

COMPARATIVE COGNITION & BEHAVIOR REVIEWS

2009

Volume 4, pp 116-134

Applying Signal Detection Theory to Contingency Assessment

Shepard Siegel, Lorraine G. Allan, Samuel D. Hannah
McMaster University

Matthew J. C. Crump
Vanderbilt University

In most studies of contingency assessment participants judge the magnitude of the relationship between cues and outcomes. This judgment is a conflated measure of the participant's sensitivity to the cue-outcome relationship, and his or her response bias. A psychophysical model (signal detection theory, SDT) can be used to dissect the independent contributions of sensitivity and bias to contingency judgment. Results of an experiment concerning cue-interaction (blocking) illustrate the utility of applying SDT to understanding contingency assessment. Most accounts of such assessment are associative (derived primarily from Pavlovian conditioning experiments with non-human animals). A psychophysical analysis of contingency assessment is not an alternative to such associative accounts. The SDT analysis supplements (not replaces) learning principles with psychophysical principles.

Keywords: associative learning, blocking, contingency assessment, cue interaction, psychophysics, signal detection theory.

In studies of contingency assessment a participant is asked to judge the relationship between events, a cue and an outcome. Although there are various forms of the task, a discrete trial format with a single cue and a single outcome often is used. On each trial, a cue may, or may not, be presented, following which an outcome may, or may not, be presented. Various cues and outcomes have been used. For example, the cue may consist of information that clouds have or have not been seeded, and the outcome consists of information that it has or has not rained (Ward & Jenkins, 1965); the cue may consist of information that a putative leavening agent

has or has not been added to bread dough, and the outcome consists of information that the dough has or has not risen (Shaklee & Mims, 1981); the cue may consist of information that a patient has or has not a particular symptom, and the outcome consists of information that a patient does or does not have a particular disease (Smedslund, 1963). More generally, the stimuli presented to a participant can be summarized as a 2x2 matrix (see Table 1a). On each trial the cue either is presented (C) or is not presented ($\sim C$), and then the outcome either does occur (O) or does not occur ($\sim O$). The letters in the cells (a, b, c, d) represent the joint frequency of occurrence of the four cue-outcome combinations in a block of trials.

Table 1a. The 2x2 matrix for the cue and outcome presentations in the contingency assessment task. The letters in the cells represent the joint frequency of occurrence of the four cue-outcome combinations in a block of trials.

	O	$\sim O$
C	a	b
$\sim C$	c	d

Preparation of this paper was supported by research grants from the Natural Sciences and Engineering Research Council of Canada to SS and LGA. We thank Rolfe Morrison for assistance with the proof in Appendix 1, and Jennifer Heisz for her insightful comments during laboratory meetings. Correspondence concerning this manuscript should be addressed to either Shepard Siegel (siegel@mcmaster.ca) or Lorraine Allan (allan@mcmaster.ca). The address for Shepard Siegel, Lorraine G. Allan, and Samuel D. Hannah is: Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton ON L8S 4K1, Canada. The address for Matthew J. C. Crump is: Psychology Department, Vanderbilt University, PMB 407817 2301, Vanderbilt Place, Nashville, TN 37240-7817, USA.

Table 1b. The 2x2 matrix for the CS and US presentations in Pavlovian conditioning. The letters in the cells represent the joint frequency of occurrence of the four CS-US combinations in a block of trials.

	US	~US
CS	a	b
~CS	c	d

Although there are various ways of summarizing the experimenter-programmed relationship between the cue and the outcome (see Allan, 1980), one that has been particularly useful is ΔP :

$$\Delta P = P(O|C) - P(O|\sim C) = \frac{a}{a+b} - \frac{c}{c+d} \quad (1)$$

The strength of the relationship between the cue and the outcome is denoted by the magnitude of ΔP . If the outcome always occurred when the cue was presented, and never in the absence of the cue, $\Delta P = 1$. If $\Delta P = 0$, the outcome is equally likely in the presence and absence of the cue.

In general, early contingency assessment researchers assumed that people extract rules from the sequence of cue and outcome presentations; that is, participants are innate statisticians, computing and comparing the conditional probabilities of the outcome in the presence, and in the absence, of the cue (e.g., Peterson & Beach, 1967). Some of the early studies of contingency assessment were designed to ascertain which statistical information in the 2x2 matrix was used to arrive at the participant's contingency judgment (e.g., which cells were most important, see Arkes & Harkness, 1983). Others were designed to ascertain how participants of different ages (e.g., Shaklee & Mims, 1981), or with different psychological disorders (e.g., Alloy & Abramson, 1979), differed on the contingency assessment task. In general, the topic of contingency assessment was rather esoteric and atheoretical until the mid 1980s. Research concerning the topic then grew exponentially (see Figure 1)⁽¹⁾. We suggest that it was in the mid 1980s that experimental psychologists in general, and those interested in basic associative processes in particular, were attracted to the study of contingency assessment.

