7 implies that participants discriminated the perfect and zero contingencies. In addition, ratings of the perfect contingency do not appear to be influenced by the blocked contingencies. Judgments of the zero contingency, on the other hand, appear lower when it was paired with a moderately positive than with a zero contingency.

The ANOVA confirms that participants discriminated the two blocking contingencies. The main effect for the blocking contingencies was significant, F (1, 23) = 291.53. However, the main effects for the blocked contingencies and time failed to attain significance, F (1, 23) = 3.41, F (1, 23) = 3.28, respectively. None of the interactions attained significance, largest F (1, 23) = 1.01. Thus, judgments of the blocking contingency were influenced only by the strength of the blocking contingencies.

Discussion

The results from this experiment are consistent with previous reports in which strong contingencies lower judgments of weaker ones when causes are presented before effects (e.g., Baker et al., 1993; Shanks & Lopez, 1996; Van Hamme et al., 1993; Van Hamme & Wasserman, 1993; Waldmann & Holyoak, 1992; Wasserman, 1990a). The findings from this experiment replicate those of Baker et al. (1993) whereby the moderate and weak blocked contingencies were discriminated when they were paired with a weak blocking contingency. Furthermore, judgments of both blocked contingencies were lowered when they were paired with a strong blocking contingency. These results are consistent with aspects of both the R-W model and causal model theory. Both models predict blocking of the 0.5 but not the 0 contingency. However, blocking of 0 contingencies was found in this experiment and has been observed in prior experiments (e.g., Baker et al., 1993; Baker et al., 1996). This finding is thus not unique to the preparation used in this experiment.

In addition, analysis of the blocking contingencies revealed that judgments of a strong contingency were not influenced by weaker ones. This result was consistent with

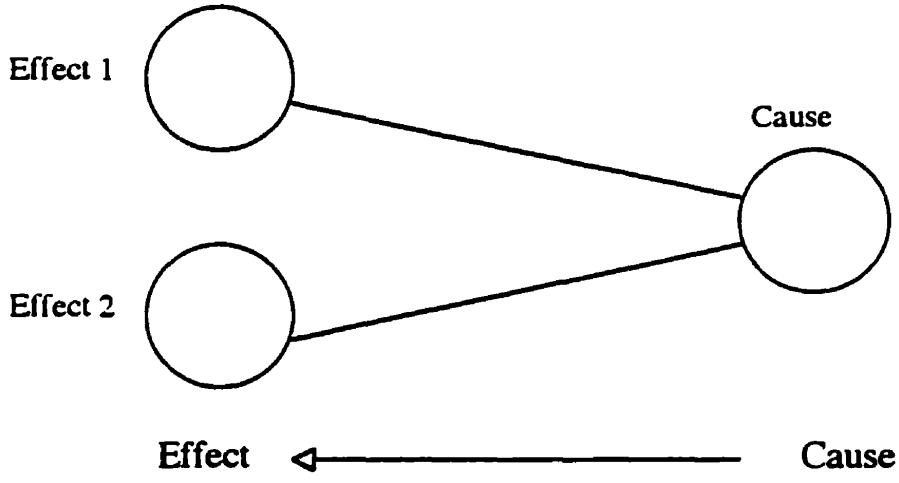
the predictions of both the R-W model and causal model theory. According to the R-W model, judgments of a strong contingency should not be lowered by a weak one because the strong one should acquire associative strength whereas the weaker ones should not. Similarly, according to causal model theory, the conditional Δps for the blocking contingencies are 0, 1, 0 and 1 for the 0.5/0, 0.5/1, 0/0 and 0/1 conditions, respectively. People's judgments should therefore be influenced by the blocking but not the blocked contingencies. Thus, both models predict that the perfect contingencies should be rated higher than the zero contingencies and that the zero contingencies should not be influenced by whether they are paired with a weak or moderate contingency. In short, the results from this experiment are consistent with previous findings and are consistent with both the R-W model and causal model theory.

EXPERIMENT 2

In Experiment 1, cue competition was observed in 2C-1E. Experiment 2 was an attempt to detect whether cue competition would be detected in 2E-1C, that is when effects were presented before causes. The cues and method of presentation were the same as those described in Experiment 1. However, the nature of how the cues interacted and the order in which they were presented were different. Rather than evaluating how the chemicals influenced the bacteria, participants evaluated the extent to which a strain of bacteria affected the production of two chemicals. Hence, the bacterial strain was the cause and the production of the two chemicals were the effects. Furthermore, the presence or absence of the two effects were presented before the presence or absence of the cause. This order of presentation, inferring cause from effects, conforms to Waldmann and Holyoak's (1992) definition of diagnostic reasoning.

The R-W model and causal model theory make opposite predictions about the experimental outcome (Figure 5). Whereas R-W predicts that cue competition should be observed, causal model theory predicts its absence. According to the associative approach, the effects presented first would be represented as input cues and the cause would be encoded as the outcome in a simple network. Since input cues compete for

Causal Model Theory's Predictions for 2E - 1C



R-W model's Predictions for 2E -1C

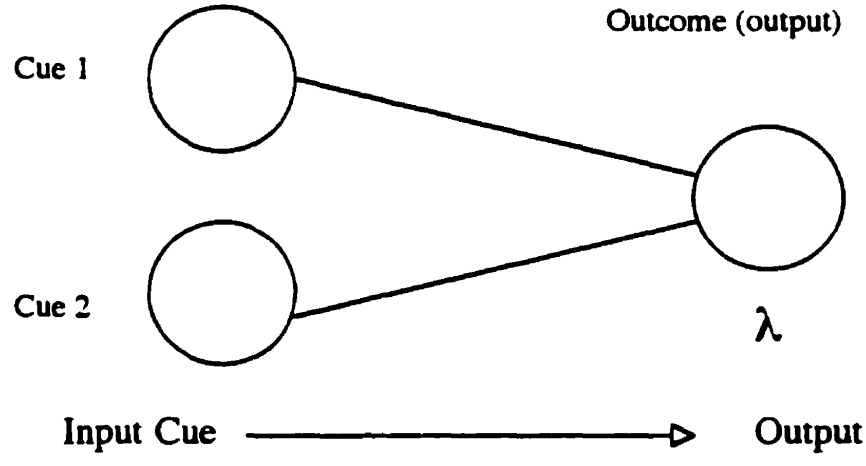


Figure 5. Models displaying the assumptions in causal model theory and the Rescorla - Wagner model for Experiment 2 (2E - 1C).

association with an outcome, cue competition is expected with this approach. Since the frequency of events, apart from whether the cues are defined as causes or effects, is the same as Experiment 1, the associative model predicts the same order of means. That is, for the blocked contingencies, the order of means should be 0.5/0 > 0.5/1 = 0/0 = 0/1. Hence, participants should discriminate the 0.5 and 0 contingencies when they are paired with a 0 contingency, and blocking would be expected for the 0.5 but not the 0 contingency.

However, causal model theory rests on the assumption that people reason from cause to effect even if the information is presented in the reverse order (Figure 5). Furthermore, only causes compete to be linked with effects. Since there is only one cause, each contingency should be represented as being independent of the other. Cue competition would therefore not be expected. The predicted order of means for the blocked contingencies should thus be (0.5/0 = 0.5/1) > (0/0 = 0/1). In other words, people should discriminate the 0.5 and 0 contingencies but blocking should not occur. Causal model theory, however, states that one important condition must be satisfied. Participants should not have prior biases or causal models that would influence their perception of the contingencies. As will be explained in the Procedure section, steps were taken to minimize the influence of prior conceptions and alternative causal models on how participants perceived the presented contingencies.

Method

Participants

Twenty-four McGill University students were paid $5 for their participation in this study. Fourteen were male and ten were female.

Procedure

After reading and signing the consent form, participants were seated in front of the computer. They were then told that the instructions would appear on the computer screen and to press the space-bar to advance from one screen of instructions to the next.

They were also told that they could ask the experimenter questions at any time during the instructions phase. The five screens of instructions are displayed in the Appendix (section II). The first screen explained that there were four strains of bacteria being investigated, and that for each strain scientists were carrying out experiments to evaluate whether these bacteria affected the production of certain pairs of chemicals. For each experiment, the scientists obtained blood samples from laboratory rats and verified the presence or absence of each of the two chemicals. The scientists later tested for the presence or absence of the bacteria. In an attempt to minimize the influence of alternative causal models, participants were informed that rats, instead of people, were used so that variables such as diet, drug history, lifestyle factors, and socio-economic status could be controlled for. Furthermore, participants were informed that the study was done under aseptic and highly controlled conditions. Namely, they were told that the animals were kept in filtered cages to prevent contaminants in the air from entering the cage; the room temperature was kept constant; and all animals were fed a similar diet. The second and third screens were designed to explain that the contingency between a chemical and bacteria could be positive, negative or zero. That is, the bacteria might make it more likely, less likely, or might not affect the likelihood that a chemical would be produced. The fourth screen explained what would happen on each trial. Participants would first be informed about the presence or absence of the two chemicals. They would then enter whether they thought that the bacteria were present or absent, after which they would be informed whether they were correct or incorrect, and would read what outcome had occurred. Participants were also encouraged to keep track of what happened in both the bacteria's presence and absence. The fifth screen stated that participants would be asked to rate how strongly each chemical's production was affected by the bacteria on a scale ranging from -100 to +100. The format of this question conforms to Waldmann and Holyoak's (1992) definition of a diagnostic inference.

After answering any questions that the participants might have had about the instructions, the experimenter left the room and allowed the participants to proceed with the task. During each trial, participants were informed about the presence or absence of

the two chemicals and were then asked whether they thought the bacteria was present or absent. After entering their choice, they were informed whether they were correct or incorrect and read what outcome had taken place. They then pressed the space-bar to begin the next trial. The participants' ratings were obtained after the second block and at the end of each game.

Results

Ratings of the Blocked Contingencies

Experiments investigating blocking of moderate and weak contingencies assume that participants can discriminate strong and weak covariation relationships. One participant gave a rating of zero to the perfect contingency in both the 0.5/1 and 0/1 conditions and was therefore excluded from the analysis as an outlier.

Table 8 contains the means and standard errors for the first and final estimates of the blocked contingencies, and Figure 6 displays the final ratings for the blocked contingencies. These results are similar to those in Experiment 1. Participants could differentiate a 0.5 contingency from a 0 contingency when these contingencies were paired with a zero contingency, although the ratings of the 0 blocked contingency were somewhat elevated above 0. Furthermore, these ratings were lowered when the blocking contingency was perfect.

Statistical analyses confirm the presence of these effects. The main effects for the blocked and blocking contingencies were significant, F (1, 22) = 6.46, F (1, 22) = 30.20, respectively, but the blocked by blocking interaction was not significant, F (1,22) = 1.94. Contrary to the results of Experiment 1, the main effect for time was significant, F (1, 22) = 5.45. Inspection of the table reveals that the final ratings of the blocked contingencies were higher than the initial ones. However, the time at which the ratings were collected did not interact with any other variable, highest F (1, 22) = 2.22. In short, participants discriminated the two blocked contingencies but gave significantly lower ratings when they were paired with a strong blocking contingency.

Table 8

Experiment 2: 2E-1C: ?EC

Mean Estimates of the Blocked and Blocking Contingencies after 32 (first rating) and 48 trials (final rating). The number before the slash refers to the blocked chemical's contingency and the number after the slash to the blocking chemical's contingency.

BLOCKED CONTINGENCIES				
	First Ratings		*Final Ratings*	
Condition	**Mean**	**SEM**	**Mean**	**SEM**
0.5/0	35.5	8.3	36.4	8.8
0.5/1	- 22.8	11.4	- 24.1	10.6
0/0	- 1.5	8.3	16.3	9.2
0/1	-39.6	10.2	-30.0	10.4
BLOCKING CONTINGENCIES				
0.5/0	- 5.4	8.0	- 4.7	10.2
0.5/1	90.4	5.1	93.4	3.6
0/0	13.6	7.9	15.6	7.6
0/1	90.4	4.9	87.8	5.4

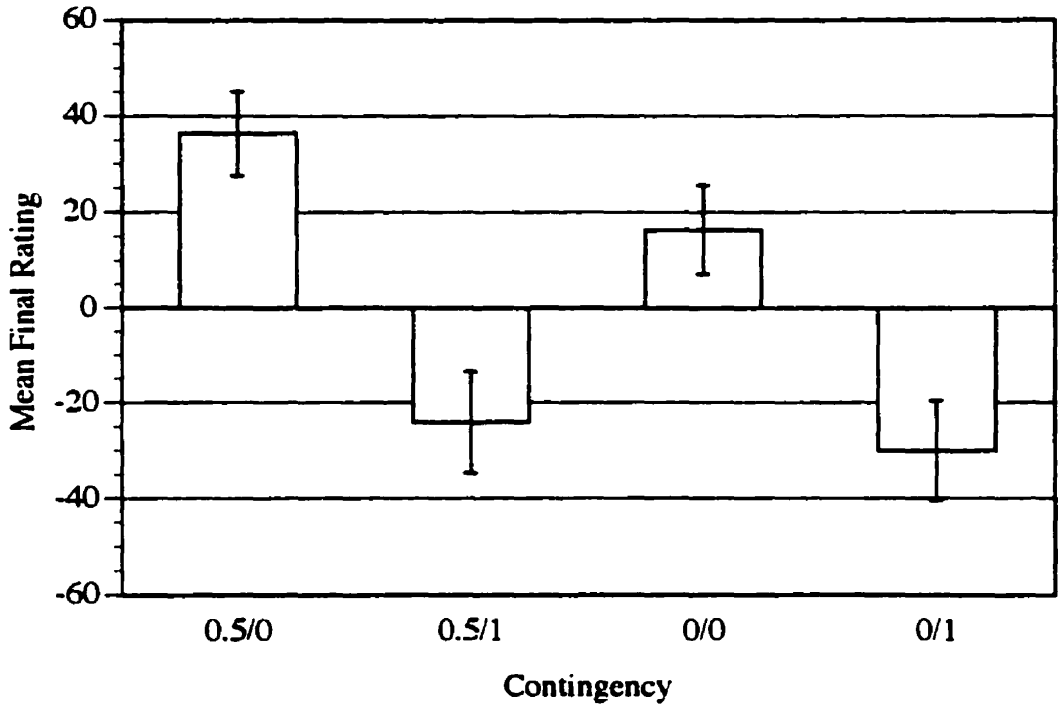


Figure 6. Final judgments and standard errors of Blocked contingency in Experiment 2 (2E - 1C).

Ratings of the Blocking Contingencies

The results in Table 8 imply that participants discriminated a perfect contingency from a weak one. However, the ratings for the zero contingency appear to be lower when the blocked contingency was moderately positive than when it was zero.

Statistical analyses confirm these impressions. Although the main effect for the blocked contingency was not significant, F (1,22) = 2.51, the main effect for the blocking contingencies and the blocked by blocking interaction were significant, F (1, 22) = 124.77, F (1, 22) = 7.84, respectively. A Bonferroni corrected rejection criterion of 0.017 was used for the follow-up analyses. These tests revealed that participants reliably discriminated the perfect and zero blocking contingencies when they were paired with a moderate blocked contingency and with a zero blocked contingency [F (1, 23) = 124.77, F (1, 23) = 128.21, respectively]; however, ratings of the zero blocking contingency were significantly higher when it was paired with a zero than with a 0.5 blocked contingency, F (1, 23) = 6.92. The main effect for time was not significant, F (1, 22) < 1, and this factor did not interact with any other variable, all Fs (1, 22) < 1.

Discussion

Participants discriminated the blocked contingencies when they were paired with a zero contingency. Judgments of these contingencies were lowered when they were paired with a perfect contingency. The size of this effect was similar for both of the blocked contingencies. Although these results are somewhat different from Experiment 1, in which the cue competition effect was smaller for the zero contingency, they are similar to those of Baker et al.'s (1993) Experiment 1. Furthermore, the results are generally consistent with previous reports of cue competition between effects when they were presented before causes (Chapman, 1991; Matute et al., 1996; Price & Yates, 1993; Waldmann & Holyoak, 1992, Experiment 2; Shanks, 1991; Shanks & Lopez, 1996; however, see Van Hamme et al., 1993; Waldmann & Holyoak, 1992, Exp 3). They are also consistent with the finding of Vallée-Tourangeau et al. (1994), in which cue competition was observed when two geometric shapes predicted the occurrence of a third

shape. In this latter experiment, it is difficult to argue that the cue competition was due to participants inferring a causal relationship among the cues. These experiments attest to the generality of cue competition between cues when they are presented before a single outcome, regardless of whether they are defined as causes preceding an effect or effects preceding a cause.

The results appear to provide clear support for an associative approach. In Experiments 1 and 2, when the cover stories involved two predictors of a single outcome, judgments of moderate and weak contingencies were lowered when they were paired with a strong one, regardless of whether the predictors were defined as causes or effects. Thus, at a superficial level, the results would provide apparently indisputable support for an associative approach and not causal model theory. On the other hand, some ambiguity remains. It could be argued that cue competition occurred because participants had either misunderstood the causal model or had used alternative competing causal models (Waldmann & Holyoak, 1992, 1997).

According to the first argument, participants in Experiment 2 would have interpreted the chemicals as causes and the bacteria as the effect, rather than as intended. This argument does not seem plausible given that participants were explicitly informed that they were evaluating the extent to which the bacteria influenced the production of the two chemicals. It appears that the bacteria would clearly be the cause and the two chemicals would be the effects. However, these arguments are speculative and this experiment provides no evidence either in favour or against these approaches because the participants' understanding of the instructions was never directly measured. This issue will be addressed later.

The second argument, that blocking occurred due to alternative causal theories, seems less plausible. As mentioned, cue competition has been found when cues were effects that were presented before causes. Furthermore, these results are consistent with Shanks and Lopez (1996), who observed cue competition in their multiple effect-single cause conditions despite various attempts to manipulate focal sets and causal models. Finally, this effect has been found even in scenarios in which causal models could not

plausibly be invoked (Vallée-Tourangeau et al., 1994). This converging evidence implies that the cue interaction effect observed is not likely due to competing causal models. However, direct measurement of these models is needed and would provide stronger evidence in favour of this argument.

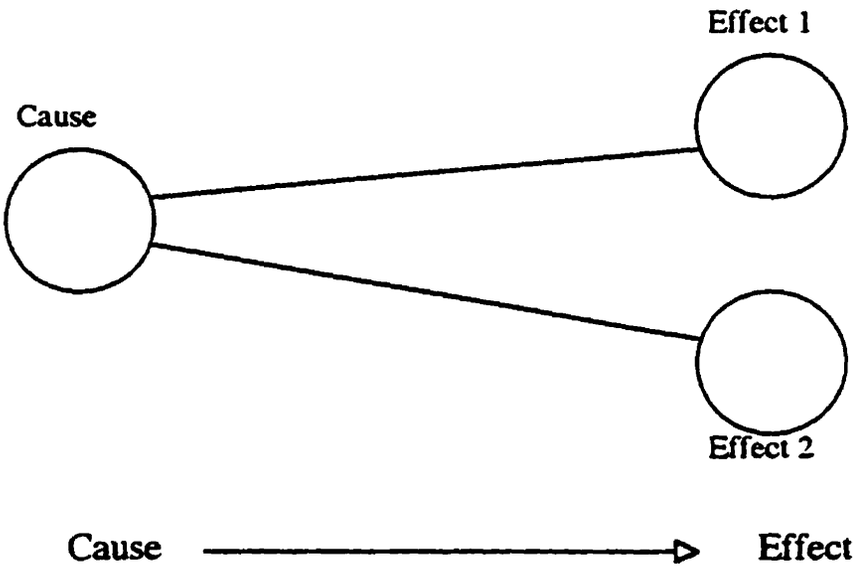
To summarize, these results are consistent with an associative account but do not definitively exclude causal model theory. Although blocking was observed, it is difficult to speculate on the reasons for its occurrence. On one hand, it could have been the product of an associative mechanism. On the other hand, it is possible that participants did not comprehend the causal model in the cover story or that they were influenced by alternative causal models based on their knowledge of chemicals and bacteria. These issues were investigated in Experiments 5, 6, and 7.

EXPERIMENT 3

Experiments 1 and 2 were replications of previous work investigating cue selection effects when two cues preceded a single outcome. In these experiments, blocking was observed when two cues preceded a single cue, regardless of whether the antecedent cues were defined as causes or effects. Experiment 3 was an attempt to test whether blocking would be observed when a single cue predicted two outcomes. Similar to Experiment 2, a single bacterial strain produced two chemicals. However, information about the presence or absence of the bacteria was presented before information about the production of the two chemicals. Thus, one cause preceded two effects (1C-2E).

Both the R-W model and causal model theory predict the absence of cue competition effects with this preparation. According to the R-W model (Figure 7), the cause should be encoded as an input of a simple associative network and the two effects should be the outcomes. Both outcomes would be able to support independent amounts of associative strength (λ_1 and λ_2). Since outcomes do not compete for association with predictor cues, a blocking effect would not be expected. Causal model theory makes the same prediction. The underlying assumptions are that participants reason from cause to effect and that only causes compete to be linked with effects. Since the single cause has

Causal Model's Theory's Predictions for 1C - 2E



R-W Model's Predictions for 1C - 2E

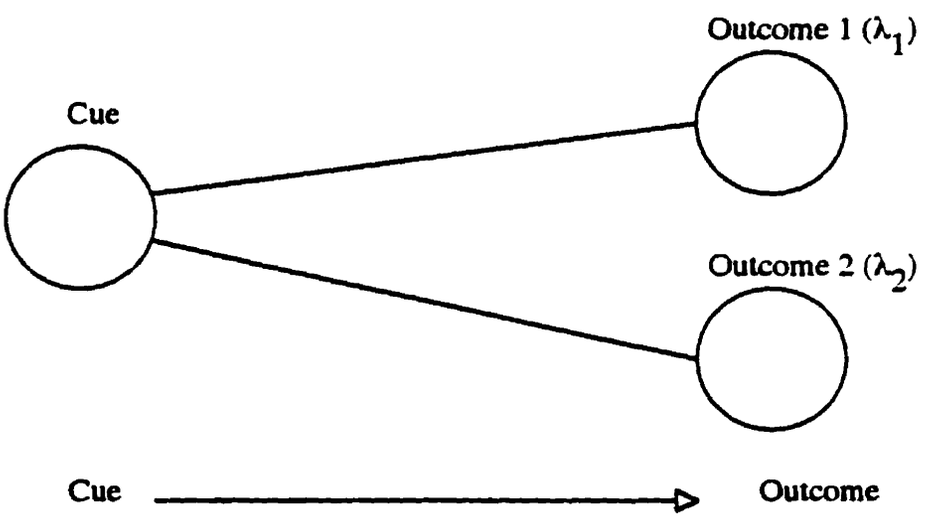


Figure 7. Assumptions of causal model theory and R-W model for Experiment 3 (1C - 2E)

nothing to compete with, blocking would not be expected. Thus, both theories predict that participants should discriminate the 0.5 and 0 contingencies, but that blocking should not be observed.