The Growth of Contingency Assessment Research

The reason for the rapid growth of research on contingency assessment is of considerable historical interest. Coincidentally, at the time that some researchers were investigating how humans decide that cues and outcomes are related to

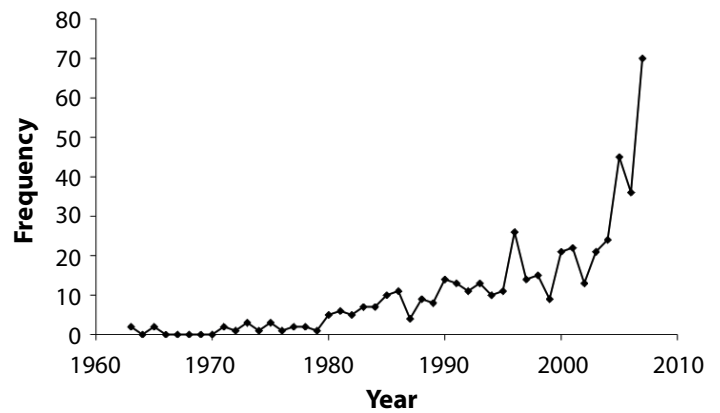


Figure 1. Number of papers concerned with contingency assessment published each year. Papers were selected by searching the PsycINFO® database for studies that used human participants and contained phrases in the title or abstract such as: causal (or contingency) learning, correlation (or covariation or causality or contingency) judgment, contingency assessment, cue interaction, forward blocking, backward blocking, and power PC. The exact search query is described in Endnote 1.

each other, others were investigating new phenomena in the area of nonhuman animal learning. Challenges to the prevalent view that pairing was sufficient to establish an association between events arose from many quarters. Leon Kamin reported intriguing cue-interaction effects: blocking and overshadowing (Kamin, 1968, 1969a, 1969b). Both were demonstrations that a simple pairing analysis of classical conditioning was apparently inadequate. That is, learning about the relationship between a target conditional stimulus (CS) and an unconditional stimulus (US) depended not only upon pairings of the target CS and the US, but also on the associative history and salience of other CSs that were presented in compound with the target CS. Allan Wagner and colleagues reported another phenomenon, cue validity, that made essentially the same point – there are conditions under which CSs and USs are paired, but apparently little is learned about the relationship between them (Wagner, Logan, Haberlandt, & Price, 1968). Perhaps most famously, Robert Rescorla published an influential paper suggesting that the contingency (or correlation) between events, rather than contiguity (or pairing), was the crucial factor in establishing associations (Rescorla, 1967). That is, learning about the relationship between a CS and a US depends not only on trials involving CS-US pairings, but also on presentations of the US during the intertrial interval (when no CS is presented). Rescorla collaborated with Wagner in the development of a model that integrated and made sense of these (and other) then-recent findings about Pavlovian conditioning – the Rescorla-Wagner model (Rescorla & Wagner, 1972, Wagner & Rescorla, 1972).

The Rescorla-Wagner model has been very influential in understanding basic associative processes (studied primarily in nonhuman animals). Furthermore, it inspired the development of other models of Pavlovian conditioning that permitted simulation of results of various conditioning manipulations (e.g., Pearce & Hall, 1980). In addition, “the Rescorla-Wagner model has been the primary export of traditional learning theory to other areas of psychology” (Miller, Barnet, & Grahame, 1995, p. 363); That is, the model has had substantial impact in a variety of disciplines (see Siegel & Allan, 1996). One of these disciplines is contingency assessment

Although early researchers of contingency assessment noted similarities between contingency assessment and Pavlovian conditioning (e.g., Alloy & Abramson, 1979), the Rescorla-Wagner model (and the Pearce-Hall model) was applied to contingency assessment by Dickinson, Shanks, and Evenden (1984). According to this associative analysis, the way we judge whether or not cloud seeding and rain are associated is the same way that a rat judges whether or not a tone and a shock are associated. Consider the situation in which a shock (the US) sometimes is presented in the presence of the tone (the CS), and sometimes in its absence. We can construct a contingency matrix for these CS and US presentations (see Table 1b). Such a CS-US contingency matrix is like the contingency matrix for cue and outcome presentations in a contingency assessment task (compare Table 1b with Table 1a). Dickinson et al. suggested that if contingency assessment was like Pavlovian conditioning, then the most influential models of such conditioning could be useful in understanding contingency assessment.

As suggested by the data depicted in Figure 1, Dickinson et al.’s (1984) insight had a galvanizing effect on contingency assessment research. Many researchers with a background in basic learning research with nonhuman animals broadened their investigations to include the new species and the procedures used in the contingency assessment laboratory. Could a simple associative model, like the Rescorla-Wagner model, be useful in understanding contingency assessment? The answer seemed to be yes. In 1993, Allan reviewed the burgeoning literature and concluded: “Associative models in general, and the R-W [Rescorla-Wagner] model in particular, can account for much of the human contingency data ... Not infrequently, these models have prompted the examination of issues unlikely to have been explored outside the framework of associative models” (Allan, 1993, p. 446). Those “issues” prompted by associative accounts of contingency assessment primarily involved cue-interaction effects (Kamin, 1968, 1969a, 1969b).

Cue-interaction effects originally were delineated in studies of Pavlovian conditioning with non-human animals. The

most influential such effect has been two-phase blocking. If a particular CS (e.g., a tone, generally termed a companion cue, C_C) has been associated with a US, and the C_C subsequently is compounded with a second CS (e.g., a light, more generally termed a target cue, C_T), with this C_C - C_T compound still being paired with the US, little seems to be learned about the C_T -US relationship (despite many C_T -US pairings). That is, prior training with one component of a compound appears to block the conditioning of a second component; C_C blocks the C_T -US association. As would be expected on the basis of an associative analysis, such two-phase blocking has been demonstrated using traditional contingency assessment procedures (see Allan, 1993; Shanks, 2007 for reviews). For example, in an experiment by Chapman (1991), participants received information about fictitious medical patients. Each patient exhibited symptoms, and the participant’s task was to predict the likelihood that the patient suffered from “morolis” (a fictitious disease). In the first phase of the experiment, participants received many pairings of a particular symptom (e.g., coughing) with morolis. In a second phase, this symptom was compounded with a second symptom (e.g., dizziness), with the syndrome still being paired with morolis. In this example, coughing is the C_C and dizziness is the C_T . Blocking was demonstrated; participants reported only a minimal relationship between dizziness and morolis (compared to various within-subject control conditions).

According to the Rescorla-Wagner and similar models, blocking results because the participant does not associate C_T with the US. Applying such reasoning to contingency assessment, little is learned about the relationship between a particular target cue and outcome if this C_T is presented in the presence of a C_C that is a more reliable predictor of the outcome. Research inspired by the associative analysis of contingency assessment has progressed since Allan’s (1993) summary. Cue-interaction effects have repeatedly been demonstrated using Kamin’s two-phase blocking procedure, as well as other, related cue-interaction preparations: one-phase blocking (e.g., Baker, Mercier, Vallee-Tourangeau, Frank, & Pan, 1993; Spellman, 1996a), relative cue validity (e.g., Wasserman, 1990), and overshadowing (e.g., Waldmann, 2001). These and other results continued to support the Rescorla-Wagner interpretation. However, some investigators have reported findings inconsistent with the original version of this model, and have suggested modifications, or alternative associative models (see Allan & Tangen, 2005; Shanks, 2007).

In contrast with findings that are consistent with an associative analysis of contingency assessment, we (e.g., Allan, Siegel, & Tangen, 2005; Allan, Hannah, Crump, & Siegel, 2008) and others (e.g., Perales, Catena, Shanks, & González, 2005) have reported results that are not readily incorporated

into the Rescorla-Wagner model (or any other such associative account of contingency assessment). A problem with such accounts can be illustrated with a hypothetical example. Consider the contingency assessment task where the cue consists of information that a patient does or does not have a particular symptom, and the outcome consists of information that a patient does or does not have a particular disease. If the consequences of not correctly diagnosing the disease are catastrophic for the patient (e.g., debilitating chronic illness), and the effects of erroneously concluding that an illness-free patient has the disease are minimal (e.g., needless administration of innocuous medication), the diagnostician may well be biased to conclude that the contingency is high, even though the actual ΔP value is small. This is an example of how the costs and benefits of various conclusions about the relationship between a cue and outcome (termed the payoff matrix) affect the contingency assessment, independently of the ΔP value. Indeed, we (Allan et al., 2008), and others (Perales et al., 2005), recently have demonstrated that payoff matrix manipulations affect contingency assessment in the laboratory.

The fact that contingency assessment is affected by variables other than the programmed relationship between cue and outcome (such as the payoff matrix) has led us to conclude that a complete analysis of such assessment should incorporate psychophysical principles (Allan et al., 2005, 2008; Allan, Siegel, & Hannah, 2007). One purpose of this article is to suggest that psychophysics in general, and one psychophysical model in particular (signal detection theory, SDT), can profitably be applied to understanding contingency assessment. Since cue-interaction effects have played such a central role in theorizing about the mechanisms of contingency assessment, we present new data demonstrating the utility of applying SDT to understanding cue-interaction effects in contingency assessment. Finally, we make the case that a signal detection analysis of contingency assessment is not an alternative to associative models. Rather, it inspires an analysis of learning that incorporates psychophysical as well as associative principles.

Psychophysics and Contingency Assessment

“Psychophysics is the study of the relationship between physical events and our internal experience of these physical events” (Allan & Siegel, 2002, p. 419). In the contingency judgment task, the presentation of a series of cues and outcomes with a particular statistical relationship between them constitute physical events. We judge the strength of this relationship between the cues and outcomes based on our internal experience of these events. It would seem that contingency assessment would be a topic of considerable interest to psychophysicists. It is not. With very few excep-

tions, research concerned with contingency assessment and research concerned with psychophysics have progressed independently, each with its own traditions and each motivated by different theoretical perspectives and models. As discussed, most contingency assessment research has been motivated by the theoretical perspectives and traditions derived from associative learning. Before making the case that psychophysical concepts should be applied to contingency assessment, we briefly review the basics of a psychophysical approach.

Separating Sensitivity and Response Bias

As indicated above, the participant’s response about the relationship between the cue and the outcome, while partially determined by the participant’s sensitivity to the actual contingency, often is influenced by other variables that bias the participant towards judging a particular contingency as either stronger or weaker than the actual ΔP value. The payoff matrix is one example of such a biasing effect. It is as if the participant asks himself or herself two questions before offering an estimate of the cue-outcome relationship: “What do I perceive to be the magnitude of the relationship?”, and, given that perception, “how should I respond?”. The participant’s final judgment of the contingency is determined both by the participant’s sensitivity to the physical contingency and by *biasing* variables. The traditional contingency task confounds these two sources of information. The participant’s unitary judgment of the cue-outcome relationship is comprised of a combination of the participant’s sensitivity to the contingency and the participant’s response bias.