Previous experimental evidence is consistent with this prediction. Baker and Mazmanian (1989) used a preparation in which participants could choose to press a space-bar to produce two outcomes, that is, changing the colour of a ball and a box. The contingency between pressing the space-bar and changing the ball's colour was 0.5, 0, or -0.5, whereas the contingency between pressing the space-bar and the box's colour change was 0, 1 or -1. Rather than being blocked, estimates of the non-zero ball contingencies were in fact *enhanced* when the box's contingency was strong, regardless of whether it was perfectly positive or negative. Similar results have also been observed in animal learning. For example, Rescorla (1991) trained rats to associate an operant response with multiple outcomes. Learning of one response-outcome did not influence associations of that response with other outcomes. Furthermore, Colwill and Rescorla (1985) found that extensive training of one response with a particular outcome did not interfere with associating that response with a novel effect. Based on this evidence, it was hypothesized that blocking would not be observed.

Baker and Mazmanian's (1989) procedure was adapted for this experiment. Rather than using a free-operant technique in which the outcome's occurrence was dependant on the participant's action, this experiment used a discrete-trial technique. Similar to Experiments 1 and 2, the conditions were programmed so that each contingency was attained by the end of a sixteen-trial block.

The methodological issue of how to structure the task during the acquisition of the contingencies presented some complexity. In Experiments 1 and 2, the two antecedent cues were presented; participants then made a prediction about the trial's outcome and were informed about what event had occurred. This experiment, however, differed from the first two experiments in that a single cue preceded two outcomes. To address this difference, two methods were devised by which participants could predict the two events that took place on each trial. The first method was to have people make a guess about the

presence or absence of each outcome separately; in other words, to predict the presence or absence of each outcome independently. This group was labelled the *two-decision group* because participants made two predictions, or decisions, on each trial. A potential problem with this method of prediction is that it implicitly forces people to monitor the two contingencies independently. In an attempt to alleviate this problem and to control for the number of responses made within each trial, a second group of participants was presented with the four possible conjunctions of the presence and absence of the two outcomes and had to guess which one they thought would occur. This group was called the *single-decision group* because participants in this condition made one prediction on each trial.

Method

Participants

Forty-eight McGill University students participated in this study. Forty-six participated for course credit and two were paid $5. Twelve participants were male and thirty-six were female. Participants were randomly assigned to the two-decision or single-decision groups.

Cover Story and Contingencies

Participants were informed that four new strains of bacteria had been discovered and that these bacteria could influence the production of certain pairs of chemicals. To test whether the bacteria influenced the production of these chemicals, the bacteria were placed in an environment that simulated the human digestive system and the production of each of the two chemicals was later verified. The contingencies and the names of the cues were the same as the previous experiments. However, rather than having information about the presence or absence of two chemicals precede information about the bacteria, participants were first informed about a single event and then about the occurrence of two outcomes (see Table 9).

Table 9

Frequency of Events in Experiments 3, 4, 5 and 6. A refers to the blocked chemical, B refers to the blocking chemical, and X refers to the outcome. The number before the slash refers to the blocked chemical's contingency and the number after the slash refers to the blocking chemical's contingency.

	Experimental Treatment			
Event	**0.5/0**	**0.5/1**	**0/0**	**0/1**
X => A & B	9	18	6	12
X => A	9	0	6	0
X => B	3	6	6	12
X => no A & no B	3	0	6	0
no X => A & B	3	0	6	0
no X => A	3	6	6	12
no X => B	9	0	6	0
no X => no A & no B	9	18	6	12
Total Number of Trials	48	48	48	48
Δp (A)	0.5	0.5	0	0
Δp (B)	0	1	0	1
Δp (A) | no B	0.5	0	0	0

Procedure

After reading and signing the consent form, participants were seated in front of the computer. They were then told that the instructions would appear on the computer screen and to hit the space-bar to advance from one screen of instructions to the next. They were also told that they could ask the experimenter questions at any time during the instructions phase. There were five screens of instructions, and they are displayed in the Appendix (section III). The first screen explained that there were four strains of bacteria being investigated, and that for each strain scientists were carrying out experiments to evaluate whether these bacteria affected the production of certain pairs of chemicals. For each experiment, the bacteria were either placed or not placed in the simulated human digestive system. The scientists then verified whether one chemical, the other chemical, both chemicals, or neither chemical was produced. The second and third screens explained that the bacteria might make it more likely, less likely, or might not affect the likelihood that a chemical would be produced. The fourth screen explained what would happen on each trial, namely that participants would first be informed about the presence or absence of the bacteria, after which they would enter whether they thought each of the two chemicals would be produced. People in the two-decision group were informed that they would guess about the presence or absence of each chemical separately, whereas those in the one-decision group read that they would be presented with the four possible outcomes that could take place and had to guess which one occurred. Participants in both groups were encouraged to keep track of what happened in both the presence and absence of the bacteria. The fifth screen stated that participants would be asked to rate how strongly the bacteria affected the production of each chemical on a scale ranging from -100 to +100.

After answering any questions that the participants might have had about the instructions, the experimenter left the room and allowed the participants to proceed with the task. During each trial, participants were informed about the presence or absence of the bacteria. The two-decision group was then asked - in two separate questions - whether each of the two chemicals was produced. After entering whether an individual chemical

was produced, participants were informed whether their guess was correct or incorrect, and read what outcome occurred. On the other hand, participants in the one-decision group were first informed about the presence or absence of the bacteria, and were presented with the four possible conjunctions of the two outcomes. Choice A was that one chemical was produced; choice B was that the second chemical was produced; choice C was that both chemicals were produced; choice D was that neither chemical was produced. After selecting the conjunction that they expected would occur, participants were given feedback about the correctness of their answer and were informed what outcome had occurred. At the end of the trial, participants in both groups pressed the space-bar to begin the next trial. Participants estimated the effectiveness of the bacteria on each chemical after 32 and after 48 trials.

Results

Ratings of Blocked Contingency

Data from one participant were excluded because that person reported being confused during the task and because inspection of that person's results revealed that zeros were entered for all contingencies.

Table 10 displays both group's first and final ratings for the blocked contingencies and Figure 8 shows their final ratings. The present experiment tested the hypothesis that blocking should not occur when one cause precedes two effects. The evidence in consistent with this notion. Although the ratings of the two-decision group are higher than those of the single-decision group, both groups discriminated the moderately positive and zero contingencies. However, unlike Experiments 1 and 2 in which a strong contingency reduced judgments of a weaker one, the perfect contingency appears to differentially influence ratings of the blocked contingencies in the two groups. In the two-decision group, ratings of both blocked contingencies were higher when they were paired with a perfect contingency, although this enhancement appears smaller for the final set of ratings. The final estimates of the zero contingency appear similar regardless of whether it was paired with a zero or perfect contingency. On the other hand, in the single-choice

Table 10

Experiment 3: 1C-2E: ?CE
Mean Estimates of the Blocked Contingencies after 32 (first rating) and 48 trials (final rating) for both the two-decision and single-decision groups. The number before the slash refers to the blocked contingency and the number after the slash refers to the blocking contingency.

TWO - DECISION GROUP				
	First Ratings		*Final Ratings*	
Condition	**Mean**	**SEM**	**Mean**	**SEM**
0.5/0	37.4	9.4	48.0	6.4
0.5/1	55.2	7.0	61.6	6.5
0/0	15.0	6.2	11.7	6.7
0/1	25.7	8.9	7.5	6.2
ONE - DECISION GROUP				
0.5/0	25.1	8.5	19.9	8.6
0.5/1	27.2	11.4	32.3	7.6
0/0	15.7	6.3	5.7	7.1
0/1	7.1	11.8	-14.0	8.3

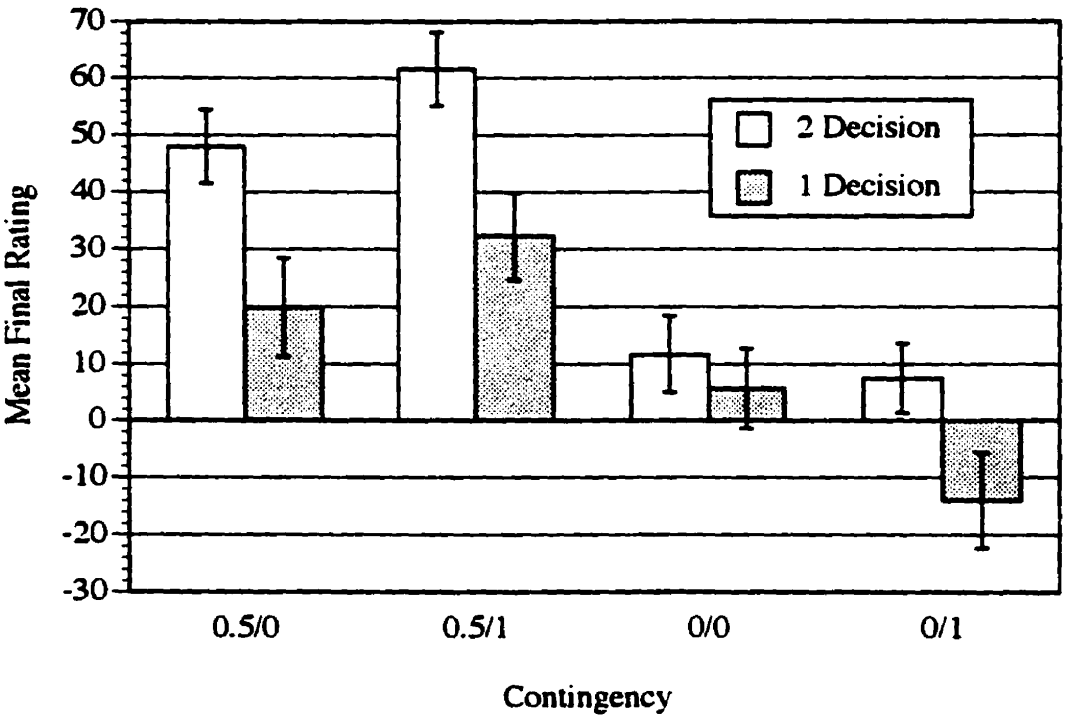


Figure 8. Final judgments and standard errors of Blocked contingency in Experiment 3 (1C - 2E).

group, the judgments are different for both sets of ratings. For the first estimates, the 0.5 contingency was rated similarly regardless of whether it was paired with a zero or perfect contingency. But for the final estimates, judgments appear enhanced when the 0.5 contingency was paired with a perfect one. Ratings of the zero contingency, on the other hand, are somewhat lower when it was paired with a perfect contingency, and this effect appears more pronounced after 48 trials.

Upon analysis of the blocked contingencies, the main effect for Group was significant, F (1, 45) = 9.36, supporting the view that the two-decision group's ratings were higher than those of the one-decision group. However, the effect of group did not interact with any other variable, highest F (1, 45) = 3.09. Thus, both groups gave similar types of ratings. The main effect for the blocked contingencies was significant, F (1, 45) = 34.41, supporting the contention that participants discriminated the 0.5 and 0 contingencies. The main effects for the blocking contingencies and of time, however, were not statistically significant, largest F (1, 45) = 1.61. The blocked by blocking interaction and the blocked by time interactions were both significant, smallest F (1, 45) = 4.11. But the blocking by time and blocked by blocking by time interactions were not significant, largest F (1, 45) = 1.78. A Bonferroni correction was applied for the following four follow-up analyses. Thus, the adopted rejection criterion was 0.013.

Inspection of Table 10 suggests the blocked by time interaction arose because participants were able to discriminate the 0.5 and 0 contingencies but that the difference in ratings was more pronounced after 48 than after 32 trials. Tests of simple main effect of the blocked contingency after 32 trials and 48 trials support this interpretation: the simple main effect was significant for both sets of ratings but was larger for the final ratings; F (1, 45) = 7.78 for the first set of ratings, F (1, 45) = 57.83 for the final judgments.

The blocked by blocking interaction, taken together with inspection of Figure 8, implies that ratings of the 0.5, but not the 0, contingency were elevated when it was paired with the perfect contingency. However, follow-up analyses do not support this view. A test of the blocking contingencies on the 0.5 contingency was not significant, F

(1, 45) = 3.51. Neither was a test of the simple effect of the blocking contingencies on the zero contingency, F (1, 45) = 1.21.

To summarize, participants discriminated a moderately positive contingency from a zero contingency. However, a strong blocking contingency did not influence judgments of the blocked contingency, regardless of whether participants had to make two separate decisions or only one decision about the outcome of each trial.

Ratings of Blocking Contingencies

Unlike Experiments 1 and 2 in which judgments of the perfect contingency were quite high (i.e., about 90 or greater), judgments of the perfect contingency were only moderately high. As shown in Table 11, the mean final rating for the perfect contingencies in the two-decision group was approximately 80; in the one-decision group, the means for the perfect contingency were even lower. Both groups, however, discriminated the perfect and zero contingencies, although judgments appear lower when the blocking contingencies were paired with the 0 than the 0.5 contingency.

Statistical analyses corroborate these impressions. The main effect of Group was not significant, F (1, 45) = 2.78, and this factor did not interact with any other main or interaction effect, highest F (1, 45) = 3.57. The main effects for both the blocked and blocking contingencies were significant, [F (1, 45) = 5.80, F (1, 45) = 106.62, respectively], but the main effect for time was not, F (1, 45) = 3.01. The blocking by time interaction was significant, F (1, 45) = 8.34. Follow-up analyses revealed that participants' discrimination of the 0 and 1 contingency was better on the final ratings than the first, F (1, 45) = 68.64, F (1, 45) = 111.48, for the first and final ratings, respectively. A Bonferroni corrected rejection criterion of 0.025 was used for these tests of simple main effects. No other interaction effect attained significance, highest F (1, 45) = 2.54.

Discussion

In this experiment, participants were presented with a single cue that preceded two outcomes. The antecedent cue was defined as a cause and the two consequent cues were

Table 11

Experiment 3: 1C-2E: ?CE
Mean Estimates of the Blocking Contingencies after 32 (first rating) and 48 trials (final rating) for both the two-decision and single-decision groups. The number before the slash refers to the blocked contingency and the number after the slash refers to the blocking contingency.

TWO - DECISION GROUP				
	First Ratings		*Final Ratings*	
Condition	**Mean**	**SEM**	**Mean**	**SEM**
0.5/0	11.7	7.9	12.9	8.5
0.5/1	88.0	6.0	82.3	7.7
0/0	6.3	8.4	0.7	6.6
0/1	63.2	8.6	78.2	7.8
ONE - DECISION GROUP				
0.5/0	12.4	9.8	6.0	7.6
0.5/1	58.0	9.1	75.6	5.6
0/0	10.8	5.6	11.6	4.9
0/1	39.0	10.7	61.6	6.7

defined as effects (1C-2E). The main finding from this experiment is that participants discriminated the 0.5 and 0 contingencies, but judgments of the blocked contingencies were not lowered by the blocking contingencies. If anything, judgments of the 0.5 contingency were somewhat enhanced when it was paired with a perfect contingency. However, this effect did not attain significance.

These results replicate previous results from human contingency learning (Baker & Mazmanian, 1989) and animal instrumental conditioning (Rescorla, 1991), and are consistent with both the R-W model and causal model theory. According to the R-W model (Baker et al., 1996), cue competition should not occur in this scenario since there are two outcomes - each of which can independently support different amounts of associative strength. Similarly, according to Waldmann and Holyoak (1992), cue competition should not be observed because effects do not compete to be linked with causes.

A second unexpected finding from this experiment was that contingency judgments were influenced by presentation format. That is, the type of intervening question asked between the presentation of the cause and the two effects appears to be an important variable in contingency judgements. Although blocking was not observed in either group, ratings for the blocked contingencies were higher when participants made two, rather than one, outcome predictions per trial. Furthermore, judgments of the blocking contingency were also somewhat higher when two predictions were required on each trial. This trend did not attain significance, but a comparison of the judgments of the perfect contingency from participants in this experiment with those in Experiments 1 and 2 reveals that these ratings were lower in this experiment. This decrease is more pronounced in the one-decision group, in which the final judgments of the blocking contingency in 0/1 was approximately 60. Judgments of the perfect contingency were higher when two decisions were made on each trial, the means being close to 80. A possible explanation for this decrease lies in the difficulty of the task. Unlike the participants in Experiments 1 and 2, those in the present experiment spontaneously commented that the task was challenging and that they had difficulty monitoring the

events presented. This issue of the type of intervening questions asked between the presentation of causes and effects was further investigated in Experiments 6 and 7.

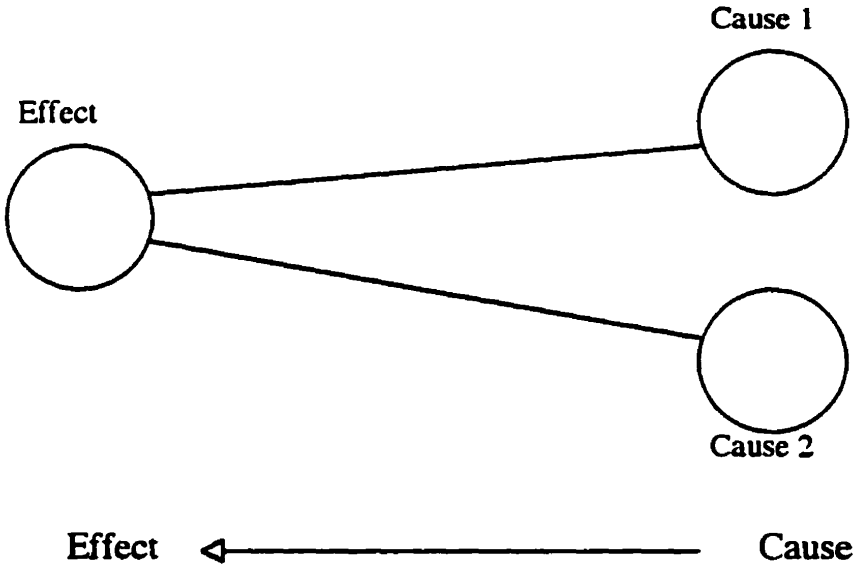
EXPERIMENT 4

Cue competition was not detected in Experiment 3 when one cause preceded two effects. The present experiment was designed to test whether the absence of cue competition would be observed when the cue presented first was defined as an effect and the outcomes as causes (i.e., 1E-2C). Unlike Experiment 3, in which the associative and statistical models predicted the same experimental outcome, the two approaches make opposite predictions for this preparation (Figure 9). According to the R-W model, the scenario entails one cue and two outcomes, with each cue being capable of supporting independent amounts of associative strength. Regardless of whether stimuli are defined as causes or effects, cues compete to be linked with an outcome, but the reverse does not occur. Thus, cue competition should not be observed for either contingency.

On the other hand, the statistical approach predicts that blocking should be observed in the 0.5 contingency. According to causal model theory, even though the contingencies are presented in the order effect-cause, people inherently reason from cause to effect. Furthermore, causes compete to be linked with effects but effects do not compete for association with causes. When causes are linked to effects, the conditional Δp for the 0.5 contingency in the absence of the blocking contingency is 0.5 when that contingency is 0, but is 0 when the blocking contingency is perfect. The conditional Δp for the 0 contingency is 0 both when the blocking contingency is 0 and when it is perfect. It thus follows that the model predicts cue competition for the 0.5 but not the 0 contingency.

Experiment 3 of Price and Yates (1995) evaluated these hypotheses and found support for the R-W model. In this experiment, participants were presented with a moderate contingency ($\Delta p = 0.44$). In the weak blocking condition, this contingency was paired with a zero contingency ($\Delta p = 0$) and in the strong blocking condition it was paired with a strong contingency ($\Delta p = 0.77$). In their 2C-1E condition, the causes were

Causal Model's Theory's Predictions for 1E - 2C



R-W Model's Predictions for 1E - 2C

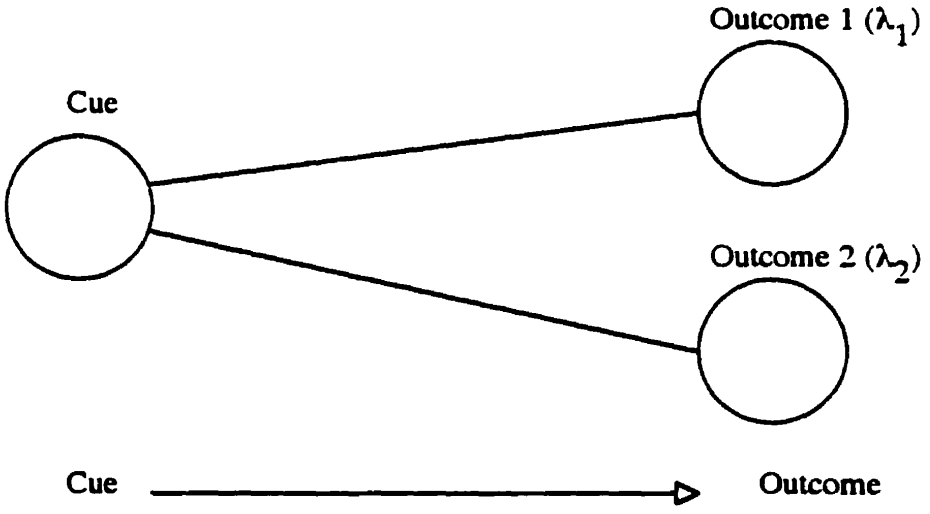


Figure 9. Assumptions of causal model theory and R-W model for Experiment 4 (1E-2C)

presented first and participants predicted the presence or absence of the effect; in 1E-2C, the effect was presented first and participants predicted whether each of the two causes were present. At the end of 36 trials, participants estimated the conditional probability of each effect in the presence and absence of each cause. Judgments of the moderate contingency were lowered when the blocking contingency was strong in 2C-1E, but not in 1E-2C. Although Price and Yates's results are consistent with the R-W model, further experimentation is needed to test whether this effect is replicable when participants are asked judgments of contingencies as opposed to conditional probabilities. Hence, the goal was this experiment was to replicate the results of Price and Yates using test questions that asked about the contingency between the effect and each cause.

In Experiment 3, it was found that ratings were higher when participants made two separate decisions about the outcome of each trial. Specifically, participants in the *two-decision* group were generally better able to discriminate the blocked contingencies. This group also appeared more likely to use larger numbers when judging the perfect blocking contingency. Accordingly, the two-decision format was employed for this experiment.

Method

Participants

Twenty-four McGill University students participated in this study. Three participated for course credit and twenty-one were paid $5. Twelve participants were male and twelve were female.