Psychophysical methodology and models provide the tools for isolating the effects of sensitivity and bias. For example, consider the *method of constant stimuli*, a psychophysical procedure initially used to evaluate the participant’s ability to discriminate among stimuli varying along a dimension (e.g., tones differing in intensity). On each trial, one of the tones is randomly selected and presented to the participant, who is required to make a binary judgment – is the tone “loud” or “soft?” Each tone is presented many times. At the end of the experiment, the probability of one of the responses, for example “loud,” $P(R_1)$, is determined for each intensity. The function relating $P(R_1)$ to the physical intensity values is referred to as a psychometric function and is shown schematically in Figure 2. Usually the empirical psychometric function is ogival in shape and can be described by two parameters: the slope and the point-of-subjective equality, PSE. PSE is defined as the value on the x-axis at which $P(R_1) = .5$.

One account of the psychometric function is provided by SDT (Green & Swets, 1966). On the basis of SDT, there are two separable factors that contribute to the participant’s

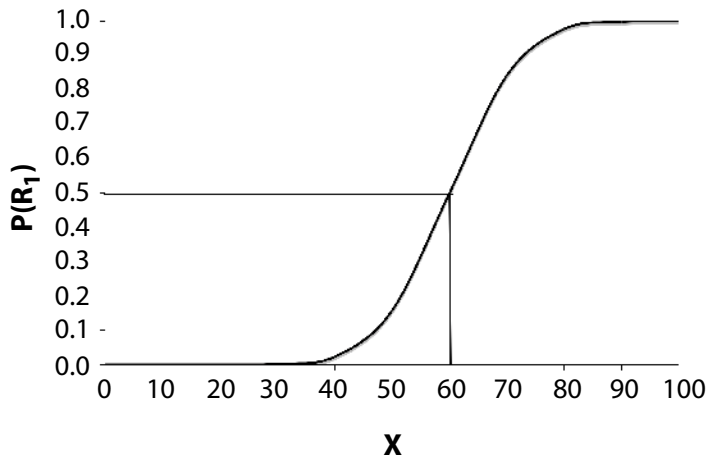


Figure 2. An ogival psychometric function. The probability of one of two responses (e.g., the probability of categorizing a tone as “loud” rather than “soft”) is $P(R_1)$. The physical intensity of the stimulus (e.g., intensity of the tone) is X . The PSE (point of subjective equality) [the value of X at which $P(R_1) = .5$] is 60.

response that the tone was loud or soft. One factor is the participant’s sensitivity to detecting the signal. The second is the participant’s response bias, or criterion.

Signal Detection Theory and Contingency Assessment

The SDT model is presented schematically in Figure 3. Repeated presentations of a constant stimulus value do not result in a constant internal value. Rather, the resulting internal value is variable. The x-axis in Figure 3 is a random variable X representing these internal values (e.g., loudness). The left y-axis shows values of the probability density, $f(X)$, for the distributions generated by each physical value. Figure 3 illustrates the basic version of SDT which assumes that the distribution of internal values generated by a constant physical value is normal, with a mean equal to the physical value and a standard deviation, σ , that is constant across all values. The participant’s task is to place the internal value experienced on each trial into one of two categories, R_1 or R_2 . The participant does so by setting a decision criterion value. If the internal value is larger than this criterion value, the response is R_1 and if the internal value is less than this criterion value, the response is R_2 . The area to the right of the criterion under each distribution represents the probability that the internal value was larger than the criterion. Thus, the value of $P(R_1)$ generated by a particular stimulus value provides an estimate of the proportion of the area under the distribution to the right of the criterion. The dashed function in Figure 3 that plots $P(R_1)$ on the right y-axis is the psychometric function. When the underlying distributions are normal and have a constant variance, the psychometric function is the cumulative normal function (see Killeen, Fet-

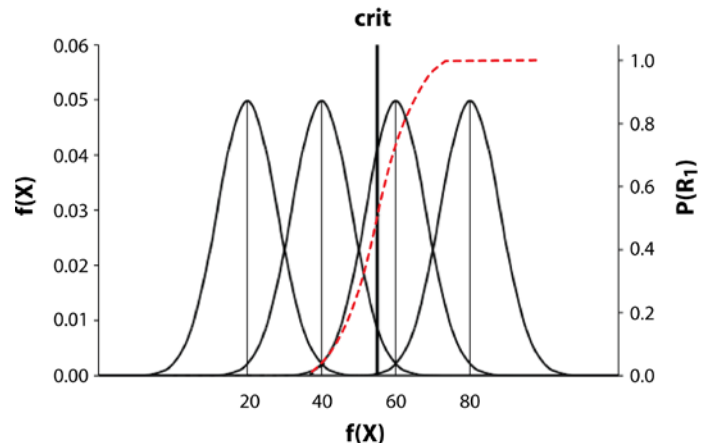


Figure 3. Schematic of the signal detection theory model. Probability density, $f(X)$, is on the left y-axis; probability of a strong response, $P(R_1)$, is on the right y-axis. The dashed line is the psychometric function (From *The psychophysics of contingency assessment*, by L. G. Allan, S. D., Hannah, M. J. C. Crump, and S. Siegel, 2008, *Journal of Experimental Psychology: General*, 137, p. 230. Copyright 2008 by the American Psychological Association. Reprinted with permission).

terman, & Bizo, 1997). The slope of the cumulative normal function is $1/\sigma$ and provides an estimate of the participant’s sensitivity to discriminating among the physical values. The PSE provides an estimate of the value of the criterion.

According to SDT, the participant’s response is determined by the *independent* actions of sensitivity and criterion. Factors that influence the participant’s willingness to say “loud” should affect the placement of that participant’s criterion but leave the slope of the psychometric function unaffected. For example, consider the case of a man who does not want to admit that he is losing his hearing. He would adopt a low criterion value *resulting* in many “loud” responses. The man would say “loud” to many loud tones, a strategy that might deceive the casual observer into thinking that the man’s hearing was intact, as he seemed to detect many actual loud tones. That is, he had many *hits*. The careful observer, however, would note that the man also said “loud” to many soft tones as well—he had many *false alarms*. Given his poor hearing, there would be little change in $P(R_1)$ with changes in the objective intensity of the tone. As a result, the slope of the psychometric function would be shallow (weak sensitivity). When that individual acquires a hearing aid, the slope would be steeper, indicating increased sensitivity as he can now better discriminate among the various tone intensities.

Originally SDT was developed to understand how organisms respond to auditory signals. Subsequently, it has been found to be applicable to many areas in addition to this limited domain. For example, SDT has been applied to medical

diagnoses (reviewed by Swets, 1996), clinical psychological assessment (reviewed by McFall & Treat, 1999), responding by depressed people (reviewed by Allan et al, 2007), and the placebo effect (reviewed by Allan & Siegel, 2002). In all these cases, a participant's judgment is determined both by the participant's sensitivity to the stimuli and by the participant's decision criterion.

We have suggested that the contingency assessment task can be recast as a signal detection task (Allan et al., 2005, 2007, 2008). The participant is exposed to a sequence of cue and outcome presentations, and, at the end of the presentations, must evaluate the strength of the contingency. The participant in a contingency assessment task, like the participant in a standard signal detection experiment, is required to make a judgment under conditions of uncertainty. There will be uncertainty because of "noise" caused by (for example) vagaries in the participant's memory of the events in the series, and shifts in attention during the cue and outcome presentations. The participant's evaluation of the strength of the contingency might be biased by a number of variables (for example, the payoff matrix).

Challenges to applying SDT to contingency assessment

Although the contingency task can conceptually be recast as a signal detection task, actually applying psychophysical procedures to the contingency task presents several challenges. In the typical implementation of the contingency assessment task the participant does not make a categorical response, but rather indicates the strength of the contingent relationship on an analogue scale. The rating response has been interpreted as an index of the participant's perception of the contingency. However, as discussed, this rating measure conflates the participant's sensitivity to the contingency and the participant's response bias, or criterion. To distinguish these two putatively independent variables a categorical response is needed. For example, the task should be one in which the participant makes a binary response: the cue-outcome relationship either is "strong" or "weak."

SDT traditionally uses psychophysical procedures that demand extensive within-subject measures of performance. A second challenge in evaluating a SDT analysis of contingency assessment is the development of such procedures for presenting a participant with a sequence of cues and outcomes. The discrete-trial contingency task is poorly suited for a SDT approach. Many cue-outcome presentations must be presented to the participant in order to ensure that sufficient information is provided about the actual contingency. Depending on the nature of the visuals used to represent cues and outcomes, a series of trials can take many minutes. For example, with presentation times of 3 sec for both the cue and 3 sec for the outcome and a 2 sec inter-pair in-

terval, a block of 40 pairings takes over 5 min. Thus, few ratings can be obtained from a participant during a typical session, greatly limiting the experimenter's ability to make within-participant comparisons. Fortunately, a new version of the contingency task, the streamed-trial procedure, is especially suited to examining contingency assessment with psychophysical procedures (Crump, Hannah, Allan, & Hord, 2007).

The streamed-trial procedure.

With the streamed-trial procedure, it takes seconds, rather than minutes, to define a contingency value. By the rapid sequential presentation of cue-outcome pairs, an entire block of trials is telescoped into a single streamed-trial. The streamed trial used by Crump et al. (2007) is depicted schematically in Figure 4. The cue and the outcome were colored geometric forms. The cue was a blue square and the outcome was a red circle. Each 100-ms presentation consisted of one of four cue-outcome combinations, and presentations were separated by a black screen of 100-ms duration. A stream of these cue-outcome combinations defined the contingency value⁽²⁾.

Psychophysics and the Streamed Trial Procedure

Allan et al. (2008) reported the results of a series of experiments that used the streamed-trial procedure, in conjunction with the method of constant stimuli, to generate psychometric functions under a variety of conditions. Streams with different ΔP values were presented and, at the end of each stream, the participant had to categorize the relationship as "strong" (R_S) or "weak" (R_W). The streamed-trial procedure yielded orderly psychometric functions that were well described by the normal ogive. Allan et al. also showed that several manipulations that have been shown to be bias effects in the psychophysical literature, the payoff matrix being one, had similar biasing effects in the streamed-trial contingency task. Moreover, as in the psychophysical literature, these manipulations had little effect on the participant's sensitivity to the contingencies.