Cover Story

This experiment employed a 1E-2C cover story. Participants assessed the extent to which two chemicals affected the survival of a strain of bacteria. However, they were first informed about the presence or absence of the bacteria and then about the presence or absence of each of the two chemicals. This preparation conforms to Waldmann and Holyoak's (1992) definition of a diagnostic reasoning task.

Participants were informed that four new strains of bacteria were discovered and that the survival of the bacteria could be affected by certain pairs of chemicals. To test whether the chemicals affected the bacteria's survival, the bacteria were placed in petri dishes containing one chemical, the other chemical, both chemicals, or neither chemical. The presence or absence of each chemical was later verified.

Procedure

After reading and signing the consent form, participants were seated in front of a computer. The instructions that appeared on the computer screen are included in the Appendix (section IV). In the first screen, participants read that there were four strains of bacteria under investigation, and that for each strain scientists would carry out a series of experiments to evaluate whether the bacteria's survival was affected by certain pairs of chemicals. For each experiment, the bacteria were placed in a culture containing one chemical, the other chemical, both chemicals, or neither chemical. The scientists then checked if the bacteria survived or did not survive and verified which chemical or chemicals were in the culture. Participants read that the experiment was done this way so that the experiments could be done blindly and so that the scientists' biases would not influence the results of these experiments or how they were interpreted. The second and third screens explained that a chemical might make it more likely, less likely, or might not affect the likelihood that the bacteria would survive. The fourth screen explained that on each trial participants would first be informed about the presence or absence of the bacteria. They also read that they would then have to enter whether they thought each of the two chemicals was in the culture and that they would be given feedback about the correctness of their predictions. Participants were also encouraged to keep track of what happened in both the presence and absence of the bacteria. The fifth screen stated that participants would be asked to rate how strongly the bacteria's survival was affected by each chemical on a scale ranging from -100 to +100.

After answering any questions that the participants might have had about the instructions, the experimenter left the room and allowed the participants to proceed with

the task. Similar to the previous experiments, each game had 48 trials divided into three 16 trials blocks. Participants gave their estimates after the second block and at the end of each game.

During each trial, participants were informed about the presence or absence of the bacteria and were then asked in two separate questions whether they thought each of the two chemicals was in the culture. After entering whether an individual chemical was in the culture, participants were informed whether their guess was correct or incorrect, and read what outcome had occurred. At the end of the trial, participants pressed the space-bar to begin the next trial.

Ratings

Ratings of the Blocked Contingencies

Table 12 displays the means and standard errors for the first and final ratings of the blocked and blocking contingencies, and Figure 10 contains the final ratings. Participants discriminated a moderately positive contingency from a zero contingency when the blocking contingency was zero. However, the perfect contingency appears to have lowered ratings of the moderately positive but not the zero contingency.

Statistical analyses only partially support these impressions. The main effect for the blocked contingencies was significant, F (1, 23) = 88.00, but the main effect for the blocking contingencies and the blocked by blocking interaction were not significant, F (1, 23) = 1.41, F (1, 23) = 1.30, respectively. To directly test causal model theory's prediction that blocking would be observed in the 0.5 contingency, the final judgments of this contingency were compared when it was paired with a zero and perfect contingency. This analysis, however, failed to attain significance, t (23) = 1.57. Finally, the difference between the first and final ratings was not significant, F (1, 23) = 2.13, and this factor did not interact with any other variable, all Fs < 1.

The main effect for the blocked contingencies supports the contention that participants reliably discriminated the moderately positive contingency from the zero contingency. However, the lack of a main effect for the blocking contingencies, the lack

Table 12

Experiment 4: 1E-2C: ?EC

Mean Estimates of the Blocked and Blocking Contingencies after 32 (first rating) and 48 trials (final rating). The number before the slash refers to the blocked chemical's contingency and the number after the slash to the blocking chemical's contingency.

BLOCKED CONTINGENCIES				
	First Ratings		*Final Ratings*	
Condition	**Mean**	**SEM**	**Mean**	**SEM**
0.5/0	46.9	6.4	48.5	6.1
0.5/1	29.0	9.6	32.7	8.3
0/0	- 5.4	8.1	0.4	6.8
0/1	- 6.5	9.8	- 0.6	8.6
BLOCKING CONTINGENCIES				
0.5/0	- 4.8	9.0	- 4.0	9.2
0.5/1	78.5	8.0	76.9	9.4
0/0	14.7	8.1	7.9	7.3
0/1	74.2	8.4	75.0	8.2

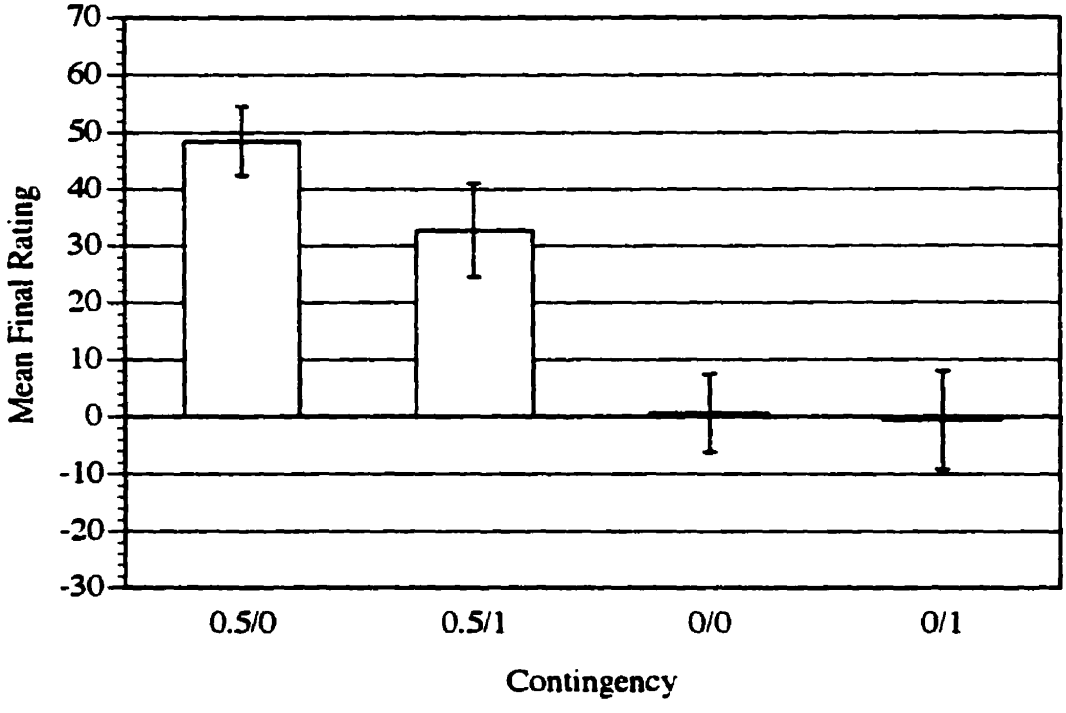


Figure 10. Final judgments and standard errors of blocked contingency in Experiment 4 (1E - 2C).

of a blocked by blocking interaction, and the results of a direct test of blocking in the 0.5 contingency suggest that ratings were not influenced by the blocking contingencies. In other words, no significant blocking effect was observed in this experiment; this finding is consistent with the R-W model.

Ratings of the Blocking Contingencies

The results in Table 12 imply that participants discriminated the perfect and zero contingencies. Ratings of the zero contingency, however, appear to be lower when the blocked contingency was moderately positive than when it was zero. Similar to Experiment 3, judgments of the perfect contingency were below 90 and were lower than the ratings given in Experiments 1 and 2. Inspection of the data, however, failed to reveal any participants giving low (i.e., zero or lower) ratings to the perfect contingency.

Statistical analyses only partially confirm the described impressions. The main effect for the blocking contingencies was significant, F (1, 23) = 37.49. However, the main effects for the blocked contingencies and time were not statistically significant, F (1, 23) = 2.36, F (1, 23) < 1, respectively. Furthermore, none of the interactions attained significance, highest F (1, 23) = 3.65. Thus, participants discriminated strong contingencies from weak ones, but ratings of the blocking contingencies were not influenced by the blocked contingencies.

Discussion

Two factors were varied in the four experiments described: causal scenario and causal order. As opposed to previous research, the same cues were used in all experiments but varied in whether they were a) defined as two causes of one effect or one cause of two effects, and b) whether information about the cause was presented before the effect or the effect before the cause. These variables were manipulated in an attempt to discover conditions under which blocking would be observed. In Experiments 1 and 2, cue competition occurred when two cues were presented before a single outcome regardless of whether how the cues were defined. On the other hand, cue competition was

not observed in Experiments 3 and 4 when a single cue preceded the presentation of two outcomes, which is consistent with prior findings (Baker & Mazmanian, 1989; Price & Yates, 1995; Rescorla, 1991). The initial results thus imply that blocking is influenced by the number of antecedent cues, and is consistent with the predictions from the R-W model.

However, the number of antecedent cues was confounded by the nature of the intervening questions during contingency acquisition. When two cues preceded a single outcome (i.e., in 2C-1E and 2E-1C), participants were presented with information about the two antecedent cues and *made a single decision or prediction* about whether the outcome occurred. However, when a single cue preceded two outcomes (i.e., in 1C- 2E and 1E-2C) they were asked *two separate questions* about whether each of the two outcomes had taken place. The problem with this method of having participants predict the two outcomes in that it encourages them to keep track of the blocked and blocking contingencies independently.[3] In other words, the absence of blocking in 1C-2E and 1E-2C could be an artefact of having participants make two outcome predictions on each trial whereas blocking was observed in 2C-1E and 2E-1C because only one prediction was made on a trial. This concern was examined in Experiments 6 and 7.

In addition to controlling for the intervening questions presented between the cue and outcome, two other issues need to be addressed. The first is whether participants encoded the causal model correctly. This issue becomes important when causal order is manipulated (Waldmann & Holyoak, 1997). Otherwise, it could be argued that the cue competition observed in Experiment 2 and the absence of this effect in Experiment 4 could be due to participants encoding the causes as effects or effects as causes. This topic was addressed in Experiment 5.

---

[3] This argument could also apply to the single-decision group in Experiment 3. Rather than making two separate decisions, participants had to predict which conjunction of four possible outcomes occurred. Although the single-decision format was adopted as a control for the number of predictions made on each trial, it seems reasonable to assume that this format also implicitly requires that the volunteers monitor the blocked and blocking contingencies individually.

A second issue is the role of the test question. As shown in Table 5, the cover story was confounded with the test question. That is, in each experiment the test question corresponded with the order in which the cues were presented. This issue is less critical given previous evidence that similar judgments will be found if people are asked ?EC or ?CE causal questions (Matute et al., 1996). This methodological issue was also investigated in Experiment 5.

EXPERIMENT 5

In Experiment 4, blocking was not detected when one effect preceded two causes. The present experiment attempted to replicate this absence of cue competition using a ?CE (predictive) instead of an ?EC (diagnostic) test question. If the results obtained are similar to those observed in Experiment 4, this evidence would support Matute et al.'s (1996) argument that ?CE and ?EC are interpreted as being the same question. An additional motivation for the present experiment was to test whether participants had appropriately encoded the causal model, that is as 1E-2C as opposed to 1C-2E. If participants understood that the chemicals affected the bacteria's survival despite information about the bacteria being presented before each of the two chemicals, then this would exclude the possibility that the lack of cue competition observed in Experiment 4 could be attributed to participants not comprehending the described scenario and would address a criticism raised in Waldmann and Holyoak (1997).

The final goal of the present experiment was to obtain a list of alternative causes. As will be described later, one goal of Experiment 7 was to test whether the blocking effect observed in 2E-1C could be attributed to alternative causal models. This objective was accomplished by a) first obtaining a list of alternative causes in this experiment and in Experiment 6, and b) informing participants that these causes were controlled for in Experiment 7. These goals were accomplished by having participants in this experiment fill out a questionnaire in which they specified 1) which cues were causes and which ones were effects and 2) whether they believed there were alternative causes that could have

influenced the events they observed. This information was collected after completing the experimental task so that it would not influence participants' ratings.

Method

Participants

Twenty-four McGill University students were paid $5 to participate in this study. Six were male and eighteen were female.

Cover Story and Procedure

The cover story and procedure were similar to those used in Experiment 4. The instructions are included in the Appendix (section V). As can be seen, the main difference between Experiments 4 and 5 was in the question asked when requesting the participants' ratings. Whereas Experiment 4 used an ?EC question ("Rate how strongly the production of [chemical] was affected by [bacterial strain]"), this experiment used a ?CE question, namely "Rate how strongly [bacterial strain] affected [chemical's] production." The second difference between the present experiment and Experiment 4 is that participants were administered a post-experimental questionnaire containing three questions. This questionnaire is displayed in the Appendix (sectionVI). The first question asked whether the scientists were examining whether the chemicals were affecting the bacteria, the bacteria were affecting the chemicals, or neither of these two options. If participants chose the third option, they were requested to state what they thought was the correct answer. The second question stated that a white ball on a pool table was hit by a pool cue, which caused the ball to move; the participant was asked to state whether the chemicals were analogous to the pool cue, the white ball, or neither of these two options. Again, if participants chose the third option they were requested to state what they thought was the correct answer. Finally, the third question attempted to assess participants' alternative causal models. Participants were asked if there were any variables not taken into account that could have influenced the results that the scientists obtained. In this experiment and Experiment 4, the instructions on the computer screen stated only that the scientists were

blind to the experimental conditions but did not state whether any steps were taken to control for alternative causes. Thus, the third question was asked in an attempt to independently obtain participants' alternative causal models so that they could be controlled for in Experiment 7.

Results

Ratings of Blocked Contingency

Data from four participants were removed because they gave zero or negative judgments to the perfect contingency in both the 0.5/1 and 0/1 conditions. Table 13 contains the first and final estimates of the blocked and blocking contingencies, and Figure 11 contains the final ratings of the blocked contingencies. The results imply that participants generally discriminated moderately positive and zero contingencies. For the first set of ratings, judgments of the moderate but not the zero contingency appear elevated when it was paired with the perfect blocking contingency. However, the opposite pattern is observed for the final ratings: judgments of the 0 contingency appear elevated when it was paired with a perfect contingency.

Statistical analyses confirm these impressions. The main effect for the blocked contingencies was significant, $F(1, 19) = 46.34$, but the main effects for the blocking contingencies and time were not, highest $F(1, 19) = 1.01$. Although the blocked by blocking, blocked by time, and blocking by time interactions were not statistically significant, all Fs $(1, 19) < 1$, the blocked by blocking by time interaction was significant, $F(1, 19) = 5.39$. Two tests of simple interaction effects were conducted to investigate the nature of this three-way interaction; with a Bonferroni correction, the rejection criterion was 0.025. The blocked by blocking interaction failed to attain significance at either the first or final ratings, $F(1, 19) = 1.87$, $F(1, 19) < 1.4$, respectively. Thus, the three-way interaction arose from a crossover effect of having judgments of the 0.5 contingency being slightly elevated during the first ratings, whereas the opposite pattern was observed in the final ratings. Taken together, there is little evidence that participants' ratings of the blocked contingencies were significantly lowered by the blocking contingencies.

Table 13

Experiment 5: 1E - 2C: ?CE

Mean Estimates of the Blocked and Blocking Contingencies after 32 (first rating) and 48 trials (final rating). The number before the slash refers to the blocked chemical's contingency and the number after the slash to the blocking chemical's contingency.

BLOCKED CONTINGENCIES				
	First Ratings		*Final Ratings*	
Condition	**Mean**	**SEM**	**Mean**	**SEM**
0.5/0	40.3	8.7	53.8	7.0
0.5/1	56.8	6.9	56.4	6.4
0/0	-1.4	10.2	-11.0	7.1
0/1	-4.0	10.9	2.3	8.2
BLOCKING CONTINGENCIES				
0.5/0	6.5	8.5	-1.3	5.7
0.5/1	91.3	4.0	92.0	3.6
0/0	-7.9	7.0	11.3	6.6
0/1	78.5	9.6	91.5	3.0

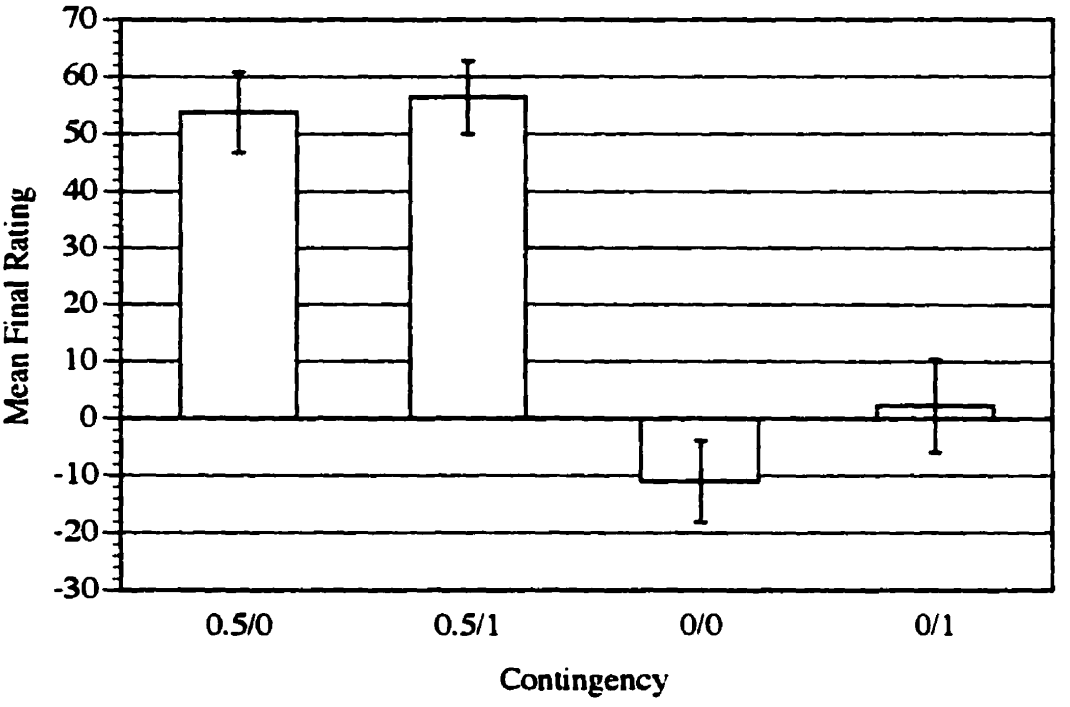


Figure 11. Final judgments and standard errors of Blocked contingency in Experiment 5 (1E - 2C: ?CE)

Ratings of Blocking Contingencies

Inspection of Table 13 reveals that participants discriminated the perfect contingencies from the zero contingencies. For the first set of ratings, judgments of the zero contingency appear higher when it was paired with a zero blocked contingency. However, for the final ratings, judgments of the zero contingency appear higher when it was paired with a moderate blocked contingency.

Statistical analyses confirm these impressions. The main effect for the blocking contingencies was significant, F (1, 19) = 222.44, but the main effects for the blocked contingencies and for time were not significant, highest F (1, 19) = 2.44. The blocked by blocking, blocking by time, and blocked by blocking by time interactions did not attain significance, highest F (1, 19) = 1.43. However, the blocked by time interaction was significant, F (1, 23) = 6.64. Two tests of simple main effects were conducted using a Bonferroni corrected rejection criterion of 0.025. These follow-up analyses revealed that the blocked by time interaction arose because the blocked contingencies had a non-significant influence on the first set of judgments, F (1, 23) = 4.22, although this effect would have been statistically reliable as a planned comparison. The blocked contingencies did not have a detectable effect on the final ratings, F (1, 19) = 1.68. Thus, the blocked contingencies appear to have a modest influence on the blocking contingencies, although this effect is not apparent for the final judgments.

Taken together, these results are consistent with the previous experiments in which participants readily discriminated perfect and zero blocked contingencies, but were minimally influenced by the blocking contingencies.

Questionnaire

The first question asked whether the scientists were evaluating whether the chemicals were affecting the bacteria, the bacteria were affecting the chemicals, or neither of these two possibilities. Nineteen of the twenty participants chose the first option (i.e., the correct answer), and one chose the second option. Inspection of this person's questionnaire revealed that the correct answer was chosen for the second question even

though the response for question 1 was wrong. Furthermore, this participant's ratings were similar to the rest of the sample. Since this person's judgments were representative of the sample, these data were included in the analyses for the blocked and blocking contingencies.

For the second question, seventeen participants stated that the chemicals were analogous to the pool cue and three participants stated that they were analogous to the white ball. The three people who had chosen the incorrect answer to this question had chosen the correct answer for the first question. Thus, the majority of participants could make the analogy between the causes and effect in this experiment and those in another situation. Taken together, the responses to these two questions imply that participants understood which variables were the causes and effect.

For the third question, thirteen participants stated that there were alternative uncontrolled causes that were not taken into account whereas seven did not. As shown in Table 14, the responses fell into four categories. The majority of participants listed environmental conditions in the laboratory and interactions between chemicals as alternative causes. Examples of environmental conditions were temperature and the age of the bacterial sample whereas interactions between chemicals referred to whether the two chemicals together would form a compound that is more or less potent than either chemical alone. One participant wrote that the chemicals in the present experiment should have been compared to ones whose actions are known, and two participants gave responses that did not fall into the prior three categories; these responses were classified as "Other."

Discussion

This experiment used a preparation in which one effect preceded two causes (i.e., 1E-2C) and participants were asked a ?CE test question at the time of judgment. Similar to Experiment 4, a strong cause did not block ratings of a weaker one. Apart from replicating the results of 1E-2C: ?EC, this experiment demonstrated that the lack of blocking could not be accounted for by the argument that participants had not understood

Table 14

Categories of alternative causes listed by participants in Experiment 5. The four categories were Laboratory Conditions, Chemical Interactions, Experimental Design, and Other. Laboratory Conditions consisted of a series of examples; the number of respondents listing each example is shown. Other consists of causes that were either vague or were not classified in the other categories.

Category	**Examples**	**Number of Respondents**
Laboratory Conditions		*Total: 7*
	Temperature of culture	5
	Time of day or age of culture	3
	Contaminants in the air or in the culture	3
	pH of chemicals	2
	Nutrients in the culture	1
Interactions Between Chemicals		*Total: 4*
1) "interactions [between] the two chemicals which neutralizes their properties 2) "possible reactions between the drugs resulting in the formation of chemical C" 3) "interaction of the two chemicals" 4) "combination of chemicals"		
Experimental Design		*Total: 1*
"They should have compared it to other chemicals whose actions are well known"		
Other		*Total: 2*
1) "they didn't consider the presence of other things." 2) "Having not been given more details of the study, it is my assumption that any number of variables could have possibly skewed the results."		

the causal model. Almost all of the participants correctly stated that the chemicals had acted on the bacteria. Furthermore, the majority of participants could extend this comprehension to a novel situation in which they had to state the analogy between the cause and effect in this experiment and a pool cue causing a ball to move. Participants were thus able to state which variables were causes and effects on two different measures. This evidence thus suggests that participants had appropriately encoded the causal model as two causes of one effect, even though the contingencies were presented as effects before the causes.