Cue Interaction and Contingency Assessment

As discussed earlier, contingency assessment research was reinvigorated in the mid-1980s by suggestions that simple associative models could be applied to the contingency assessment situation. A great strength of such associative models is that they can explain cue interaction effects, such as blocking, seen with simple conditioning preparations. In a short time these cue-interaction effects were reported by contingency assessment researchers. That is, participants minimize the contingency between a target cue (C_T) and outcome if the C_T is presented in compound with a companion cue (C_C) that is highly predictive of the outcome. These findings seemed to establish the utility of an associative in-

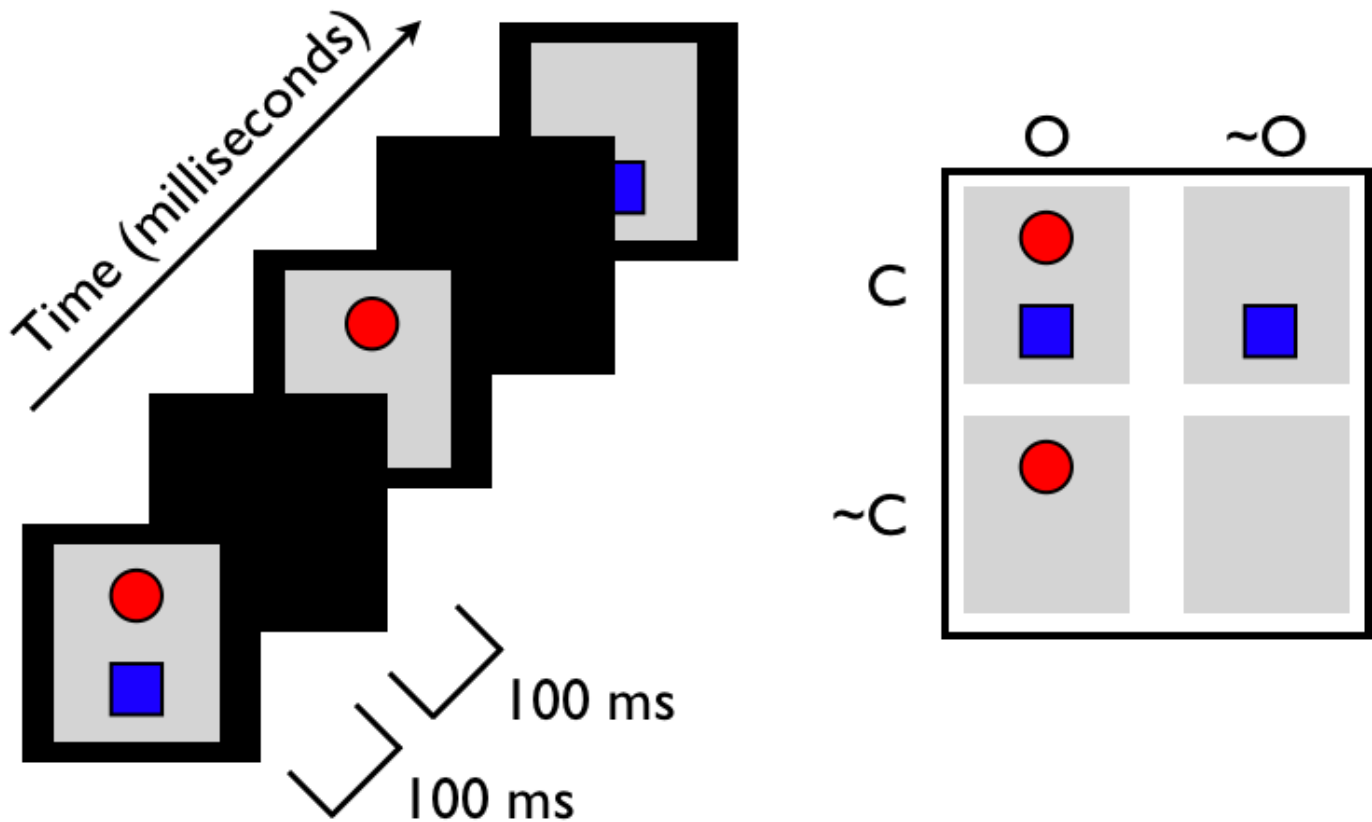


Figure 4. On the left is a schematic illustrating the structure of a streamed trial in Crump et al. (2007). On the right are the four possible cue-outcome combinations in a streamed trial. Squares are cues and were presented in blue. Circles are outcomes and were presented in red (From *The psychophysics of contingency assessment*, by L. G. Allan, S. D., Hannah, M. J. C. Crump, and S. Siegel, 2008, *Journal of Experimental Psychology: General*, 137, p. 228. Copyright 2008 by the American Psychological Association. Reprinted with permission).

terpretation of contingency assessment. On the basis of most associative theories, blocking indicates that the participant learns little about a blocked C_T ; thus, the blocking seen in contingency assessment typically has been attributed to the participant learning little about the C_T -outcome relationship (e.g., Allan & Tangen, 2005). Inasmuch as we are proposing that contingency assessment can be best understood from the perspective of SDT, it behooves us to demonstrate why an analysis of cue interaction effects in contingency assessment that incorporates SDT is better than one that does not. That is, perhaps the simultaneous presence of a C_C that is highly predictive of the outcome does not attenuate learning about C_T , but rather alters the participant's criterion for responding.

We have already reported that cue interaction effects can be demonstrated with the streamed trial procedure. Hannah, Crump, Allan, and Siegel (2009) adapted the streamed-trial procedure for both the one-phase blocking paradigm and the two-phase blocking paradigm. They showed that the streamed-trial procedure produced conventional cue interaction effects with both paradigms. In the Hannah et al. experi-

ments, as in most other cue-interaction experiments, the participants' judgments of contingency were measured with the typical rating response. However, the rating response used by Hannah et al. (and almost everyone else) does not separate the individual contributions of the participant's sensitivity to the cue-outcome relationship, and his or her response bias. We previously suggested that the streamed-trial procedure is especially suited to examining the potentially separable contributions of sensitivity and bias to cue interaction effects, and we actually described the design of an experiment that would accomplish this goal, using the streamed-trial version of the one-phase blocking paradigm (Allan et al., 2008). Nobody accepted our challenge to use this procedure to evaluate a signal detection analysis of cue interaction, so we did the study and report the results here.

One-Phase Blocking Experiment

In the one-phase blocking paradigm, two cues, C_T and C_C , are paired with a common outcome across trials. The two cues result in four possible cue combinations: both the target and the companion cues may be present ($C_T C_C$), both the

target and the companion cues may be absent ($\sim C_T \sim C_C$), the target cue may be present and the companion cue absent ($C_T \sim C_C$), or the target cue may be absent and the companion cue present ($\sim C_T C_C$). For each cue combination, the outcome either occurs (O) or does not occur ($\sim O$), resulting in eight possible cue-outcome combinations, as is depicted in Table 2. Of primary interest is the effect of companion cue contingency on the participant's assessment of the target cue contingency. The usual finding is that ratings of C_T depend on the contingency between C_C and the outcome (e.g., Baker et al., 1993; Hannah et al., 2009; Spellman, 1996; Tangen & Allan, 2003, 2004). Hannah et al., for example, showed that for a fixed contingency of 0.5 between C_T and the outcome, ratings of C_T were lower when the contingency between C_C and the outcome was perfect ($\Delta P = 1.0$) than when there was no contingency between C_C and the outcome ($\Delta P = 0.0$). Thus, the simultaneous presence of a C_C that was highly predictive of the outcome blocked the apprehension of the C_T -outcome contingency. The present experiment was designed to evaluate the separable contributions of contingency sensitivity and response bias to such cue interaction.

Table 2. The 4x2 matrix for cue and outcome presentations in the one-phase blocking paradigm. The letters in the cells represent the joint frequency of occurrence of the eight cue-outcome combinations.

	O	$\sim O$
$C_T C_C$	a	b
$C_T \sim C_C$	c	d
$\sim C_T C_C$	e	f
$\sim C_T \sim C_C$	g	h

Method

Participants, stimuli and apparatus

As discussed elsewhere (e.g., Wickens, 2002, p. 78), signal-detection measures are best obtained by collecting many observations from a small number of participants (rather than a few observations from a large number of participants). Moreover, participants in signal-detection experiments often are experienced with the task requirements (in the present case, categorizing a contingent relationship presented in a streamed trial either as "strong" or "weak"). There were six participants in this experiment (by first and last initial: XG, MC, AS, AB, SA, & CH), all members of the McMaster University community, who were paid \$10.00 per session. Four had prior experience with the streamed trial procedure. A session was completed in about 45 min. Participants were

restricted to a maximum of two sessions per day with at least an hour break between sessions.

Figure 5 shows the streamed-trial procedure, as adapted by Hannah et al. (2009) for the one-phase blocking paradigm. Cues were blue squares and blue triangles and the outcome was a red circle. The eight possible cue-outcome pairs are depicted in Figure 5. A cue-outcome pair was presented for 100 ms, and cue-outcome frames were separated by a black 100-ms frame. The location of each shape, when it was present, was constant across all frames: the red circle was centered above the blue shapes, the blue square was on the left, and the blue triangle was on the right. The geometric forms were presented on a grey background (8.8 cm in height and 7.0 cm in width). The blue square measured 2.1 cm in height and width; the blue triangle measured 2.7 cm at its base and extended 2.3 cm in height; the red circle measured 2.5 cm in diameter.

Procedure

The participants were told that their task was to categorize the strength of association between a blue shape (cue) and a red circle (outcome) as weak or strong. Streams were composed so that one blue shape, C_T , had (on different streams) a contingency of either 0.2, 0.4, 0.6, or 0.8 with the circle, while its C_C had either a contingency of 1.0 (*companion₁ stream*) or 0.0 (*companion₀ stream*) with the circle.

Table 3. The eight contingency matrices in the experiment.

$C_C \Delta P = 0$	$C_T \Delta P$							
	0.2		0.4		0.6		0.8	
	O	$\sim O$	O	$\sim O$	O	$\sim O$	O	$\sim O$
$C_T C_C$	8	6	9	5	10	4	10	2
$C_T \sim C_C$	4	2	5	1	6	0	8	0
$\sim C_T C_C$	2	4	1	5	0	6	0	8
$\sim C_T \sim C_C$	6	8	5	9	4	10	2	10

$C_C \Delta P = 1$	$C_T \Delta P$							
	0.2		0.4		0.6		0.8	
	O	$\sim O$	O	$\sim O$	O	$\sim O$	O	$\sim O$
$C_T C_C$	12	0	14	0	16	0	18	0
$C_T \sim C_C$	0	8	0	6	0	4	0	2
$\sim C_T C_C$	8	0	6	0	4	0	2	0
$\sim C_T \sim C_C$	0	12	0	14	0	16	0	18