In addition to testing whether the causal models were appropriately encoded, this experiment was designed to evaluate the role of the test question. According to causal model theory, blocking should occur in this preparation if an ?EC test question, as opposed to a ?CE question, is asked. On the other hand, Matute et al. (1996) would argue that similar results should be obtained in both cases because ?CE and ?EC are synonymous. The blocking effect did not attain significance regardless of which test question was asked, which suggests that the directionality of the test question may not play as prominent a role in determining whether blocking is observed. This analysis is consistent with Matute et al. (1996).

Thus far, the issue of the people's understanding of the causal model when effects are presented before causes has been addressed, as well as the role of the test question. The most important issue, however, with regard to 1C-2E and 1E-2C is that they differ from the 2C-1E and 2E-1C in terms of how the two sets of tasks implicitly encourage people to keep track of the contingencies. Experiments 6 and 7 examined this issue.

EXPERIMENT 6

In the experiments described, blocking was observed when two cues preceded a single outcome but was not observed when one cue preceded two outcomes. These initial results from Experiments 1, 2, 3, and 4 implied that blocking is influenced by the number of antecedent cues. However, the number of cues is confounded by the intervening predictions between presentation of the antecedent and consequent cues. When two cues

were presented, participants made a single prediction about the trial's outcome. But when one cue was presented, they had to predict each of the two outcomes separately. This method of asking two separate questions on each trial might implicitly coax participants to keep track of the two cues independently because they have to make predictions about them.

The primary motivation for this experiment was to examine whether similar results would be obtained using a preparation that does not encourage people to monitor each contingency separately. This goal was attained by having participants passively observe the contingencies. Thus, rather than making two separate decisions on each trial, participants were informed about the events on each trial but did not have to predict what would happen. If the results with this preparation are similar to those of Experiments 3 and 4, then the lack of blocking in those experiments a) would not be due to the task imposing an artificial constraint on the participants and b) would provide stronger support for the R-W model. On the other hand, if cue competition did occur, then the results would imply that the role of the intervening questions between antecedent and consequent cues would be an important variable to consider when testing for blocking effects. In this case, further experimentation would be needed in which both causal order and causal scenario are manipulated in the same experiment, but in which all participants observe the contingencies in a similar manner.

A second motivation for this experiment was to further ensure that the lack of blocking observed in Experiments 4 and 5 could be attributed to participants not comprehending the causal model. Although the results in Experiment 5 imply that participants had properly encoded the causal model as two causes of one effect, it is important to replicate this finding.

Method

Participants

Twenty-four McGill University students participated in this study. Twenty-three were paid $5 and one participated for course credit. Fifteen participants were female and nine were male.

Cover Story and Procedure

The same set of instructions used in Experiment 5 was used here; the one screen of instructions that differed, the fourth one, is included in the Appendix (section VII). Similar to Experiments 4 and 5, participants were presented a 1E-2C scenario in which two chemicals affected a bacterial strain's survival and information about the effect was presented before the causes. Furthermore, this experiment was similar to Experiment 5 in that participants were presented with an ?CE question when their judgments were requested. This experiment, however, differed from Experiments 4 and 5 in terms of the events that took place on each trial between presenting the effect and the two causes. Namely, participants were first informed about the presence or absence of the bacteria. After pressing the space-bar, they were informed about the presence or absence of each of the two chemicals, without having to make any predictions about the outcome on each trial.

Participants were administered the same post-experimental questionnaire used in Experiment 5, which is included in the Appendix (section VI). Briefly, the first question asked participants if they understood whether the chemicals were affecting the bacteria or if the bacteria were affecting the chemicals. The second question asked if the chemicals were akin to a pool cue or the movement of a ball on a pool table. Finally, the third question asked whether participants believed that there were alternative causes that could account for the results observed.

Results

Ratings of Blocked Contingencies

Table 15 contains the first and final estimates of the blocked and blocking contingencies and Figure 12 displays the final ratings of the blocked contingencies. Participants discriminated the moderate and zero contingencies. But judgments were lowered when the blocked contingencies were paired with a perfect contingency, and this effect was larger for the moderate contingency. Finally, the perfect contingency's influence on ratings of the zero contingency appears smaller in the final ratings.

Statistical analyses support these impressions. The main effects for the blocked and blocking contingencies were significant, F (1, 23) = 19.26; F (1, 23) = 26.48, respectively, but the main effect for time was not, F (1, 23) = 1.36. Furthermore, the blocked by time and blocked by blocking by time interactions were not significant, highest F (1, 23) = 1.89. However, the blocked by blocking and blocking by time interactions were significant, F (1, 23) = 10.99, F (1, 23) = 4.59, respectively. A Bonferroni corrected rejection criterion of 0.0125 was used for the four tests of simple main effects. From the table and graph, it is apparent that the blocked by blocking interaction arose because the cue competition effect was larger for the 0.5 than for the 0 contingency, although both were significant, F (1, 23) = 8.61, F(1, 23) = 35.42, respectively. Furthermore, the blocking by time interaction arose because the cue competition effect was smaller for the final than for the first set of estimates, F (1, 23) = 15.39, F (1, 23) = 39.19, respectively. In summary, participants discriminated the two blocked contingencies but were also influenced by the blocking contingencies.

Ratings of Blocking Contingencies

As shown in Table 15, perfect and zero contingencies were discriminated and consistent ratings were given both after 32 and after 48 trials. However, the ratings of the zero blocking contingency appear lower when it was paired with the moderate than with a zero blocked contingency.

Table 15

Experiment 6: 1E-2C: ?CE (Passive Observing)

Mean Estimates of the Blocked and Blocking Contingencies after 32 (first rating) and 48 trials (final rating). The number before the slash refers to the blocked chemical's contingency and the number after the slash to the blocking chemical's contingency.

BLOCKED CONTINGENCIES				
	First Ratings		*Final Ratings*	
Condition	**Mean**	**SEM**	**Mean**	**SEM**
0.5/0	33.1	5.4	33.5	5.7
0.5/1	- 22.4	6.5	- 19.3	6.7
0/0	6.7	3.6	2.1	3.0
0/1	- 27.1	7.1	- 14.8	10.0
BLOCKING CONTINGENCIES				
0.5/0	- 14.6	6.0	- 15.6	5.3
0.5/1	87.7	4.9	88.9	4.8
0/0	4.0	3.8	- 0.4	2.8
0/1	90.6	3.9	80.8	8.3

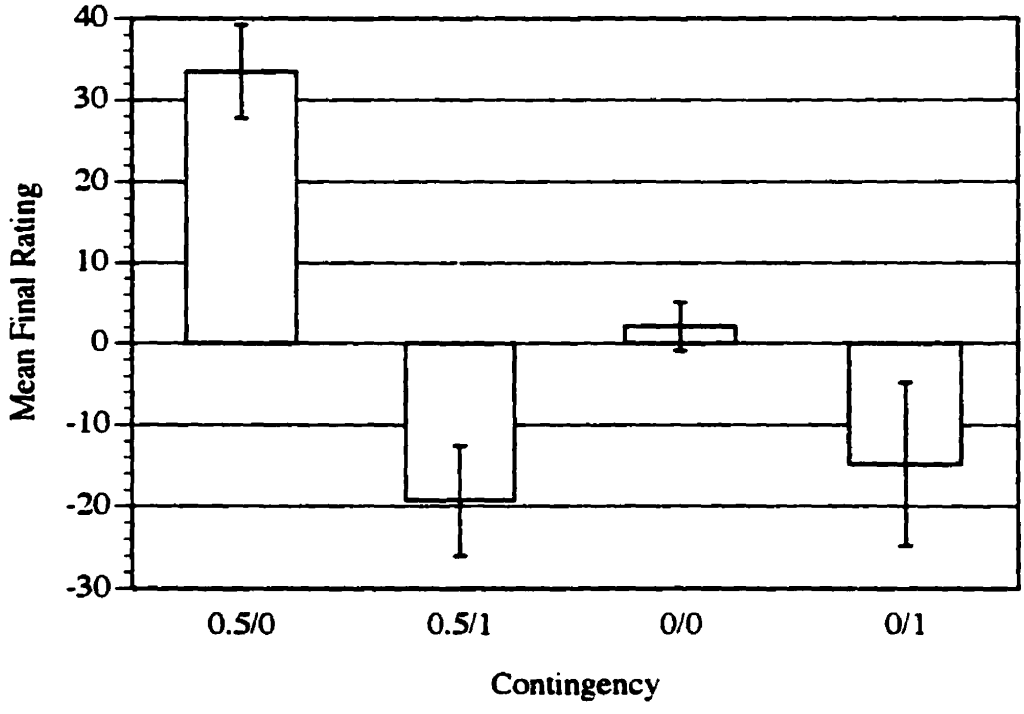


Figure 12. Final judgments and standard errors of Blocked contingency in Experiment 6 (1E - 2C: ?CE, Passive Observing).

Statistical analyses only partially confirm these impressions. The main effect for the blocking contingencies was significant, F (1, 23) = 486.77. However, neither the main effect for the blocked contingencies nor for time attained significance, highest F (1, 23) = 2.98. Furthermore, none of the interactions attained significance, highest F (1,23) = 3.72. Thus, perfect and zero contingencies were discriminated, but were not influenced by the blocked contingencies.

Questionnaire

In response to the first question, all participants correctly stated that the scientists were investigating how the chemicals were affecting the bacteria. Similarly, for the second question, twenty-three stated that the chemicals were analogous to the pool cue; only one subject thought there was no similarity between the variables in the computer task and the pool cue and ball. Thus, participants clearly understood which variable was the cause and which was the effect.

For the third question, eighteen participants stated that there were alternative uncontrolled variables that could account for the results observed and six who did not. As shown in Table 16, the vast majority listed laboratory conditions as alternative causes, with temperature and nutrients in the culture being mentioned most frequently. Two participants wrote about possible interactions between the chemicals, and two referred to the methodology in the experiment. In particular, one person questioned the validity of classifying the bacteria as surviving or not surviving as opposed to counting the number of bacteria in the sample; the other mentioned taking into account the culture's normal survival time before adding or not adding the chemical or chemicals. Finally, four participants listed causes that did not fall into the previous three categories. Two responses were vague: one referred to the base rate of the bacterial sample's survival and one questioned how generalizable the results would be to humans if the bacteria were studied in isolation.

Table 16

Categories of alternative causes listed by participants in Experiment 6. The four categories were Laboratory Conditions, Chemical Interactions, Experimental Design, and Other. Laboratory Conditions consisted of a series of examples; the number of respondents listing each example is shown. Other consists of causes that were either vague or were not classified in the other categories.

Category	**Examples**	**Number of Respondents**
Laboratory Conditions		*Total: 14*
	Temperature of culture	9
	Nutrients in the culture	4
	Concentration or relative amount of chemicals	3
	Contaminants in the air or in the culture	2
	Strength or quantity of bacteria	2
	Light or dark room	2
	pH of chemicals	1
Interactions Between Chemicals		*Total: 2*
"Two chemicals together can react." "Interactions (positive and negative) between the chemical combinations that resulted in an overall positive or negative effect on the bacteria."		
Experimental Design		*Total: 2*
"Counting bacteria [as opposed to classifying as present or absent]" "Normal survival time of bacteria"		
Other		*Total: 4*
"chance of bacteria surviving on its own" "The bacteria are in a culture isolated. If they were in the human body, their survival may be influenced by other factors." "presence of other chemicals which may inhibit growth" "other chemicals that could have affected the development (or not) of the bacterias [sic]"		

Discussion

This experiment examined whether blocking would be observed in 1E-2C when participants were not implicitly encouraged to monitor the two causes independently. Contrary to Experiments 4 and 5, blocking occurred. Similar to Experiment 5, participants had comprehended the causal model and thus the lack of blocking in Experiments 4 and 5 can not be accounted for by the simple explanation that causes were encoded as effects and vice versa.

It seems reasonable to conclude that an important variable influencing blocking is whether or not participants actively try to predict the presence or absence of the consequent cues. The R-W model does not anticipate and does not readily account for this finding. Causal model theory, however, can provide a plausible explanation. From an information processing perspective it could be argued that participants in Experiments 4 and 5 would have reasoned cause to effect, but the task's cognitive demands were quite large and thus prevented them from using this strategy when acquiring the contingencies. But in Experiment 6, the cognitive demands were lowered and participants could then reason from cause to effect; since causes compete, blocking would then be expected.

The R-W model, however, can be modified to accommodate the results from Experiment 6. It is assumed that the link between cues and outcomes is unidirectional (Matute et al., 1996). That is, the directionality of reasoning is from cues to outcomes but not from outcomes to cues. However, Matute et al. (1996) have argued the link between cues and outcomes need not be linked solely in one direction. If this constraint is removed so that the associative bond between cues and outcomes is bidirectional, then this modified version of the R-W model shares a property similar to Miller's comparator hypothesis (Miller & Matzel, 1988). By having a bidirectional link, the modified R-W model anticipates blocking when participants are presented with 1E-2C. This prediction is made because the theory assumes that, after acquisition, the contingency between two events (e.g., contingency A) are compared to the contingencies of other cues (e.g., contingencies B and C). If A is the strongest contingency, then it will be judged as strongly positive. But if contingencies B or C are higher, A will be judged as being weak

because its contingency is small relative to these other contingencies. In 1E-2C, cue competition would be expected in 0.5/1 and 0/1 because both the moderate and zero blocked contingencies would be weak when compared to the perfect blocking contingency.

By having cues and outcomes linked bidirectionally, the comparator approach makes some interesting predictions about whether cue competition should be observed. As mentioned in the Introduction, the comparator hypothesis states that all associations are learned about during acquisition and blocking is determined by the post-acquisition treatment. Matute et al. (1996) have argued that the critical treatment is the wording of the test question. However, the results from the experiments outlined here were not consistent with this line of reasoning. Although there was some evidence that the test question might influence blocking, it was not as strong as the authors claim. Instead, the results from these experiments imply that the processing of information between the cues and outcomes is an important factor.

The modified R-W model can account for the results by stating that competition between contingencies is less likely to be observed when the task is set up so that it demands that participants monitor contingencies independently during acquisition. This method of presentation was used in Experiments 3 to 6, as well as in other experiments in the literature such as Baker and Mazmanian (1989) and Price and Yates (1995). In all of these experiments, blocking was not observed. It is possible that blocking is more likely to be observed when the demands of the task are minimized, for example by passively observing the contingencies. An implication of this assertion is that cue competition should be observed regardless of both causal scenario and causal order, when contingencies are acquired in this manner. This possibility was examined in the following experiment.

EXPERIMENT 7

There were two goals for this investigation. The first was to contrast the predictions of three models of how causal scenario and causal order influence cue

competition. To allow for a direct comparison of their influence, these variables were manipulated within the same experiment rather than between experiments. The second motivation was to use the same method of acquiring the contingencies in all conditions. It was important that this variable be equated across conditions. In an attempt to rule out the possibility that alternative causal models would account for any blocking effects, particularly in 2E-1C, participants were informed that various measures were taken to ensure that the events they observed could not be attributed to extraneous variables. These variables were chosen based on the alternative causes listed by participants in Experiments 5 and 6.

As shown in Table 17, causal model theory, the R-W model and the modified R-W model each anticipates a different pattern of results. To recapitulate, causal model theory states that blocking is dependent on both the causal scenario and causal order. When there are two causes of one effect, blocking should occur regardless of causal order. This pattern is predicted because participants putatively reason from cause to effect. Blocking is anticipated because there are multiple causes and because causes compete for association with an effect. But when there is one cause of two effects, blocking is not expected. An important criterion, however, has to be met has to be met in 2E-1C: that participants do not resort to alternative causal models.

The R-W model predicts that blocking is dependent only on the number of predictive cues. Thus, blocking is expected when two predictors precede a single outcome but is not anticipated when a single cue precedes two outcomes, regardless of the cues' causal status. According to the revised R-W model, however, blocking is expected in all four conditions. In other words, blocking is anticipated regardless of causal scenario and causal order.

Design

This experiment was designed to test how ratings of the blocked contingency would be influenced by whether causes preceded effects or vice versa, and by whether two cues predicted the presence or absence of a single cue as opposed to one cue

Table 17

Predictions of how causal scenario and causal order influence blocking. The anticipated results are derived from causal model theory, the Rescorla-Wagner model, and a modified version of the Rescorla-Wagner model.

Model	**Experimental Condition**			
	2C - 1E	1C - 2E	2E - 1C	1E - 2C
Causal Model Theory	Blocking	No Blocking	No Blocking	Blocking
R - W model	Blocking	No Blocking	Blocking	No Blocking
Modified R -W model	Blocking	Blocking	Blocking	Blocking

predicting the presence or absence of two cues. Thus, a four factor mixed design was employed and is shown in Table 18. The first factor, causal order, was a between-groups factor containing two levels. This factor refers to whether participants observed the contingencies in the order CE or EC. The second factor, predictor, was a within-participants factor containing two levels. Predictor refers to whether participants were presented with two predictors of a single outcome or one predictor of two outcomes, regardless of whether cues were causes or effects. Thus, participants in the CE group were presented with both a 2C-1E scenario and a 1C-2E scenario whereas those in the EC group were presented with 2E-1C and 1E-2C. Within each group, the order in which the scenarios was administered was counterbalanced across participants.

In each of these conditions, a moderately positive blocked contingency ($\Delta p = 0.5$) was paired with a blocking contingency that was either weak or perfect. Thus, the third factor was the strength of the blocking contingency. A blocked contingency of zero was not used because participants were clearly able to distinguish moderate and weak contingencies in all of the previous experiments. To minimize the number of conditions, this control was not included in this experiment. The final factor was time, that is whether ratings were collected after 32 or 48 trials.

Method

Participants

The participants were 48 undergraduates from McGill University. Forty-one were paid $5 and seven participated for course credit. Participants were randomly assigned to the CE or EC groups. Both groups contained 24 participants.

Procedure

After reading and signing the consent form, participants were seated in front of the computer. They were then told that the instructions would appear on the computer screen and to hit the space-bar to advance from one screen to the next. Furthermore, they were told that they could ask the experimenter questions at any time during the

Table 18

Design for Experiment 7

Between-Groups Factor: **Causal Order**	**Within-Groups Factors**		
	Predictor	*Contingency*	*Time*
CE	2 cues => 1 outcome	0.5/0	Ratings collected after 32 and 48 trials
EC	1 cue => 2 outcomes	0.5/1	

instructions phase. Six screens of instructions were presented on the screen. The instructions from each condition are included in the Appendix (section VIII).

In the CE group, half of the participants were presented with 2C-1E followed by 1C-2E and the others received the reverse order. In 2C-1E, the first screen stated that two pairs of chemicals and two strains of bacteria were discovered and that for each strain scientists were conducting experiments to evaluate whether a pair of chemicals affected the bacteria's survival. To do this, the bacteria were placed in petri dishes and either one chemical, the second chemical, both chemicals, or neither chemical was added. The sample was later classified as having survived or not survived. The second and third screens explained that it was not known what effect each chemical may have on the bacteria. That is, each chemical could have no effect or could make it more or less likely that the bacteria would survive. The fourth screen stated that various steps were taken to ensure that the results would be reliable. Namely, participants read that the chemicals did not interact, that similar concentrations of bacteria and chemicals were used in all conditions, that optimal environmental conditions such as temperature and lighting were established prior to the study, that all equipment was sterile, and that the test classifying the bacteria as surviving or not surviving was as reliable as counting the bacteria before and after adding the chemical or chemicals. In the fifth screen, participants read that, on each trial, they would press the space-bar to initiate a trial and that they would be told which chemical or chemicals were added. After pressing the space-bar, they would be told whether or not the bacteria survived. In the final screen, participants were told that they would be asked to rate how strongly each chemical affected the bacteria's survival. This screen was the same as that used in Experiment 1.

In the 1C-2E condition, the instructions were similar to those used in Experiment 3. Namely, participants were first informed that scientists were evaluating the extent to which each of two strains of bacteria produced two chemicals. This goal was accomplished by adding or not adding the bacteria to an environment that simulated the mammalian digestive system and then testing for the presence or absence of each of the two chemicals. In the second and third screens, participants read that the bacteria could

have no effect on each chemical's production or could make it more or less likely that each chemical would be produced. In the fourth screen, participants read that steps were taken to ensure that the results would be reliable. This screen was similar to that used in the 2C-1E condition. The fifth screen stated that on each trial, participants would first be informed about whether the bacteria were added or not added. After pressing the space-bar, they would be informed which chemical or chemicals were produced. The final screen explained that they would be asked to rate how strongly the bacteria affected the production of each chemical. This screen was the same as that used in Experiment 3.

In the 2E-1C condition, participants read that two strains of bacteria had been discovered that existed in the mammalian digestive system and that scientists were studying whether the bacteria aided in, interfered with, or had no effect on the production of pairs of chemicals. This was accomplished by obtaining blood samples from laboratory rats and verifying the presence or absence of each chemical. The second and third screens explained that the bacteria might make it more or less likely that a chemical would be produced or could have no effect on a chemical's production. The fourth screen explained that laboratory rats were used instead of people so that variables such as diet, drug history, lifestyle factors, socio-economic status could be controlled; that the study was conducted under aseptic conditions and that environmental conditions such as temperature, the rats' age, time since the last meal, and diet were constant; that hormonal levels were constant; that the tests for the bacteria and chemicals were done at the same time each day; that the chemicals did not interact; and that the test used was reliable. The fifth screen explained that on each trial participants would first be informed about which chemicals were produced. After pressing the space-bar, they would then be informed whether the bacteria were present or absent. The final screen stated that participants would be asked to rate how strongly the bacteria affected each chemical's production. This screen was the same as that used in Experiments 5 and 6.

In the 1E-2C condition, participants read that scientists were investigating whether certain pairs of chemicals affected the survival of different strains of bacteria. This was achieved by placing the bacteria in petri dishes containing one chemical, the other

chemical, both chemicals, or neither chemical. The scientists later verified whether the bacteria survived and which chemical or chemicals were in the petri dish. The second and third screens explained that each chemical could make it more likely or less likely that the bacteria would survive, or could have no effect on the bacteria's survival. The fourth screen stated that alternative causes such as environmental conditions and the reliability of the test were controlled (similar to that described in the previous three conditions). The fifth screen stated that on each trial, participants would first be informed about the presence or absence of the bacteria and after pressing the space-bar, the presence or absence of the two chemicals. The final screen explained that participants would rate how strongly each chemical affected the bacteria's survival.

After answering any questions that the participants might have had about the instructions, the experimenter left the room and allowed the participants to proceed with the task. Each game had 48 trials that were divided into three 16 trials blocks. On each trial, participants were informed about the events that occurred on each trial as detailed above. At the end of the trial, participants pressed the space-bar to begin the next trial. Participants gave their estimates after the second block and at the end of each game. At the end of each game, participants were administered the post-experimental questionnaire used in Experiments 6 and 7.

Results

Data from two participants were removed, both of whom were in the EC group. The first participant was removed because that person was administered the wrong experimental conditions. The second participant was removed based on her responses on the questionnaire and because, during debriefing, the participant revealed that she was convinced that there was deception in the experiment. The results of the blocked and blocking contingencies were similar regardless of whether this person's data were included. However, the following analyses were conducted with these data excluded because a central assumption in causal model theory is that participants learning about effects prior to the cause do not believe that there are multiple alternative causes for the

effect (Waldmann & Holyoak, 1992, 1997). Thus, the following analyses are based on a sample size of 22 for EC and 24 for CE.

Ratings of the Moderate Contingency

The first and final ratings of the moderate contingency are shown in Table 19, and the final judgments are displayed in Figure 13. In the CE group, judgments of the 0.5 contingency were moderately positive when the blocking contingency was zero, regardless of whether the cover story was 2C-1E or 1C-2E. In 2C-1E, judgments were lowered when the blocking contingency was perfect. This cue competition effect was consistent for both sets of ratings. However, in 1C-2E, the pattern of judgments was different for the two sets of ratings. For the first set, judgments were only somewhat lower when the blocking contingency was perfect. For the final ratings, this effect was larger, albeit not as large as in the 2C-1E condition. In the EC group, the judgments of the 0.5 contingency were moderately positive in both 2E-1C and 1E-2C when it was paired with a zero contingency. In both conditions, judgements were lowered when the 0.5 contingency was paired with a perfect contingency. These results were consistent for both sets of ratings. However, unlike the CE group in which the cue competition effect was larger when 2 predictors preceded one outcome, in the EC group this effect was larger when 1 cue preceded two outcomes.

Statistical analyses confirm these impressions. Of the within-subject main effects and interactions, only the blocking factor attained significance, $F(1, 44) = 88.64$, next largest $F(1, 44) = 3.28$. Similarly, the main effect for the between-subject factor, causal order, was not significant, $F(1, 44) < 1$. However, this factor interacted with the main effect of predictor, $F(1, 44) = 8.17$, the predictor by blocking interaction, $F(1, 44) = 7.17$, and the predictor by time interaction, $F(1, 44) = 5.32$. Causal order did not interact with either the main effects of blocking or time, nor the blocking by time or predictor by blocking by time interactions, largest $F(1, 44) = 2.37$.

A Bonferroni correction was used for the following six follow-up analyses. Thus, the rejection criterion was 0.008. The causal order by predictor interaction arose because

Table 19

Experiment 7
Mean Estimates of the Blocked Contingencies after 32 (first rating) and 48 trials (final rating) for both the CE and EC groups. The number before the slash refers to the blocked contingency and the number after the slash refers to the blocking contingency.

CE GROUP				
	First Ratings		*Final Ratings*	
Condition	**Mean**	**SEM**	**Mean**	**SEM**
2 Causes - 1 Effect (2C - 1E)				
0.5/0	32.3	8.5	35.6	7.0
0.5/1	-29.4	9.6	- 24.2	7.0
1 Cause - 2 Effects (1C - 2E)				
0.5/0	35.8	5.7	36.7	5.7
0.5/1	24.2	7.9	- 2.1	3.8
EC GROUP				
2 Effects - 1 Cause (2E - 1C)				
0.5/0	41.4	7.1	35.7	8.6
0.5/1	-6.1	11.3	-6.8	9.8
1 Effect - 2 Causes (1 E - 2C)				
0.5/0	39.5	5.3	41.8	7.2
0.5/1	-19.1	9.4	-15.9	8.7

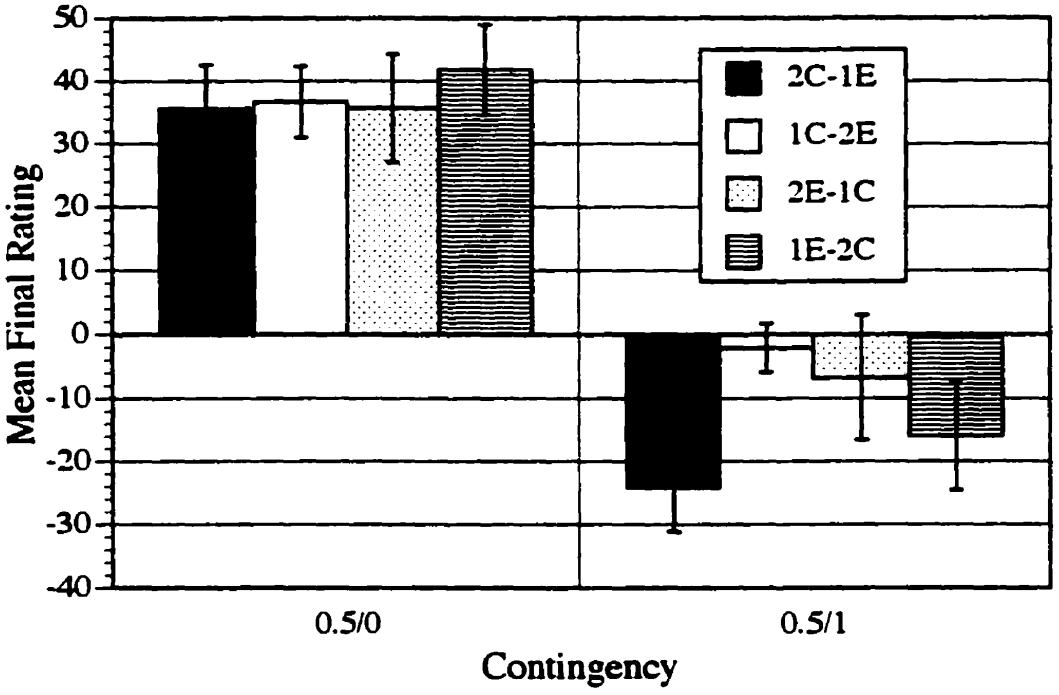


Figure 13. Final judgments and standard errors of Blocked contingency in Experiment 7.

both groups gave similar ratings when two predictors preceded a single outcome, F (1, 44) = 2.48; however, ratings were somewhat higher in 1C-2E than in 1E-2C, F (1, 44) = 3.73. Neither analysis attained significance.

The causal order by predictor by blocking interaction arose because the size of the cue competition effect was somewhat larger in 2C-1E than in 2E-1C; but the opposite occurred when one predictor preceded two outcomes. Pursuant analyses revealed that the difference in the cue competition effect was not different between 2C-1E and 2E-1C, F (1, 44) = 1.10, whereas this effect was significantly smaller in 1C-2E than in 1E-2C, F (1, 44) = 7.72. Last, the causal order by predictor by time interaction arose because, for the first set of ratings, judgments were significantly higher in 1C-2E than in 1E-2C, F (1, 44) = 9.55; but both groups gave similar judgments after 48 trials, F (1, 44) = 2.81.

Ratings of the Blocking Contingencies

As shown in Table 20, both groups rated the perfect contingency higher than the zero contingency. The size of this difference varied between the two groups, the difference being larger in EC than in CE. This discrepancy arose because EC rated the zero contingency lower than CE; furthermore, judgments of the perfect contingency were somewhat lower in 1C-2E than in the other three conditions. Last, whereas the two groups judged the contingencies differently during the first set of ratings, these differences tended to be smaller for the final ratings.

Statistical analyses support these contentions. The main effect of causal order was not significant, F (1, 44) = 3.60. Furthermore, of the within-subject factors and interaction, only the main effect of the blocking contingencies attained significance, F (1, 44) = 997.02; next largest F (1, 44) = 2.40. However, causal order interacted with the main effects of blocking and time [F (1, 44) = 9.60, F (1, 44) = 5.82, respectively]; as well, the causal order by predictor by blocking interaction attained significance, F (1, 44) = 7.66.

A Bonferroni correction was used for the six follow-up analyses. Thus, the rejection criterion was 0.008. The causal order by blocking interaction arose because EC

Table 20

Experiment 7
Mean Estimates of the Blocking Contingencies after 32 (first rating) and 48 trials (final rating) for both the CE and EC groups. The number before the slash refers to the blocked contingency and the number after the slash refers to the blocking contingency.

CE GROUP				
	First Ratings		*Final Ratings*	
Condition	**Mean**	**SEM**	**Mean**	**SEM**
2 Causes - 1 Effect (2C - 1E)				
0.5/0	- 5.8	8.4	- 6.0	7.5
0.5/1	99.0	0.9	99.6	0.4
1 Cause - 2 Effects (1C - 2E)				
0.5/0	15.2	4.3	1.9	4.9
0.5/1	84.2	6.2	81.5	6.6
EC GROUP				
2 Effects - 1 Cause (2E - 1C)				
0.5/0	-17.0	9.4	0.6	8.6
0.5/1	95.9	3.2	95.9	2.6
1 Effect - 2 Causes (1E - 2C)				
0.5/0	-25.9	6.9	-18.6	8.0
0.5/1	91.8	3.6	92.5	3.3

gave significantly lower ratings than CE to the zero contingency, F (1, 44) = 9.69, although no group differences were found for judgments of the perfect contingency, F (1, 44) < 1. This analysis, however, should be approached with caution because of judgments being close to ceiling in 2C-1E and 2E-1C, resulting in heterogeneity of variance.

The causal order by time interaction arose because group CE gave significantly higher ratings than EC for the first set of judgments, F (1, 44) = 8.53; both groups, however, gave similar judgments when they were requested after 48 trials, F (1, 44) < 1. Last, the causal order by predictor by blocking interaction arose because the size of the difference between judgments of the zero and perfect contingencies varied by both the number of predictors and whether participants reasoned from cause to effect or effect to cause. When a single predictor preceded two outcomes, the difference between the zero and perfect contingency was smaller in CE than in EC, F (1, 44) = 14.41; this is because ratings of the perfect contingency were lowest in 1C-2E. However, the size of this difference was similar in groups CE and EC when two predictors preceded a single outcome, F (1, 44) < 1. In summary, participants discriminated the perfect and zero blocking contingencies.

Questionnaire

The results from the questionnaires imply that the participants understood which variables were causes and which ones were effects. All of the participants in group CE answered the first question, which asks whether the chemicals affected the bacteria or the bacteria affected the chemicals, correctly. In EC, all participants (except the one removed from the study) answered this question correctly in 1E-2C and only one participant in 2E-1C thought that the chemicals were affecting the bacteria. In 2E-1C, the participant removed from the study wrote "whether the presence of one indicates the presence of the other" and in 1E-2C wrote "They were looking at how each may be affecting the other (no causal inferences can be drawn)." As stated, this participant felt that there was deception in the experiment and stated this during debriefing. This person's results were removed from the study because she was the only one who believed that the relationship

between the cause and two effects was correlational and not causal, which clearly violated a central tenet in causal model theory (Waldmann & Holyoak, 1997).

The majority of participants correctly stated which variables were causes and effects and most of the volunteers in CE were able to extend the analogy between cause and effect in the present experiment to a novel situation in which the cues were a pool cue and pool ball. However, some participants in EC failed to do so. In 1C-2E, all participants stated the bacteria were analogous to the pool cue, and in 2C-1E, twenty-three out of twenty-four volunteers stated the chemicals were analogous to the pool cue. However, sixteen out of twenty-two participants in 2E-1C solved this analogy correctly whereas nineteen were correct in 1E-2C. Thus, fewer participants in 2E-1C could extend the variables in this experiment to a novel situation, although the majority (73%) answered this question correctly.

Finally, few participants listed alternative causes, as shown in Table 21. Similar results for the judgments of the blocked and blocking contingencies were obtained regardless of whether these participants were included in the analyses. To avoid unnecessarily reducing the sample size, the reported analyses included these participants' data.

In CE, one participant stated, in both 2C-1E and 1C-2E, that it is impossible to control for every source of random variation in an experiment, and one person complained that the testing procedure in 2C-1E was not accurate if consistent results were not obtained on every trial. The most alternative causes were listed in 2E-1C, in which there were 3 responses given. Unlike Experiments 5 and 6, however, none of the alternative causes could plausibly fall under the category of laboratory conditions, faulty experimental design or interacting causes. One participant listed the sex of the rats whereas another was not sure if the rats' age was specified in the instructions. And one person mentioned that one or more competing strains of bacteria might have accounted for the chemicals' production or that the chemicals themselves might have interacted. In 1E-2C, two participants listed alternative causes: one was not sure if the amount of each chemical was the same when two chemicals were added to the bacterial sample as

Table 21

List of alternative causes listed by participants in Experiment 7.

2C -1E	
"Other things. It's impossible to take everything into account."	"Test was not reliable i.e. the same chemical would be added and would yield inconsistent results."
1C - 2E	
"Like I said before, it's impossible to take everything into account."	
2E - 1C		
"male or female rats?"	"Maybe. Other bacterias [sic] and factors or the combination of both chemicals together."	"The age of the rats (if it was mentioned, then no)"
1E - 2C		
"when there were 2 chemicals in the culture, did they use the same amount of <u>each</u> as they would if there were only one chemical?"	"They should have had the same number of trials for each combination of chemical"	

opposed to when only a single chemical was added, and the other complained that there should have been an equal number of trials in which one chemical, both chemicals and neither chemical was added. Upon debriefing, this person explained that an equal number of these types of trials would have made the task easier.

Discussion

To summarize, the results from Experiments 1, 2, 3 and 4 implied that the important variable influencing blocking was whether two cues preceded a single outcome or one cue preceded two outcomes. The results appeared to provide clear support for the R-W model. However, the number of antecedent cues was confounded by the fact that two sets of intervening questions were asked on each trial when one predictor was presented before two outcomes, but only one was asked when two predictors were presented before a single outcome. These questions in 1C-2E and 1E-2C encouraged participants to pay attention to the individual cues, and thus prevented blocking from being observed. However, the intervening questions were removed in the present experiment, resulting in cue competition being observed regardless of causal scenario and causal order.

The results from the questionnaires imply that this blocking effect can not be accounted for by participants not comprehending the causal model or not believing that there was a causal relationship among the cues. The cause and effect was correctly identified in almost every condition and most participants were able to extend this understanding to a novel situation involving a pool cue and ball.

Furthermore, the results in 2E-1C imply that blocking in this condition can not be plausibly explained by the argument that participants are resorting to using alternative causal models. As stated in the Introduction, this argument is circular because blocking confirms the presence of these competing causes whereas the absence of blocking confirms that they do not play a role in the experiment. In this experiment, data about alternative causes were collected independently of people's causal judgments. The number of alternative causes listed was similar across all four conditions. In addition,

only one person in 2E-1C stated that there might have been competing strains of bacteria that could have influenced the chemicals' production. This evidence is clearly not strong enough to support the contention that blocking in 2E-1C is accounted for by alternative competing causes.

These results are arguably most consistent with the modified R-W model, although they do not provide overwhelming support for any one of the theories shown in Table 17. On one hand, they are consistent with the modified R-W model because blocking was found in all four experimental conditions. For example, that blocking even occurred in 1C-2E is inconsistent with both causal model theory and the R-W model. According to causal model theory (Waldmann & Holyoak, 1992, 1997), people should readily accept that one cause is capable of producing multiple effects. Since effects do not interact, cue competition would not be expected. Similarly, the R-W model predicts the absence of blocking because each outcome should be capable of supporting independent amounts of associative strength. Although cue competition was observed, more trials were needed in 1C-2E compared to 2C-1E, 2E-1C, and 1E-2C, in which blocking was observed after 32 trials.

Blocking was also largest in 2C-1E and 1E-2C, which is inconsistent with causal model theory. According to this view, blocking should be similar in both of these conditions. The underlying assumption of causal model theory is that people reason from cause to effect, regardless of the order in which the events are presented. Since causes compete, blocking would be expected. However, it could be argued that the acquisition should be more difficult in 1E-2C because the contingencies were presented in the reverse order from which the events took place. The participant thereby has the additional task of having to retrospectively place the events in the order they would occur in nature. These additional demands should make the task more difficult to learn. But the evidence does not support this assertion. Judgments of the moderate contingency when it was paired with a weak one were similar in all conditions, and equivalent blocking was observed after 32 and 48 trials.

GENERAL DISCUSSION

This series of experiments investigated the fundamental question of how causal scenario and causal order influence blocking. According to statistical theories (Waldmann and Holyoak, 1992, 1997), blocking should be observed whenever there are multiple causes of a single effect but not when there are multiple effects of a single cause, regardless of whether causes are presented before effects or vice versa. On the other hand, reductionist or associative theories (Mackintosh, 1975; Pearce, 1987, 1994; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Wagner & Rescorla, 1972) view events as cues and outcomes, but not as causes or effects. The assumption underlying these theories is that cues compete for association with outcomes, but the reverse does not occur. Consequently, cue competition is expected only between antecedent events, regardless of whether they represent causes or effects.

Associative and normative theories have formed two separate camps. Researchers from the associative school of thought use testing strategies that produce results inconsistent with normative models. Similarly, normative theorists use methodologies that produce findings that are incompatible with an associative approach. The results from the present series of experiments imply that neither the R-W model nor causal model theory can capture the full range of findings. Instead, they are consistent with the notion that the link between cues and outcomes is bidirectional (Arcediano et al., 1997; Matute et al., 1996; Miller & Matzel, 1988), and that the intervening questions between the antecedent and consequent cues can modulate the occurrence of blocking. In particular, placing these questions between the antecedent and consequent cues has a greater influence when one antecedent cue predicts two outcomes, as opposed to two antecedent cues predicting one outcome.

In Experiments 1 and 2, two antecedent cues preceded a single consequent event. Blocking was observed between causes and effects both a) when participants had to predict the outcome on the individual trials (see also Baker et al., 1993; Shanks & Lopez, 1996; Vallée-Tourangeau et al., 1994; Price & Yates, 1993), and b) when participants passively acquired the learning task (see also Experiment 7; Price & Yates, 1993). Thus,

the intervening question between the antecedent cues and consequent cue did not play a prominent role when there were two predictors of a single outcome, regardless of whether the cues presented first were causes or effects. This finding is consistent with previous evidence of blocking when two cues precede an outcome.

On the other hand, when a single cue preceded two outcomes, the intervening question was an important mediator of whether cue competition would be observed. In Experiments 3, 4, and 5, participants were required to monitor each of the two outcomes independently because they either had to predict whether each one would occur or had to keep track of which conjunction of the two outcomes would take place. In this condition, blocking was not observed in either 1C-2E or 1E-2C (see also Baker & Mazmanian, 1989; Price & Yates, 1995; Rescorla, 1991). But when the contingencies were passively acquired (Experiment 6 and 7), blocking was observed in both conditions. This finding demonstrates that the intervening predictions play a more important moderating role in influencing blocking when there is a single predictor of two outcomes, as opposed to two predictors of one consequent event. The results from the present series of experiments imply that actively predicting the presence or absence of the consequent cues encourages participants to attend to the individual cues, which in turn prevents blocking from being observed. This discovery was unexpected because it was not anticipated by either causal model theory or the R-W model.

The findings from the present series of experiments and their implications for causal model theory and the R-W model will be discussed. But before doing so, it is instructive to ask whether the designs used in these experiments were appropriate for evaluating causal model theory. The section that follows will address the criticism that the preparations used did not meet the methodological criteria to compare the two models (Waldmann & Holyoak, 1997).

Addressing Objections: Do the Experiments Test Causal Model Theory?

The first requirement is that instructions in the experiment make it clear that there is a cause-effect relationship in the learning task. Discussion of this topic is usually based

on speculation. For example, Shanks and Lopez (1996) used a task in which symptoms were presented before diseases and argued that this cover story involved effects being seen before causes because diseases precede an illness; on the other hand, Waldmann and Holyoak (1997) have argued that the symptoms may have been interpreted as causes rather than effects (see also Price & Yates, 1995). However, without an independent measure of how these tasks are interpreted, it is not productive to argue about how the causal model is deciphered.

In Experiments 1, 2, 3, and 4, it was argued that there was an implied causal mechanism between the cues because the chemicals either influenced the bacteria's survival or because they were metabolic by-products of the bacteria. Direct evidence for this assertion was obtained in Experiments 5, 6 and 7. The volunteers filled out a questionnaire in which they stated the causal relationship among the cues in the experiment. In Experiments 5 and 6 (1E-2C), participants consistently stated that the two chemicals were acting on the bacteria even though information about the presence or absence of the bacteria was presented before the chemicals. And in Experiment 7, the volunteers correctly stated the causal direction in each of the four conditions (2C-1E, 2E-1C, 1C-2E, 1E-2C). Furthermore, these participants were able to extend their comprehension of the causal model to a novel situation involving a pool cue and ball. These experiments clearly satisfy the requirement that the tasks were interpreted in a cause-effect direction.

The second criterion is that the acquired causal knowledge be measured appropriately. This requirement arose because Shanks and Lopez (1996; Shanks, 1991) asked participants how strongly each cue was associated with an outcome and equated this measure with predictiveness. In objection, Waldmann and Holyoak (1997) argued that participants should be asked meaningful test questions rather than questions about the mere association among arbitrary events. These test questions would include rating how sure the participant is that a factor is a cause or effect, or rating the degree to which a cue is predictive. In the experiments described here, when two causes preceded one effect and when one effect preceded two causes, participants were asked to rate how strongly

each chemical affected the bacteria's survival or how strongly the bacteria affected each chemical's production. Furthermore, when two effects preceded one cause and when one effect preceded two causes, the volunteers were asked questions in both the ?EC (diagnostic) and ?CE (predictive) directions. In other words, when causes preceded effects participants were asked cause-to-effect (?CE) test questions, and when effects preceded causes the volunteers were asked both predictive and diagnostic test questions. It thus seems reasonable to assume that these experiments measured the causal knowledge that participants acquired.

The third requirement consists of two parts, namely a) that causal order be manipulated, and b) that participants presented with effects before causes (i.e., diagnostic conditions) do not believe that there are multiple alternative causes for the effects. The first criterion was clearly met in the described experiments. Previous experiments have confounded causal scenario with causal order. For example, they have compared only multiple causes of a single effect to multiple effects of a single cause (Matute et al, 1996; Price & Yates, 1993, 1995; Shanks & Lopez, 1996; Van Hamme et al., 1993; Waldmann & Holyoak, 1992). And in one experiment, 2C-1E was compared to 1E-2C (Price & Yates, 1995). But none of the experiments investigating these two factors used a 1C-2E preparation. In addition, different cues were sometime used in the diagnostic and predictive scenarios. For example, foods were causes of an allergy when multiple causes preceded an effect; but symptoms were presented before diseases or allergic reactions were presented before medications when multiple effects preceded a cause (e.g., Matute et al., 1996;Van Hamme et al., 1993). In the experiments described here, causal scenario was crossed with causal order. These factors were investigated in separate experiments (Experiments 1 to 6), as well as within a single experiment (Experiment 7). Furthermore, the same cues were used in all experiments: namely, all tasks used two chemicals and a bacterial strain. The differences between the predictive and diagnostic scenarios were thus minimized in terms of the cues and how they interacted with each other.

The second criterion regarding controlling for alternative causes was also achieved. In Experiments 5 and 6, participants listed causes apart from the chemicals that

could influence the bacteria's survival. In each condition of Experiment 7, the volunteers read that these causes were controlled for before they started the learning phase. At the end of the experiment, they were asked to list alternative causes. Unlike Experiment 5 and 6, few participants in any of the conditions believed that there were alternative causes. Furthermore, whereas those in Experiments 5 and 6 gave a list of causes, those in Experiment 7 wrote only one or two statements about other causes. Thus it is unlikely that the blocking observed when effects preceded causes can be attributed to the volunteers believing that there were multiple alternative causes.

Last, a fourth criterion for testing causal model theory is that the statistical relation between the cause and effect be controlled. This requirement was discussed because, in Shanks and Lopez (1996), the contingencies among the cues were different when multiple causes preceded a single effect as opposed to when multiple effects preceded a single cause. In the experiments described here, however, the frequency of events was the same both when two antecedent events preceded a single consequent and when one antecedent cue preceded two consequent events. As shown in Tables 6 and 9, the frequency of events was the same in Experiments 1 to 6 and in the 0.5/0 and 0.5/1 conditions of Experiment 7. The contingency between each chemical and the bacterial sample was thus controlled in all experiments.

Furthermore, steps were taken to help ensure that the participants were sensitive to contingency. Based on prior research, the instructions in each experiment explained the notion of contingency (Crocker, 1981) and mentioned that each cause could have a positive, negative, or no contingency with the effect (Peterson, 1980; Wasserman, 1990b). Whereas the critical comparison in Waldmann and Holyoak (1992; see also Waldmann, 2000) involved testing blocking of one perfectly positive contingency by another, these experiments investigated cue competition of moderate (see also Matute et al., 1996; Price and Yates, 1995) and zero contingencies by a strong positive blocking contingency. This issue is important because few have used a range of contingencies (Van Hamme et al., 1993). Thus, these experiments are a first step at investigating the role of contingency and causal models in blocking.

The present series of experiments also provide direction for future research. To further study the role of contingency, it would be valuable to test whether the present results can be extended to negative contingencies. In addition, follow-up studies should investigate blocking by a strong contingency of opposite polarity, that is, blocking of a moderately positive contingency by a strong negative one and vice versa. These experiments would provide a more complete description of the interaction between contingency and causal models in blocking.

Implications for Causal Model Theory & the R-W Model

The findings from the experiments in this project have important implications for both causal model theory and the R-W model. An important assumption in causal model theory is that causes compete to be linked with effects, but the reverse does not occur. Table 22 displays the general results obtained in the experiments described here and in the literature on causal scenario and order. As can be seen, cue competition appears to be dependent not only on these two variables but also on the processing involved in the gap between antecedent and consequent cues.

People's judgments tend to conform to the predictions of causal model theory (column 3) when all of the cues are presented simultaneously or when a two-stage design is used. For example, Van Hamme et al. (1993) found blocking in 3C-1E but not 3E-1C when the contingencies were presented on a list. This finding was replicated by Matute et al. (1996) when they asked questions about causality. Similarly, Waldmann and Holyoak (1992) obtained these results using a two-stage design in which participants first learned about a perfect predictor of the trial's outcome before being presented with a redundant predictor (see Table 1). Waldmann (2000) has recently published some experiments consistent with causal model theory, which are described in an endnote in order to avoid interrupting the present discourse.[1] A key difference between Waldmann's (2000) experiments and most other experiments testing causal model theory (columns 1 and 2) is that Waldmann (2000) uses a two-stage design whereas others have presented the blocked and blocking cues in either a single stage of a computer-based game or simultaneously

Table 22

Summary of findings from this thesis and prior experiments.

2 Antecedent Cues			
	One Prediction Between Cue and Outcome	*Passive Observing*	*Simultaneous Presentation of Cues or 2-Stage Designs*
2C-1E	Blocking [Experiment 1; Baker et al., 1993; Price & Yates, 1993, 1995; Shanks & Lopez, 1996; Vallée-Tourangeau et al., 1994]	Blocking [Experiment 7; Price & Yates, 1993]	Blocking [Waldmann & Holyoak, 1992, see also Matute et al., 1996; Van Hamme et al., 1993; Waldmann, 2000]
2E-1C	Blocking [Experiment 2; Shanks & Lopez, 1996]	Blocking [Experiment 7]	No Blocking [Waldmann & Holyoak, 1992; see also Matute et al., 1996; Van Hamme et al., 1993; Waldmann, 2000]
1 Antecedent Cue			
	Monitor Two Consequent Cues Independently	*Passive Observing*	*Simultaneous Presentation of Cues or 2-Stage Designs*
1C-2E	No Blocking [Experiment 3; Baker & Mazmanian, 1989]	Blocking [Experiment 7]	Not Tested
1E-2C	No Blocking [Experiments 4 & 5; Price & Yates, 1995]	Blocking [Experiments 6 & 7]	Not Tested
Results Consistent with:	R-W model	Modified R-W model	Causal Model Theory

(i.e., on a list). The results thus tend to be consistent with causal model theory in two-stage designs or when there is minimal information processing. In fact, Waldmann (2000) has argued that "causal-model theory is presently restricted to model people's competence to learn about causal structures. This competence displays itself best when potential information processing constraints are reduced (p. 72)." This statement raises the question of how generalizable this model is for real-world causal acquisition if it makes successful predictions only for extremely simple tasks.

Other experiments investigating effects of causal scenario and order have used single-stage designs. When there are two antecedent cues, the results appear to be consistent with the R-W model when the contingencies are acquired over trials. That is, blocking is observed regardless of whether the antecedent cues are causes or effects. This result has furthermore been obtained even when all of the cues were shapes and regardless of whether there were intervening questions (Vallée-Tourangeau et al., 1994).

The interpretation of the factors influencing blocking is more complex, however, when one predictor precedes two events. As described earlier and shown in the table, blocking is not observed in this situation if participants are required to monitor each of the two consequent cues separately. Price & Yates (1995) directly compared 2C-1E to 1E-2C and found blocking in 2C-1E but not in the latter condition. They argued that this result was consistent with the R-W model because causal model theory predicts blocking in both conditions. According to causal model theory, blocking should be observed in 1E-2C but not in 1C-2E. This result would most likely be obtained if the contingencies are presented simultaneously or in a two-stage design. At this time, this preparation has not been used for comparing these two conditions and thus would be a promising venue for future research. However, participants in Experiments 7 acquired the contingencies with a passive observing procedure and blocking was obtained in both conditions.

The evidence, when considered as a whole, suggests that the intervening questions impose a unidirectional form of learning, that is, from cue to outcome. Accordingly, when there is only one antecedent cue, having participants predict the presence or absence of the two consequent cues requires that the volunteers monitor each of the two

contingencies independently. This demand, however, is absent when two antecedent cues precede one consequent cue. Hence, blocking is observed when two antecedent cues precede one consequent cue but not when one cue predicts two consequent cues, regardless of whether the cues are defined as causes or effects. When participants are not required to predict the outcome or outcomes on each trial, then contingency acquisition is bidirectional. In this instance, competition is observed between contingencies regardless of whether the are defined as causes or effects, and regardless of whether causes precede effects or effects precede causes. This analysis may account for why cue competition is observed in some but not other circumstances.

The final issue to be briefly discussed is the role of the test question, which has also been proposed as an important variable influencing blocking (Matute et al., 1996; Waldmann, 2000). To limit the scope of this thesis, the experiments did not investigate this issue in detail. This issue was given less priority because the current arguments about its role have been vague and circular. For example, Waldmann and Holyoak (1992) attribute the opposing results in their 2E-1C conditions of Experiments 1 and 2 to the fact that a ?CE (predictive) test question was asked in the first experiment whereas an ?EC (diagnostic) test question was asked in the second (Table 1). But no independent evidence was provided to demonstrate that the participants' interpretation of the test question was the critical variable. The same type of reasoning has been used by Matute et al. (1996) when they argue that causality and contiguity (or conditional probability) test questions are interpreted differently. That is, the meaning assigned to the test question is speculative because no evidence was provided to show that the test questions were represented as intended. Furthermore, other researchers have found that these two types of test questions yield similar results (Price & Yates, 1993, 1995). Further research is thus required to resolve the relationship between cue competition and the test question.

Do Associative Models Have A Place in Causal Learning?

The question that thus arises is whether associative or statistical models are best able to account for the acquisition of causal knowledge. At one extreme, those in the

statistical camp have written that associative theories should be abandoned. For example, Cheng (1997; see also Cheng & Holyoak, 1995) has argued that the R-W model is simplistic. Similarly, Waldmann and Holyoak (1992) conclude their paper with the assertion that "lower-order associative learning should be reduced to high order causal induction" (p. 235).

On the other hand, others claim that associative models are needed in order to account for the full range of empirical findings in causal acquisition (e.g., Baker et al., 1996; Price & Yates, 1995; Siegel & Allan, 1996; Young, 1995). One advantage offered by the R-W model is that it provides a simple computational algorithm by which cues and outcomes are linked. Hence, it has heuristic value because it is sensitive to causal relations even though these contingencies are not explicitly calculated. A second advantage is that it includes few assumptions. It only assumes a monotonic relationship between the cue and outcome, whereas statistical theories (e.g., Cheng & Holyoak, 1995; Cheng, 1997; Waldmann, 2000) rely on two or more theoretical constructs: e.g., selecting an appropriate focal set, taking causal scenario and order into account, and calculating each conditional contingency. Consequently, empirical phenomena such as Kamin's (1969) blocking experiment are explained in more complex terms such as an undefined Δp.

The results from the experiments in this project imply that both schools of thought have a place in causal learning. Statistical theories are best at explaining cue competition effects when the learning task has minimal information processing demands. This includes learning about the cues in stages or being presented with the causes and effects simultaneously. In these conditions, people are sensitive to causal models in a way not predicted by associative theories. On the other hand, associative theories are best able to account for the occurrence of cue competition when intervening questions are asked between the cues and outcomes. In between these two types of information processing, the results are consistent with the view that the link between cues and outcomes is bidirectional rather than unidirectional (Baker et al., 1996; Shanks et al., 1996). Taken together, the results suggest that learning about cause-effect relations can occur in both

the cause-effect and effect-cause direction, and that cue competition is mediated not by the test questions asked at the time of judgment (Miller & Matzel, 1988) but by the processing involved during acquisition. This view thus takes an intermediate position between a reductionist and statistical approach. Both types of models have utility in explaining causal learning, and they complement rather than contradict each other.

STATEMENT OF ORIGINAL CONTRIBUTION

The seven experiments reported in this thesis provide original contributions to the advancement of knowledge in the field of contingency judgments. The experimental preparation of Baker et al. (1993) was adapted to test the effects of causal scenario and causal order on blocking. Unlike previous research investigating this question, causal scenario was not confounded with causal order. Furthermore, the contingency between the causes and effects, as well as the cues were used were the same in all experiments.

Experiments 1 and 2 replicated prior findings of blocking between antecedent cues, regardless of whether they were defined as causes or effects. In Experiments 3 and 4, blocking was not observed when one cause preceded two effects or when one effect preceded two causes. These results are consistent with the Rescorla-Wagner model. Experiment 5 ruled out the possibility that participants had not appropriately encoded the cues as an effect and two causes, respectively. This goal was achieved by asking participants to state which variable was the cause and which were the effects. Thus, direct evidence was obtained for the argument that participants comprehended the causal model. This is in contrast to prior research in which discussion on this issue is based on speculation.

In Experiments 3 to 5, participants had to implicitly monitor the two consequent cues separately when one predictor preceded two outcomes. This constraint was removed in Experiment 6 by allowing participants to observe the contingencies but without requiring them to predict the outcomes on each trial. Unlike Experiments 4 and 5, blocking was observed when one effect preceded two causes. This experiment thus demonstrated that the intervening question between antecedent and consequent cues plays

an important role in modulating blocking. Although the experimental finding was not anticipated by the Rescorla-Wagner model, the result would be expected if the model is modified so that the link between cues and outcomes is bidirectional and one assumes that the intervening questions impose unidirectional processing from the predictor to the outcome. A corollary that arises from this view is that contingencies should compete if this constraint is removed, irrespective of causal scenario and causal order.

Experiment 7 tested this prediction by investigating causal scenario and causal order in the same experiment. To avoid confounding these variables with the intervening questions between antecedent and consequent cues, all participants acquired the contingencies by passively observing them on the computer screen. Furthermore, based on a list of causes listed by participants in Experiments 5 and 6, the volunteers read instructions designed to minimize the role of alternative causal models. Blocking was observed irrespective of causal scenario and causal order, which was consistent with the modified Rescorla-Wagner model.

CONCLUSION

The seven experiments were designed to test the extent to which associative and statistical models can account for how people learn about contingencies. Causal scenario and causal order were combined and investigated both in separate experiments and within the same experiment, and their influence on blocking was examined. It was concluded that the link between cues and outcomes is bidirectional and that the type of processing in the gap between antecedent and consequent cues influences blocking. Thus, both models can account for empirical findings under different experimental conditions. The results are consistent with the statistical approach when the cues are presented simultaneously or when a two-stage design is used during acquisition. On the other hand, the Rescorla-Wagner model can account for blocking if participants are asked to predict the trial's outcome. In between these two conditions, competition between contingencies is observed. The results from these experiments suggest that both normative and statistical models are needed to provide a complete account of causal acquisition.

REFERENCES

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

Allan, L.G. (1993). Human contingency judgments: Rule based or associative? . *Psychological Bulletin, 114*, 435-448.

Allan, L.G., & Jenkins, H.M. (1980). The judgment of contingency and the nature of response alternatives. *Canadian Journal of Psychology, 34*, 1-11.

Allan, L.G. & Jenkins, H.M. (1983). The effect of representations of binary variables on judgment of influence. *Learning and Motivation, 14*, 381-405.

Alloy, L.B., & Abramson, L.Y. (1979). Judgment of contingency in depressed and nondepressed college students: Sadder but wiser?. *Journal of Experimental Psycholgy: General*, 108: 441-485.

Alloy, L.B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review, 91*, 112-149

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

Arkes, H.R., & Harkness, A.R. (1983). Estimates of contingency between two dichotomous variables. *Journal of Experimental Psychology: General, 112*, 117-135.

Baker, A.G., Berbrier, M.W., & Vallée-Tourangeau, F. (1989). Judgments of a 2 x 2 contingency table: Sequential processing and the learning curve. *Quarterly Journal of Experimental Psychology, 41B*, 65-97.

Baker, A.G. & Mazmanian, D. (1989). Selective associations in causality judgments II: A strong causal relationship may facilitate judgments of a weaker one. In *Proceedings of he Eleventh Annual Conference of the Cognitive Science Society* (pp. 538-545). Hillsdale: Lawrence Erlbaum Associates.

Baker, A.G., Mercier, P., Vallée-Tourangeau, F., Frank, R, & Pan, M. (1993). Selective associations and causality judgments: The presence of a strong causal factor may reduce judgments of a weaker one. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 414-432.

Baker, A.G., Murphy, R.A. & Vallée-Tourangeau, F.(1996). Associative and normative models of causal induction: Reacting to versus understanding cause. In D.R. Shanks, K.J. Holyoak, & D.L. Medin (Eds). *The Psychology of Learning and Motivation* (pp. 1-45). San Diego: Academic Press.

Baron, J. (1994). *Thinking and deciding*. Second Edition. Cambridge: Cambridge University Press.

Benedict, J.O., & Ayres, J.J. (1972). Factors affecting conditioning in the truly random control procedure in the rat. *Journal of Comparative & Physiological Psychology*, 78, 323-330.

Beyth-Marom, R.(1982). Perception of correlation reexamined. *Memory and Cognition, 10*, 511-519.

Bullock, M., & Gelman, R. (1979). Preschool children's assumptions about cause and effect: temporal ordering. *Child Development, 50*, 89-96.

Busemeyer, J.R., McDaniel, M.A. & Byun, E. (1996). The use of intervening variables in causal learning. In D.R. Shanks, K.J. Holyoak, & D.L. Medin (Eds). *The Psychology of Learning and Motivation* (pp.357-391). San Diego: Academic Press.

Catena, A., Maldonado, A., & Càndido, A. (1998). The effect of the frequency of judgment and the type of trials on covariation learning. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 481-495.

Chapman G.B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 837-854.

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

Chapman L.J., & Chapman, J.P. (1967). Genesis of popular but erroneous psychodiagnositic observations. *Journal of Abnormal Psychology, 72*, 193-204.

Chapman L.J., & Chapman, J.P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology, 74*, 271-280.

Chatlosh, D.L., Neunaber, D.L., & Wasserman, E.A . (1985). Response-outcome contingency: Behavioural and judgmental effects of appetitive and aversive outcomes with college students. *Learning and Motivation, 16*, 1-34.

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

Cheng, P.W. & Holyoak, K.J. (1995). Complex adaptive systems as intuitive statisticians: causality, contingency, and prediction. In H.L. Roitblat & J.-A. Meyer (Eds.), *Comparative Approaches to Cognitive Science*. (pp. 271-302). Cambridge: MIT Press.

Cheng, P.W. & Novick, L.R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology, 58*, 545-567.

Cheng, P.W & Novick, L.R. (1992). Covariation in natural causal induction. *Psychological Review, 99*, 365-382.

Cheng, P.W., Park, J., Yarlas, A.S. & Holyoak, K.J. (1996). A causal power theory of focal sets. In D.R. Shanks, K.J. Holyoak, & D.L. Medin (Eds). *The Psychology of Learning and Motivation* (pp.313-355). San Diego: Academic Press.

Colwill, R.M, & Rescorla, R.A. (1985). Postconditioning devaluation of a reinforcer affects instrumental responding. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 120-132.

Crocker J. (1981). Judgment of covariation by social perceivers. *Psychological Bulletin*, 90, 272-292.

Crocker, J. (1982). Biased questions in judgment of covariation studies. *Personality and Social Psychology Bulletin, 8*, 214-220.

Dickinson, A., & Charnock, D.J. (1985). Contingency effects with maintained instrumental reinforcement. *Quarterly Journal of Experimental Psychology, 37B*, 397-416.

Dickinson, A., & Shanks, D.R. (1985). Animal learning and human causality judgment. In L. -G. Nilson & T. Archer (Eds.), *Perspectives on Learning and Memory*. (pp. 167-191). Hillsdale: Lawrence Erlbaum Associates.

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

Durlach, P.J. (1983). Effects of signalling intertrial unconditioned stimuli in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes, 9*, 374-389.

Eddy, D. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, and A. Tversky (Eds.). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press. Cambridge.

Einhorn, H.J. & Hogarth, R.M. (1986). Judging probable cause. *Psychological Bulletin, 99*, 3-19.

Esmoris-Arranz, F.J., Miller, R.R., & Matute, H. (1997). Blocking of subsequent and antecedent events. *Journal of Experimental Psychology: Animal Behavior Processes, 23*, 145-156.

Fales, E., & Wasserman, E.A. (1992). Causal knowledge: What can psychology teach philosophers?. *Journal of Mind and Behavior, 13*, 1-28.

Fischoff, B. & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review, 90*, 239-260.

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

Golding, S.L., & Rorer, L.G. (1972). Illusory correlation and subjective judgment. *Journal of Abnormal Psychology, 80*, 249-260.

Gruber, H.E., Fink, C.D., & Damm, V. (1957). Effects of experience on perception of causality. *Journal of Experimental Psychology*, 53, 89-93.

Hammond, L.J. (1980). The effect of contingency upon the appetitive conditioning of free-operant behavior. *Journal of the Experimental Analysis of Behavior, 34*, 297-304.

Harris, P.L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition, 61*, 233-259.

Hart, H.L.A, & Honoré, A.M. (1959). *Causation in the law*. Oxford: Clarendon Press.

Hilton, D.J., & Slugoski, B.R. (1986). Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review, 93*, 75 - 88.

Hume, D. (1946/1739). *A Treatise of Human Nature*. Oxford: Clarendon Press.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.

Jenkins, H.M., & Ward, W.C. (1965). The judgment of contingency between responses and outcomes. Psychological Monographs, 79, No.1 (Whole No. 594).

Jennings, D.L., Amabile, T., & Ross, L. (1982). Informal covariation assessment: Data-based versus theory-based judgments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 211-230). Cambridge: Cambridge University Press.

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

Kant, I. (1965/1739). *Critique of Pure Reason.* New York: Macmillan.

Kao, S. -F., & Wasserman, E.A. (1993). Assessment of an information integration account of contingency with examination of subjective cell importance and method of information presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1363-1386.

Kelley, H.H. (1973). The process of causal attribution. *American Psychologist, 28,* 107-128.

Klayman, J. & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94,* 211-228.

Klayman, J. & Ha , J. -W. (1989). Hypothesis testing in rule discovery: Strategy, structure and content. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 596-604.

Kremer, E.F. (1971). Truly random and traditional control procedures in CER conditioning in the rat. *Journal of Comparative and Physiological Psychology, 76,* 441-448.

Kremer, E.F. (1974). The truly random control procedure: Conditioning to the static cues. *Journal of Comparative and Physiological Psychology, 86,* 700-707.

Kremer, E.F., & Kamin, L.J. (1971). The truly random control procedure: Associative or nonassociative effects in rats. *Journal of Comparative and Physiological Psychology, 74,* 203-210.

Leslie A.M. (1982). The perception of causality in infants. *Perception, 11,* 172-186.

Levin, I..P., Wasserman, E.A., & Kao, S.-F. (1993). Multiple methods for examining biased information in contingency judgments. *Organizational Behavior and Human Decision Processes, 55,* 228-250.

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,* 672-694.

Mackie, J.L. (1974). *The cement of the universe: A study of causation.* London: Oxford University Press.

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

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

McClelland, J.L., & Rumelhart, D.E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology, 114,* 159-188.

Michotte, A. (1963). *The perception of causality.* London: Methuen.

Mill, J.S. (1949/1843). *A system of logic.* London: Longmans, Green and Co.

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

Nisbett, R., and Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs: Prentice-Hall.

Pearce, J.M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review, 94,* 61-73.

Pearce, J.M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review, 101,* 587-607.

Pearce, J.M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not unconditioned stimuli. *Psychological Review, 87*, 532-552.

Peterson, C. (1980). Recognition of noncontingency. *Journal of Personality and Social Psychlogy*, 38, 727-734.

Peterson, C.R., & Beach, L.R. (1967). Man as an intuitive statistician. *Psychological Bulletin*. 68: 29-46.

Plous, S. (1993). *The psychology of judgment and decision making*. New York: Mcgraw-Hill.

Popper, K.R. (1965). *Conjectures and Refutation: The Growth of Scientific Knowledge*. New York: Harper Torchbooks.

Price, P.C., & Yates, J.F. (1993). Judgmental overshadowing: Further evidence of cue-interaction in contingency judgment. *Memory & Cognition, 21*, 561-572.

Price, P.C., & Yates, J.F. (1995). Associative and rule-based accounts of cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1638-1655.

Quinsey, V.J. (1971). Conditioned suppression with no CS-US contingency in the rat. *Canadian Journal of Psychology, 25*, 1-5.

Reed, P. (1992). Effect of a signalled delay between an action and an outcome on human judgment of causality. *Quarterly Journal of Experimental Psychology, 44B*, 81-100.

Reed, P. (1999). Role of a stimulus filling an action-outcome delay in human judgements of causal effectiveness. *Journal of Experimental Psychology: Animal Behavior Processes, 25*, 92-102.

Reichenbach, H. (1956). *The direction of time*. Berkeley: University of California Press.

Rescorla, R.A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology, 66*, 1-5.

Rescorla, R.A. (1969). Conditioned inhibition of fear resulting from negative CS-US contingencies. *Journal of Comparative and Physiological Psychology, 67*, 504-509.

Rescorla, R.A. (1984). Signalling intertrial shocks attenuates their negative effect on conditioned suppression. *Bulletin of the Psychonomic Society, 22*, 225-228.

Rescorla, R.A. (1991). Associations of multiple outcomes with an instrumental response. *Journal of Experimental Psychology: Animal Behavior Processes, 17*, 465-474.

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 theory and research* (pp. 64-99). New York: Appleton-Century-Crofts.

Richards, R.W. (1981). A comparison of signalled and unsignalled delay of reinforcement. *Journal of the Experimental Analysis of Behavior, 35*, 145-152.

Schlottmann, A., & Shanks, D.R. (1992). Evidence for a distinction between judged and perceived causality. *Quarterly Journal of Experimental Psychology, 44A*, 321-342.

Schustack, M.W., & Sternberg, R.J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General, 110*, 101-120.

Shaklee, H. (1983). Human covariation judgment: Accuracy and strategy. *Learning and Motivation, 14*, 433-448.

Shaklee, H & Mims, M. (1982). Sources of error in judging event covariations: effects of memory demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 208-224.

Shaklee, H., & Tucker, D. (1980). A rule analysis of judgments of covariation between events. *Memory and Cognition, 8*, 459-467.

Shaklee, H., & Wasserman, E.A. (1986). Judging interevent contingencies: Being right for the wrong reasons. *Bulletin of the Psychonomic Society, 24*, 91-94.

Shanks, D.R. (1986). Selective attribution and the judgment of causality. *Learning and Motivation, 17*, 311-334.

Shanks, D.R. (1989). Selection processes in causality judgment. *Memory and Cognition, 17*, 27-34.

Shanks, D.R. (1990). Connectionism and the learning of probabilistic concepts. *Quarterly Journal of Experimental Psychology, 41B*, 139-159.

Shanks, D.R. (1991). Categorization by a connectionist network. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 433 - 443.

Shanks D.R. & Dickinson A. (1987). Associative accounts of causality judgment. In G.H. Bower (Ed.), *The psychology of learning and motivation: Vol. 21* (pp. 229-261). San Diego, CA: Academic Press.

Shanks, D.R., & Dickinson, A. (1991). Instrumental judgment and performance under variations in action-outcome contingency and contiguity. *Memory and Cognition, 19,*353-360.

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

Shanks, D.R., Lopez, F.J., Darby, R.J., & Dickinson, A. (1996). Distinguishing associative and probabilistic contrast theories of human contingency judgment. In D.R. Shanks, K.J. Holyoak, & D.L. Medin (Eds). *The Psychology of Learning and Motivation* (pp. 265-311). San Diego: Academic Press.

Shanks, D.R., Pearson, S.M., & Dickinson, A. (1989). Temporal contiguity and the judgment of causality. *Quarterly Journal of Experimental Psychology, 41B,* 139-159.

Shultz, T.R. (1982). Rules for causal attribution. *Monographs of the Society for Research in Child Development, 47* (1, Serial No. 194).

Siegel, S., & Allan, L.G. (1996). The widespread influence of the Rescorla-Wagner model. *Psychonomic Bulletin and Review, 3,*314-321.

Siegler, R.S., & Liebert, R.M. (1974). Effects of contiguity, regularity, and age on children's causal inferences. *Developmental Psychology, 10,* 574-579.

Smedslund, J. (1963). The concept of correlation in adults. *Scandinavian Journal of Psychology.* 4, 165-173.

Spellman, B.A. (1996). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science, 7,* 337-342.

Trolier, T.K., & Hamilton, D.L. (1986). Variables influencing judgments of correlational relations. *Journal of Personality and Social Psychology,* 50, 879-888.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science,* 185, 1124-1131.

Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed)., *Progress in Social Psychology*. Hillsdale, NJ: Erlbaum.

Vallée-Tournageau, F., Baker, A.G., & Mercier, P. (1994). Discounting in causality and covariation judgments. *Quarterly Journal of Experimental Psychology, 47B*, 151-171.

Vallée-Tournageau, F., Hollingsworth, L., & Murphy, R.A. (1998). 'Attentional bias' in correlation judgments? Smedslund (1963) revisited. *Scandinavian Journal of Psychology, 39*, 221-233.

Vallée-Tournageau, F., Murphy, R.A., & Baker, A.G. (1998). Causal induction in the presence of a perfect negative cue: Contrasting predictions from associative and statistical models. *Quarterly Journal of Experimental Psychology, 51B*, 173-191.

Vallée-Tournageau, F., Murphy, R.A., Drew, S., & Baker, A.G. (1998). Judging the importance of constant and variable candidate causes: A test of the Power PC theory. *Quarterly Journal of Experimental Psychology, 51A*, 65-84.

Van Hamme, L.J., Kao, S. -F., & Wasserman, E.A. (1993). Judging interevent relations: From cause to effect and from effect to cause. *Memory and Cognition, 21*, 802-808.

Van Hamme, L.J., & Wasserman, E.A. (1993). Cue competition in causality judgments: The role of manner of information presentation. *Bulletin of the Psychonomic Society, 31*, 457-460.

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*, 127-151.

Van Overwalle, F. (1996). The relationship between the Rescorla-Wagner associative model and the probabilistic joint model of causality. *Psychologica Belgica, 36*, 171-192.

Van Overwalle, F. (1998). Causal explanation as constraint satisfaction: A critique and a feedforward connectionist alternative. *Journal of Personality and Social Psychology, 74*, 312-328.

von Wright, G.H. (1974). *Causality and determinism*. New York: Columbia University Press.

Wagner, A.R., Logan, F.A., Haberlandt, K., & Price T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology, 76,* 171-180.

Wagner, A.R., & Rescorla, R.A. (1972). Inhibition in Pavlovian conditioning: Application of a theory. In M.S. Halliday & R.A. Boakes (Eds.), *Inhibition and learning* (pp. 301-336). London: Academic Press.

Waldmann, M.R. (1996). Knowledge-based causal induction. In D.R. Shanks, K.J. Holyoak & D.L. Medin (Eds.), *The Psychology of Learning and Motivation: Vol. 34. Causal learning* (pp. 47-88). San Diego, CA: Academic Press.

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

Waldmann, M.R., & Holyoak, K.J. (1997). Determining whether causal order affects cue selection in human contingency learning: Comments on Shanks and Lopez (1996). *Memory and Cognition, 25,* 125-134.

Ward, W.C., & Jenkins, H.M. (1965). The display of information and the judgment of contingency. *Canadian Journal of Psychology, 19,* 231 - 241.

Wasserman, E.A. (1974). Stimulus-reinforcer predictiveness and selective discrimination learning in pigeons. *Journal of Experimental Psychology, 103,* 284-297.

Wasserman E.A. (1990a). Attribution of causality to common and distinctive elements of compound stimuli. *Psychological Science, 1,* 298-302.

Wasserman, E.A. (1990b). Detecting response-outcome relations: Toward an understanding of the causal texture of the environment. In G.H. Bower (Ed.), *The psychology of learning and motivation: Vol. 26* (pp. 27-82). San Diego, CA: Academic Press.

Wasserman, E.A., Chatlosh, D.L., & Neunaber, D.J. (1983). Perception of causal relations in humans: Factors affecting judgments of response-outcome contingencies under free-operant procedures. *Learning and Motivation, 14,* 406-432.

Wasserman, E.A., Dorner, W.W., & Kao, S. -F. (1990). Contributions of specific cell information to judgments of interevent contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 509-521.

Wasserman, E.A., Elek, S.M., Chatlosh, D.L., & Baker, A.G. (1993). Rating causal relations: Role of probability in judgments of response-outcome contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 174-188.

Wasserman, E.A., & Neunaber, D.J. (1986). College students' responding to and rating of contingency: The role of temporal contiguity. *Journal of the Experimental Analysis of Behavior, 46,*15-35.

Wasserman, E.A., & Shaklee, H. (1984). Judging response-outcome relations: The role of response-outcome contingency, outcome probability, and method of information presentation. *Memory and Cognition, 12,* 270-286.

White, P.A. (1998). The dissipation effect: A general tendency in causal judgments about complex physical systems. *American Journal of Psychology, 111,* 179-410.

Williams, B.A. (1976). The effects of unsignalled delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 26,* 441-449.

Williams, B.A., & Heyneman, D. (1982). Multiple determinants of "blocking" effects on operant behavior. *Animal Learning and Behavior, 10,* 72-76.

Wu, M. & Cheng, P.W. (1999). Why causation need not follow from statistical association: Boundary conditions for the evaluation of generative and preventative causal powers. *Psychological Science, 10,* 92-97.

Young, M.E. (1995). On the origin of personal causal theories. *Psychonomic Bulletin and Review, 2,* 83-104.

APPENDIX

I Instructions for Experiment 1 (2 C - 1E)

Screen 1

Imagine that scientists have recently discovered 4 strains of bacteria that live in the human digestive system and aid in the digestive process. The scientists are studying whether certain pairs of chemicals affect the survival of these bacteria. For each strain, the scientists test whether the chemicals aid in, interfere with, or have no effect on a strain's survival.

To do this, a strain of bacteria was first placed in culture (petri dishes). After that,

1) one chemical (e.g., chemical A)
2) the other chemical (e.g., chemical B)
3) both chemicals (e.g., chemicals A and B)
   OR
4) neither chemical

was added to the culture. A few hours later, the scientists verified whether or not the culture survived.

Screen 2

At the time of the experiment, the scientists did not know what effect each chemical might have on a culture. On one hand, either or both chemicals might make a culture MORE likely to survive, whereas the culture would be less likely to survive without the chemical (or chemicals). This would be an example of either or both chemicals having a POSITIVE effect on a culture's survival.

Alternatively, either or both chemicals might make it LESS likely that a culture will survive, whereas that culture would be more likely to survive without the chemical (or chemicals). This would be an example of either or both chemicals having a NEGATIVE effect on the culture's survival.

Finally, either or both chemicals might have NO systematic effect on a culture's survival. That is, it could be that the chemical (or chemicals) neither systematically aids in nor systematically interferes with a culture's survival.

Screen 3

To assess these possibilities, the scientists investigated what happened when one or two chemicals were added to the culture. They also tested what happened on control trials in which no chemicals were added to the culture. A comparison of what happened on these trials allowed the scientists to assess whether either or both chemicals had a POSITIVE effect on the culture's survival, a NEGATIVE effect on the culture's survival, or NO effect on the culture's survival.

Screen 4

You will be presented with the results from this study. On each trial, you will be told whether one chemical (e.g, chemical A), the other chemical (e.g., chemical B), both chemicals, or no chemicals were added to the culture. You will then decide whether you think the culture will survive. Enter "y" (for yes) on the keyboard if you think the culture will survive. Otherwise, enter "n" (for no). After pressing the key, you will be told whether your guess is correct or incorrect. Try to guess as accurately as possible.

When doing the task, try to keep track of what happened when one or both chemicals was added to the culture, as well as what happened when neither chemical was added to the culture. However, do not write down this information.

Screen 5

In each game, you will twice be asked to rate how strongly each chemical affects the survival of the culture. You will rate the strength on a scale ranging from -100 to +100. Remember that a POSITIVE number means that you think that the chemical has a positive effect on the culture's survival. That is, the culture is more likely to survive if the chemical is added than when it is not added.

A NEGATIVE number means that you think that the chemical has a negative effect on a culture's survival. That is, the culture is more likely to die if the chemical is added than when the chemical is not added.

And 0 means that the chemical does not systematically aid in the culture's survival, nor does it systematically interfere with that culture's survival.

The number you enter indicates how strongly positive or negative you think is the effect of the chemical on the bacteria. 100 means that the chemical has a very strong positive effect on the culture's survival, while a rating such as 50 means that the chemical has a moderately positive effect on the culture's survival. Similarly, -100 means that the chemical has a strong negative effect on the culture's survival, while a rating such as -25 means that the chemical has a weak negative effect on the culture's survival.

II Instructions For Experiment 2 (2E - 1C)

Screen 1

Imagine that scientists have recently discovered 4 strains of bacteria and 4 pairs of chemicals that exist in the mammalian digestive system. The scientists are studying whether the production of these chemicals is aided, interfered with, or not affected by the 4 strains of bacteria.

To do this, blood samples were obtained from laboratory rats and the presence or absence of each of the two chemicals was verified. Later, the scientists tested whether there was bacteria in the rat's blood.

Laboratory rats, instead of people, were used so that the effects of variables such as diet, drug history, lifestyle factors, and socio-economic status could be controlled for. In addition, to make the results easy to interpret, the scientists conducted the study under aseptic and highly
controlled conditions. That is, the animals were kept in filtered cages that prevented contaminants in the air from entering the cage; the room temperature was kept constant; and all the animals had the same diet.

Screen 2

At the time of these experiments, it was not known how the production of each chemical might be affected by the bacteria. On one hand, a chemical might be MORE likely to be produced if the bacteria are present, whereas that chemical would be less likely to be produced in their absence. This would be an example of the chemical's production being positively influenced by the bacteria.

Alternatively, the chemical might be LESS likely to be produced if the bacteria are present than if they are absent. This would be an example of the chemical's production being NEGATIVELY influenced by the bacteria.

Finally, the production of a chemical might not be influenced by the bacteria. In other words, the production of a chemical might not be aided or interfered with by the bacteria.

Screen 3

To assess these possibilities, the scientists investigated whether the bacteria were present or absent when either chemical or both chemicals were produced, as well as when no chemicals were produced. A comparison of what happened on these trials allowed the scientists to assess whether each chemical's production was POSITIVELY affected, NEGATIVELY affected, or NOT affected by the bacteria.

Screen 4

You will be presented with the results from this study. On each trial, you will be told whether the animal's blood contains one chemical, the other chemical, both chemicals, or neither chemical. You will then decide whether you think the rat has the bacteria. Enter "y" (for yes) on the keyboard if you think the rat has the bacteria. Otherwise, enter "n" (for no). After pressing the key, you will be told whether your guess is correct or incorrect. Try to guess as accurately as possible.

When doing the task, try to keep track of how the presence and absence of each chemical is affected by the presence or absence of the bacteria. However, do not write down this information.

Screen 5

In each game, you will twice be asked to rate how strongly the production of each chemical is affected by the bacteria. You will rate the strength on a scale ranging from +100 to -100. Remember that a POSITIVE number means that you think that a chemical's production is positively affected by the bacteria. That is, the production is more likely to be due to the presence rather than the absence of the bacteria.

A NEGATIVE number means that you think a chemical's production is negatively affected by the bacteria. That is, the production of the chemical is more likely to occur if the bacteria are absent rather than present.

And 0 means that the production of a chemical is not systematically aided or interfered with by the bacteria.

The number you enter indicates how strongly positive or negative you think a chemical's production is affected by the bacteria. 100 means that the production of a chemical is strongly aided by the bacteria, while a rating such as 50 means that the production is moderately aided by the bacteria. Similarly, -100 means that the production of a chemical's production is strongly interfered with by the bacteria, while a rating such as -25 means that its production is weakly interfered with by the bacteria.

III Instructions for Experiment 3 (1C - 2E)

Screen 1

Imagine that scientists have recently discovered 4 strains of bacteria that live in the human digestive system and aid in the digestive process. For each strain, the scientists are studying whether the bacteria aid in, interfere with, or have no effect on the production of certain pairs of chemicals.

To do this, the bacteria were either added or not added to an environment that simulated the human digestive system. A few hours later, the scientists verified the presence or absence of each of the two chemicals.

Screen 2

At the time of the experiment, it was not known how the bacteria might affect the production of each chemical. On one hand, the bacteria might make it MORE likely that a chemical will be produced, whereas that chemical would be less likely to be produced without the bacteria. This would be an example of the bacteria having a POSITIVE effect on the production of one or both
chemicals.

Alternatively, the bacteria might make it LESS likely that a chemical will be produced, whereas that chemical would be more likely to be produced without the bacteria. This would be an example of the bacteria having a NEGATIVE effect on the production of one or both chemicals.

Finally, the bacteria might have NO systematic effect on a chemical's production. That is, it could be that the bacteria neither systematically aid in nor systematically interfere with the production of one or both chemicals.

Screen 3

To assess these possibilities, the scientists investigated what happened when the bacteria were added to the simulated human digestive system, as well as what happened on control trials in which the bacteria were not added to the digestive system. A comparison of what happened on these trials allowed the scientists to assess whether the bacteria had a POSITIVE effect, a NEGATIVE effect, or NO effect on each chemical's production.

Screen 4: Two-Decision Group

You will be presented with the results from this study. On each trial, you will be told whether the bacteria were added or not added to the simulated human digestive system. You will then decide whether you think each chemical is produced. Enter "y" (for yes) on the keyboard if you think the chemical is produced. Otherwise, enter "n" (for no). After pressing the key, you will
be told whether your guess is correct or incorrect. Try to guess as accurately as possible.

When doing the task, try to keep track of what happened when the bacteria were added, as well as what happened when the bacteria were not added to the digestive system. However, do not write down this information.

Screen 4: One-Decision Group

You will be presented with the results from this study. On each trial, you will be told whether the bacteria were added or not added to the simulated human digestive system. You will then decide whether you think

a) one chemical (e.g., Chemical A) was produced
b) the other chemical (e.g., Chemical B) was produced
c) both chemicals (e.g., Chemicals A and B) were produced

OR

d) neither chemical was produced

After entering your choice of "a", "b", "c" or "d", you will be told whether your guess is correct or incorrect, as well as what happened on that trial. Try to guess as accurately as possible.

When doing the task, try to keep track of what happened when the bacteria were added, as well as what happened when the bacteria were not added to the digestive system. However, do not write down this information.

Screen 5

In each game, you will twice be asked to rate how strongly the bacteria affect the production of each chemical. You will rate the strength on a scale ranging from -100 to +100. Remember that a POSITIVE number means that you think that the bacteria have a positive effect on the chemical's production. That is, the chemical is more likely to be produced if the bacteria are added than when they are not added.

A NEGATIVE number means that you think that the bacteria have a negative effect on a chemical's production. That is, the chemical is less likely to be produced if the bacteria are added than if the bacteria are not added.

And 0 means that the bacteria do not systematically aid in nor systematically interfere with the chemical's production.

The number you enter indicates how strongly positive or negative you think is the effect of the bacteria on the chemical's production. 100 means that the bacteria have a very strong positive effect on the chemical's production, while a rating such as 50 means that the bacteria have a
moderately positive effect on the chemical's production. Similarly, -100 means that the bacteria have a strong negative effect on the chemical's production, while a rating such as -25 means that the bacteria have a weak negative effect on the chemical's production.

IV Instructions for Experiment 4 (1E - 2C: ?EC)

Screen 1

Imagine that scientists have discovered 4 strains of bacteria that live in the human digestive system and aid in the digestive process. For each strain, the scientists wish to test whether the bacteria's survival is aided, interfered with, or not affected by certain pairs of chemicals.

To do this, a strain of bacteria was first placed in cultures (petri dishes) containing

1) one chemical (e.g., chemical A)
2) the other chemical (e.g., chemical B)
3) both chemicals (e.g., chemicals A and B)

OR

4) neither chemical.

A few hours later, the scientists tested whether or not the bacteria survived, and verified which chemical or chemicals were in the cultures.

The scientists used the "blind technique" when doing these experiments. That is, they did not know which chemical or chemicals were in the cultures until after they had tested whether the bacteria survived. This step was taken to ensure that the scientists' biases would not influence the results of their experiments or how the results were interpreted.

Screen 2

At the time of these experiments, it was not known how each chemical might affect the bacteria's survival. On one hand, the bacteria might be MORE likely to survive if a chemical is present, whereas they would be less likely to survive without the chemical. This would be an example of the bacteria's survival being POSITIVELY affected by one or both chemicals.

Alternatively, the bacteria might be LESS likely to survive if a chemical is present, whereas they would be more likely to survive without the chemical. This would be an example of the bacteria's survival being NEGATIVELY affected by one or both chemicals.

Finally, the bacteria's survival might not be affected by a chemical. That is, the bacteria's survival might not be aided or interfered with by one or both chemicals.

Screen 3

To assess these possibilities, the scientists investigated whether the bacteria survived or did not survive when either or both chemicals were in the culture. They also tested what happened on control trials in which neither chemical was in the culture. A comparison of what happened on these trials allowed the scientists to assess whether the bacteria's survival was POSITIVELY affected, NEGATIVELY affected, or NOT affected by each chemical.

Screen 4

The results of each experiment will be presented to you. On each trial, you will be told whether the bacteria survived or did not survive. You will then decide whether you think each chemical was present. Enter "y" (for yes) on the keyboard if you think the chemical was present. Otherwise, enter "n" (for no). After pressing the key, you will be told whether your guess is correct or incorrect. Try to guess as accurately as possible.

When doing the task, try to keep track of how the bacteria's survival or lack of survival is affected by the presence or absence of each chemical. However, do not write down this information.

Screen 5

In each game, you will twice be asked to rate how strongly the bacteria's survival is affected by each chemical. You will rate the strength on a scale ranging from + 100 to -100. Remember that a POSITIVE number means that you think that the bacteria's survival is positively affected by the chemical. That is, the bacteria are MORE likely to survive if the chemical is present than if it is absent.

A NEGATIVE number means that you think the bacteria's survival is negatively affected by the chemical. That is, the bacteria are LESS likely to survive if the chemical is present than if it is absent.

And 0 means that the culture's survival is not systematically aided or interfered with by the chemical.

The number you enter indicates how strongly positive or negative you think the bacteria are affected by each chemical. 100 means that the bacteria's survival is strongly aided by the chemical, while a rating such as 50 means that the bacteria's survival is moderately aided by the chemical. Similarly, -100 means that the bacteria's survival is strongly interfered with by the chemical, while a rating such as -25 means that the bacteria's survival is weakly interfered with by the chemical.

V Instructions for Experiment 5 (1E - 2C: ?CE)

Screen 1

Imagine that scientists have discovered 4 strains of bacteria that live in the human digestive system and aid in the digestive process. For each strain, the scientists wish to test whether certain pairs of chemicals aid, interfere with, or do not affect the bacteria's survival.

To do this, a strain of bacteria was first placed in cultures (petri dishes) containing

1) one chemical (e.g., chemical A)
2) the other chemical (e.g., chemical B)
3) both chemicals (e.g., chemicals A and B)
OR
4) neither chemical.

A few hours later, the scientists tested whether or not the bacteria survived, and verified which chemical or chemicals were in the cultures.

The scientists used the "blind technique" when doing these experiments. That is, they did not know which chemical or chemicals were in the cultures until after they had tested whether the bacteria survived. This step was taken to ensure that the scientists' biases would not influence the results of their experiments or how the results were interpreted.

Screen 2

At the time of these experiments, it was not known how a chemical might affect the bacteria's survival. On one hand, the bacteria might be MORE likely to survive if a chemical is present, whereas they would be less likely to survive without the chemical. This would be an example of a chemical POSITIVELY affecting the bacteria's survival.

Alternatively, the bacteria might be LESS likely to survive if a chemical is present, whereas they would be more likely to survive without the chemical. This would be an example of a chemical NEGATIVELY affecting the bacteria's survival.

Finally, a chemical might not affect the bacteria's survival. That is, a chemical might neither aid nor interfere with the bacteria's survival.

Screen 3

To assess these possibilities, the scientists investigated whether the bacteria survived or did not survive when either or both chemicals were in the culture. They also tested what happened on control trials in which neither chemical was in the culture. A comparison of what happened on these trials allowed the scientists to assess whether each chemical POSITIVELY affected, NEGATIVELY affected, or DID NOT affect the bacteria's survival.

Screen 4

The results of each experiment will be presented to you. On each trial, you will be told whether the bacteria survived or did not survive. You will then decide whether you think each chemical was present. Enter "y" (for yes)on the keyboard if you think the chemical was present. Otherwise, enter "n"(for no). After pressing the key, you will be told whether your guess is correct or incorrect. Try to guess as accurately as possible.

When doing the task, try to keep track of how the bacteria's survival or lack of survival is affected by the presence or absence of each chemical. However, do not write down this information.

Screen 5

In each game, you will twice be asked to rate how strongly each chemical affected the bacteria's survival. You will rate the strength on a scale ranging from + 100 to -100. Remember that a POSITIVE number means that you think that the chemical positively affects the bacteria's survival. That is, the bacteria are MORE likely to survive if the chemical is present than if it is absent.

A NEGATIVE number means that you think the chemical negatively affects the bacteria's survival. That is, the bacteria are LESS likely to survive if the chemical is present than if it is absent.

And 0 means that the chemical does not systematically aid in or interfere with the bacteria's survival.

The number you enter indicates how strongly positive or negative you think each chemical affects the bacteria. 100 means that the chemical strongly aids in the bacteria's survival, while a rating such as 50 means that the chemical moderately aids in the bacteria's survival. Similarly, -100 means that the chemical strongly interferes with the bacteria' survival, while a rating such as -25 means that the chemical weakly interferes with the bacteria's survival.

VI Post-Experimental Questionnaire Used in Experiments 5, 6 and 7

#:____
Condition: ____

1. In this study, the scientists were examining (please circle):

A. how the chemicals were affecting the bacteria
B. how the bacteria were affecting the chemicals
C. neither A nor B (please write down what you think):

2. On a pool table, a white ball was hit by a pool cue, which caused the white ball to move. In this study, do you think the chemicals were more like the pool cue or the movement of the white ball (please circle)?

A. pool cue
B. white ball
C. neither A nor B (please write down what you think):

3. Do you think there were any variables that the scientists did not take into account that could have influenced the bacteria's survival (i.e., the results in the experiments):

YES NO

If you circled yes, please list some.

VII Instructions for Experiment 6 (1E - 2C: ?CE)

Screens 1 -3 & Screen 6 are identical to those of Experiment 5

Screen 4

The results of each experiment will be presented to you. On each trial, you will be told whether the bacteria survived or did not survive. After pressing the spacebar, you will be told whether

a) one chemical
b) the other chemical
c) both chemicals
OR
d) neither chemical

was present.

When doing the task, try to keep track of how the bacteria's survival or lack of survival is affected by the presence or absence of each chemical. However, do not write down this information.

VIII Instructions for Experiment 7

GROUP: CE

2 Causes - 1 Effect (2C - 1E)

Screen 1

Imagine that scientists have recently discovered 2 strains of bacteria that exist in the mammalian digestive system. The scientists are studying whether certain pairs of chemicals affect the bacteria's survival. For each strain, the scientists are testing whether the chemicals aid in, interfere with, or have no effect on a strain's survival.

To do this, a strain of bacteria was first placed in culture (petri dishes). After that,

1) one chemical (e.g., chemical A)
2) the other chemical (e.g., chemical B)
3) both chemicals (e.g., chemicals A and B)
   OR
4) neither chemical

was added to the bacterial culture. A few hours later, the scientists verified whether or not the bacterial sample survived. The time at which this measurement was done was constant within each experiment.

Screen 2

At the time of the experiment, it was not known what effect each chemical might have on the bacteria. On one hand, a chemical might make the bacteria MORE likely to survive, whereas the sample would be less likely to survive without the chemical. This would be an example of a chemical having a POSITIVE effect on the bacteria's survival.

Alternatively, a chemical might make it LESS likely that the bacteria will survive, whereas the sample would be more likely to survive without the chemical. This would be an example of a chemical having a NEGATIVE effect on the bacteria's survival.

Finally, a chemical might have NO systematic effect on the bacteria's survival. That is, it could be that the chemical neither aids in nor interferes with the bacteria's survival.

Screen 3

To assess these possibilities, the scientists investigated what happened when one or two chemicals were added to the bacterial sample. They also tested what happened on

control trials in which no chemicals were added. A comparison of what happened on these trials allowed the scientists to assess whether a chemical had a POSITIVE effect, a NEGATIVE effect, or NO effect on the bacteria's survival.

Screen 4

To ensure that the results were reliable, the following measures were taken:

* The scientists verified that the chemicals used did not interact. That is, when mixed, they neither neutralized each other nor formed a more potent compound.
* Similar concentrations of chemicals were used in each experiment.
* The age and concentration of the bacteria were similar in all conditions.
* The optimal conditions for the bacteria's survival were first established. That is, each strain's optimal temperature, pH, lighting, and nutrients were verified prior to beginning the experiments and these conditions were consistently used. The time at which the bacteria's presence or absence was verified was slightly longer than the baseline survival without the chemicals.
* The experiments were conducted under sterile conditions. That is, the cultures were first checked to ensure that they were not contaminated. As well, the scientists ensured that the samples were not exposed to contaminants in the air.
* The scientists verified that the test used to classify the bacteria as surviving or not surviving was reliable. In previous studies, it was found to yield results equally reliable to those obtained from counting the bacteria before and after adding chemicals whose actions were known to bacterial samples.

Screen 5

You will be presented with the results from this study. On each trial, you will be told whether one chemical (e.g., chemical A), the other chemical (e.g., chemical B), both chemicals, or no chemicals were added to the bacterial sample. After pressing the spacebar, you will be told whether the bacteria survived or did not survive.

When doing the task, try to keep track of what happened when one or both chemicals was added, as well as what happened when neither chemical was added. However, do not write down this information.

Screen 6

In each game, you will twice be asked to rate how strongly each chemical affects the bacteria's survival. You will rate the strength on a scale ranging from -100 to +100. Remember that a POSITIVE number means that you think that the chemical has a

positive effect on the bacteria's survival. That is, the bacteria are more likely to survive if the chemical is added than if it is not added.

A NEGATIVE number means that you think that the chemical has a negative effect on the bacteria's survival. That is, the bacteria are less likely to survive if the chemical is added than if it not added.

And 0 means that the chemical does not systematically aid in nor interfere with the bacteria's survival.

The number you enter indicates how strongly positive or negative you think is the chemical's effect on the bacteria. 100 means that the chemical has a very strong positive effect, while a rating such as 50 means that the chemical has a moderately positive effect on the bacteria's survival. Similarly, -100 means that the chemical has a strong negative effect on the
bacteria's survival, while a rating such as -25 means that the chemical has a weak negative effect on the bacteria's survival.

1 Cause - 2 Effects (1C - 2E)

Screen 1

Imagine that scientists have recently discovered 2 strains of bacteria that exist in the mammalian digestive system. For each strain, the scientists are studying whether the bacteria aid in, interfere with, or have no effect on the production of certain pairs of chemicals.

To do this, the bacteria were added or not added to an environment that simulated the mammalian digestive system. A few hours later, the scientists verified the presence or absence of each of the two chemicals.

Screen 2

At the time of the experiment, it was not known how the bacteria might affect the production of each chemical. On one hand, the bacteria might make it MORE likely that a chemical will be produced, whereas that chemical would not be produced without the bacteria. This would be an example of the bacteria having a POSITIVE effect on the production of a chemical.

Alternatively, the bacteria might make it LESS likely that a chemical will produced, whereas that chemical would be more likely to be produced without the bacteria. This would be an example of the bacteria having a NEGATIVE effect on a chemical's production.

Finally, the bacteria might have no systematic effect on a chemical's production. That is, it could be that the bacteria neither aid in nor interfere with a chemical's production.

Screen 3

To assess these possibilities, the scientists investigated what happened when the bacteria were added, as well as what happened on control trials in which the bacteria were not added to the digestive environment. A comparison of what happened on these trials allowed the scientists to assess whether the bacteria had a POSITIVE effect, a NEGATIVE effect, or NO effect on a chemical's production.

Screen 4

To ensure that the results were reliable, the following measures were taken:

* The age and concentration of the bacteria were similar in all conditions.
* An environment that simulated the mammalian digestive system was used so other factors such as time since eating, time of day, hormonal levels, temperature, pH, lighting, and nutrients could be controlled and maintained at consistent levels in all experiments.
* The experiments were conducted under sterile conditions. That is, the digestive environment was first checked to ensure that it was not contaminated.
* The scientists verified that the test used to classify the presence or absence of each chemical was reliable.

Screen 5

You will be presented with the results from this study. On each trial, you will be told whether the bacteria was added or not added to the digestive system. After pressing the spacebar, you will be told whether

a) one chemical (e.g., chemical A)
b) the other chemical (e.g., chemical B)
c) both chemicals
OR
d) neither chemical

was produced.

When doing the task, try to keep track of what happened when the bacteria was added, as well as what happened when the bacteria were not added. However, do not write down this information.

Screen 6

In each game, you will twice be asked to rate how strongly the bacteria affect the production of each chemical. You will rate the strength on a scale ranging from -100 to +100. Remember that a POSITIVE number means that you think that the bacteria have a positive effect on the chemical's production. That is, the chemical is more likely to be produced if the bacteria are added that when they are not added.

A NEGATIVE number means that you think that the bacteria have a negative effect on a chemical's production. That is, the chemical is less likely to be produced if the bacteria are added that if they are not added.

And 0 means that the bacteria do not systematically aid in nor interfere with the chemical's production.

The number you enter indicates how strongly positive or negative you think is the effect of the bacteria on the chemical's production. 100 means that the bacteria have a very strong positive effect on the chemical's production, while a rating such as 50 means that the bacteria have a moderately positive effect on the chemical's production. Similarly, -100 mean that the bacteria have a strong negative effect on the chemical's production, while a rating such as -25 means that the bacteria have a weak negative effect on the chemical's production.

GROUP: EC

2 Effects - 1 Cause

Screen 1

Imagine that scientists have recently discovered 2 strains of bacteria and that exist in the mammalian digestive system. For each strain, the scientists are studying whether the bacteria aid in, interfere with, or have no effect on the production of certain pairs of chemicals.

To do this, blood samples were obtained from laboratory rats and the presence or absence of each of the two chemicals was verified. Later, the scientists tested whether there was bacteria in the rat's blood.

Screen 2

At the time of these experiments, it was not known how the bacteria might affect the production of each chemical. On one hand, a chemical might be MORE likely to be produced if the bacteria are present, whereas that chemical would be less likely to be produced in their absence. This would be an example of the bacteria POSITIVELY influencing a chemical's production.

Alternatively, the chemical might be LESS likely to be produced if the bacteria are present than if they are absent. This would be an example of the bacteria NEGATIVELY influencing a chemical's production.

Finally, the bacteria might not influence a chemical's production. In other words, the bacteria might neither aid nor interfere with a chemical's production.

Screen 3

To assess these possibilities, the scientists investigated whether the bacteria were present or absent when either chemical or both chemicals were produced, as well as when no chemicals were produced. A comparison of what happened on these trials allowed the scientists to assess whether each chemical's production was POSITIVELY affected, NEGATIVELY affected, or NOT affected by the bacteria.

Screen 4

To ensure that the results were reliable, the following measures were taken:
* Laboratory rats, instead of people, were used so that the effects of variables such as diet, drug history, lifestyle factors, and socio-economic status could be controlled for.
* The study was conducted under aseptic conditions. That is, the animals were kept in filtered cages that prevented contaminants in the air from entering the cage; the room temperature was kept constant; and all the animals were of similar age and had the same diet.
* Before testing for the bacteria and chemicals, the scientists verified that hormonal levels were constant.
* The tests for the bacteria and chemicals were done at the same time every day, between 8 a.m. and 9 a.m.
* Food was removed from the cages on the evening prior to the day on which the chemicals and bacteria were tested for. This measure was taken so that the time since the last meal would be constant.
* Prior studies had shown that the two chemicals did not interact. That is, the two chemicals did not neutralize each other or form a different compound.
* The scientists verified that the tests used to test for the presence or absence of the chemicals and bacteria were reliable.

Screen 5

You will be presented with the results from this study. On each trial, you will be told whether the animal's blood contains one chemical, the other chemical, both chemicals, or neither chemical. After pressing the spacebar, you will be told whether the rat has or does not have the bacteria.

When doing the task, try to keep track of how the presence and absence of the bacteria affects the presence or absence of the chemicals. However, do not write down this information.

Screen 6

In each game, you will twice be asked to rate how strongly the bacteria affect the production the production of each chemical. You will rate the strength on a scale ranging from +100 to -100. Remember that a POSITIVE number means that you think that the bacteria POSITIVELY affects a chemical's production. That is, the chemical is more likely to be produced if the bacteria are present than when they are absent.

A NEGATIVE number means that you think the bacteria have a negative effect on a chemical's production. That is, the chemical's production is more likely to occur if the bacteria are absent rather than present.

And 0 means that the bacteria does not systematically aid in nor interfere with a chemical's production.

The number you enter indicates how strongly positive or negative you think the bacteria affects a chemical's production. 100 means that the bacteria strongly aids in a chemical's production, while a rating such as 50 means that the bacteria have a moderately positive effect on the chemical's production. Similarly, -100 means that the bacteria have a strong negative effect on a chemical's production, while a rating such as -25 means that the bacteria have a weak negative effect on the chemical's production.

1 Effect - 2 Causes

Screen 1

Imagine that scientists have discovered 2 strains of bacteria that exist in the mammalian digestive system. For each strain, the scientists wish to test whether certain pairs of chemicals aid in, interfere with, or do not affect the bacteria's survival.

To do this, a strain of bacteria was first placed in cultures (petri dishes) containing

1) one chemical (e.g., chemical A)
2) the other chemical (e.g., chemical B)
3) both chemicals (e.g., chemicals A and B)
OR
4) neither chemical.

A few hours later, the scientists tested whether or not the bacteria survived, and verified which chemical or chemicals were in the cultures.

Screen 2

At the time of these experiments, it was not known how a chemical might affect the bacteria's survival. On one hand, the bacteria might be MORE likely to survive if a chemical is present, whereas they would be less likely to survive without the chemical. This would be an example of a chemical POSITIVELY affecting the bacteria's survival.

Alternatively, the bacteria might be LESS likely to survive if a chemical is present, whereas they would be more likely to survive without the chemical. This would be an example of a chemical NEGATIVELY affecting the bacteria's survival.

Finally, a chemical might not affect the bacteria's survival. That is, a chemical might neither aid nor interfere with the bacteria's survival.

Screen 3

To assess these possibilities, the scientists investigated whether the bacteria survived or did not survive when either or both chemicals were in the culture. They also tested what happened on control trials in which neither chemical was in the culture. A comparison of what happened on these trials allowed the scientists to assess whether each chemical POSITIVELY affected, NEGATIVELY affected, or DID NOT affect the bacteria's survival.

Screen 4

To ensure that the results were reliable, the following measures were taken:

* The scientists verified that the chemicals used did not interact. That is, when mixed, they neither neutralized each other nor formed a more potent compound.
* Similar concentrations of chemicals were used in each experiment.
* The age and concentration of the bacteria were similar in all conditions.
* The optimal conditions for the bacteria's survival were first established. That is, each strain's optimal temperature, pH, lighting, and nutrients were verified prior to beginning the experiments and these conditions were consistently used. The time at which the bacteria's presence or absence was verified was slightly longer than the baseline survival without the
chemicals.
* The experiments were conducted under sterile conditions. That is, the cultures were first checked to ensure that they were not contaminated. As well, the scientists ensured that the samples were not exposed to contaminants in the air.
* The scientists verified that the test used to classify the bacteria as surviving or not surviving was reliable. In previous studies, it was found to yield results equally reliable to those obtained from counting the bacteria before and after adding chemicals whose actions were known to bacterial samples.

Screen 5

The results of each experiment will be presented to you. On each trial, you will be told whether the bacteria survived or did not survive. After pressing the spacebar, you will be told whether

a) one chemical
b) the other chemical
c) both chemicals
OR
d) neither chemical

was present.

When doing the task, try to keep track of how the bacteria's survival or lack of survival is affected by the presence or absence of each chemical. However, do not write down this information.

Screen 6

In each game, you will twice be asked to rate how strongly each chemical affected the bacteria's survival. You will rate the strength on a scale ranging from + 100 to -100. Remember that a POSITIVE number means that you think that the chemical positively affects the bacteria's survival. That is, the bacteria are MORE likely to survive if the chemical is present than if it is absent.

A NEGATIVE number means that you think the chemical negatively affects the chemical's survival. That is, the bacteria are LESS likely to survive if the chemical is present than if it is absent.

And 0 means that the chemical does not systematically aid in or interfere with the bacteria's survival.

The number you enter indicates how strongly positive or negative you think each chemical affects the bacteria. 100 means that the chemical strongly aids in the bacteria's survival, while a rating such as 50 means that the chemical moderately aids in the bacteria's survival. Similarly, -100 means that the chemical strongly interferes with the bacteria' survival, while a rating such as -25 means that the chemical weakly interferes with the bacteria's survival.

ENDNOTES

1. Waldmann (2000) reported four experiments. For brevity, only the first two that are most relevant to the present exposition will be described. In the first experiment, participants were presented with 4 causes of a single effect (4C-1E) or four effects of a single cause (4E-1C). In both cases, the antecedent cues were coloured lights and the consequent cue was a lamp being turned on or off (4C-1E) or a switch being turned on or off (4E-1C). In the first phase, P was established as a perfect predictor of the cue's outcome (i.e., P+). That is, the outcome occurred when P was on and did not occur when P was off. After 6 trials, judgments of P were requested. In the second phase, P was presented with a redundant second cue; this compound was followed by the outcome (PR+). Two other cues, C and D, were also presented in this phase. These cues occurred in compound and were also perfect predictors of the outcome (CD+). Thus, the four events in the second phase were a) P and R were on and the outcome occurred, b) P and R were off and the outcome did not take place, c) C and D were on and the outcome was observed, and d) C and D were off and the outcome did not follow. When information about one of the compounds was presented (e.g., P and R), none was presented about the other 2 cues (e.g., C and D). After 12 trials, judgments about all four cues were requested. Blocking of R by P was defined by comparing R to C and D; if R was lower than C and D, then blocking was observed. According to causal model theory, judgments of R should have been lower than C or D in 4C-1E but not in 4E-1C, whereas the R-W model would predict that R would be lower than C or D in both 4C-1E and 4E-1C. Ratings of P were close to 100 in both 4C-1E and 4E-1C. R was blocked by P in 4C-1E but not in 4E-1C. This result was interpreted as being consistent with causal model theory.

In the Introduction to Experiment 2, Waldmann (2000) reasoned that the lack of blocking in 4E-1C occurred because the first phase of Experiment 1 failed to provide any information about R because this cue was not mentioned until Phase 2. If participants then assumed that the redundant effect occurred throughout Phase I but was merely not mentioned, then they would have inferred that the cause had a second effect. But if people interpreted the absence of information about R in Phase I as implying that the cue itself

was absent, then blocking should initially occur because the unconditional contingency for R was interpreted as being lower than for P. But with additional training trials, R's unconditional contingency should increase and thus the blocking effect should get smaller. Experiment 2 tested the hypothesis that blocking would be observed in a diagnostic condition when R is absent in Phase 1 but not when participants lack knowledge about it. Two groups of participants performed a 3E-1C learning task. The No Information Group was told at the beginning of the task that there are three effects of a single cause. No Information signifies that this group was not provided with any information about the presence or absence of R until Phase 2. On the other hand, the Explicit Absence Group was initially informed that there were two effects of a cause, but at the beginning of the second phase
was told that there were three effects. Explicit Absence thus refers to the fact that R was not present during Phase 1.

In the 3E-1C learning task, substances in the blood of rats (P,R,U) was presented prior to information about the virus Midosis. The substances were effects of the virus. In the first phase, P became a perfect predictor of the cause and U was always absent (uncorrelated). In the second phase, R became a redundant effect of Midosis. Participants gave judgments of P and U after the first phase, and of all three cues in the second phase. In the second phase, the ratings were obtained after 20 and after 40 trials. Consistent with causal model theory, P blocked judgments of R in the Explicit Absence group but not in the No Information group. Furthermore, the size of the blocking effect was smaller after forty Phase 2 trials than after twenty.