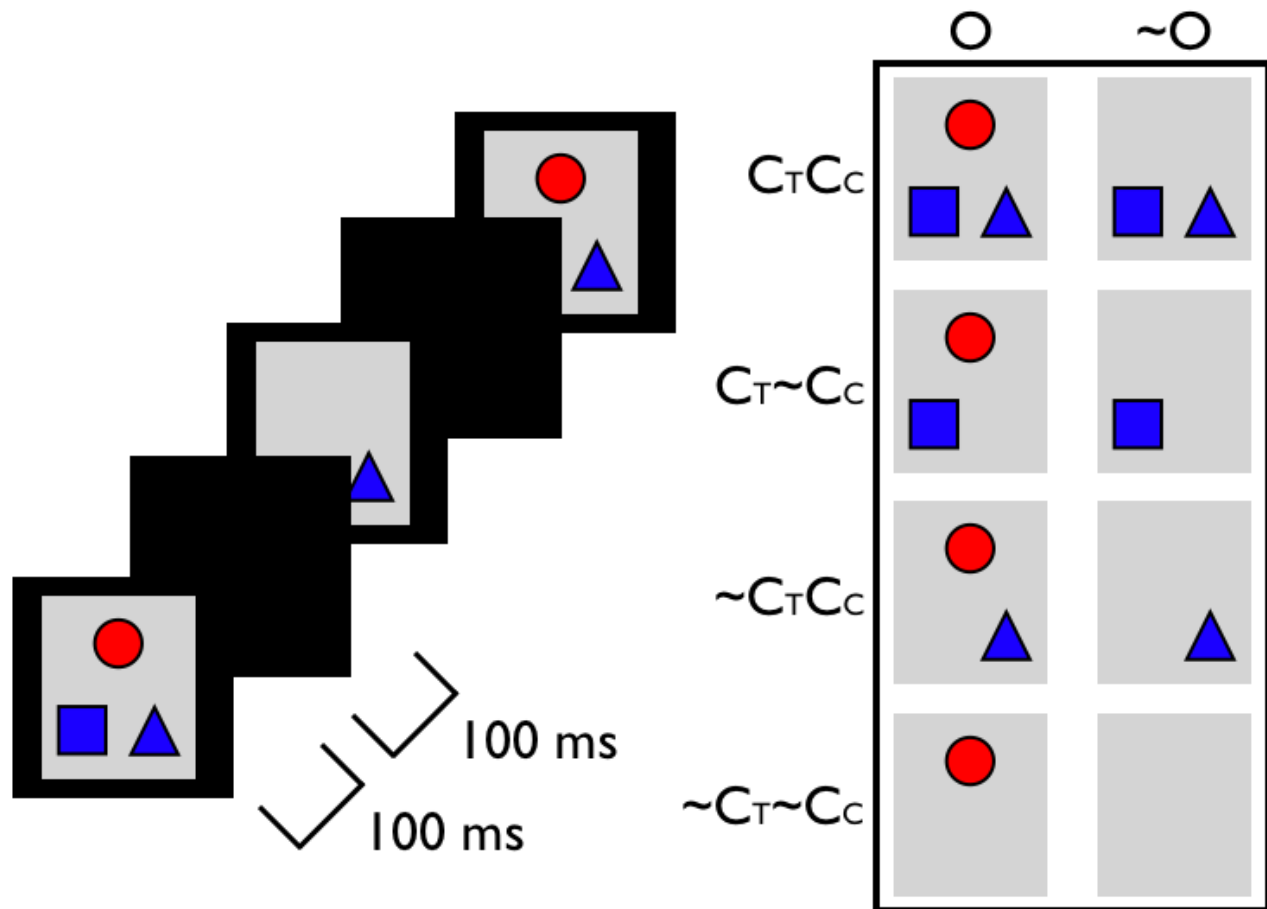


Figure 5. On the left is a schematic illustrating the structure of a streamed trial for the one-phase blocking paradigm. On the right are the eight possible cue-outcome combinations in a streamed trial. Squares and triangles are cues and were presented in blue. Circles are outcomes and were presented in red (From *The psychophysics of contingency assessment*, by L. G. Allan, S. D., Hannah, M. J. C. Crump, and S. Siegel, 2008, *Journal of Experimental Psychology: General*, 137, p. 242. Copyright 2008 by the American Psychological Association. Reprinted with permission).

The two values of C_C contingency (0.0, and 1.0) were crossed with the four values of C_T contingency (0.2, 0.4, 0.6, 0.8), resulting in eight 4 x 2 contingency matrices which are available in Table 3. Each streamed trial consisted of a sequential display of randomly ordered presentations of the eight cue-outcome combinations defined by one of the contingency matrices. The duration of a streamed trial was approximately 8 seconds.

An experimental session consisted of five blocks of 48 streamed-trials. Each of the eight contingency matrices occurred six times in a block in a randomly determined order. At the end of each stream, the participant was required to make a binary response about the relationship between *one* of the cues and the outcome by clicking one of two buttons (“weak” or “strong”) on the computer monitor. The participant was informed which cue-outcome relation to judge by a small picture appearing on the screen at the end of the stream showing one of the two cues and the outcome. For each of the eight contingency matrices, C_T was probed at the

end of half the streams and C_C was probed on the remaining streams. Each participant completed 15 sessions.

Results and Discussion

$P(R_S)$ to the probed cue was determined for each of the eight contingency matrices. $P(R_S)$ on C_T -probed streams is shown in Figure 6 for each of the six participants. $P(R_S)$ is plotted as a function of target ΔP separately for the two C_C contingencies. $P(R_S)$ is higher for companion₀ streams than for companion₁ streams, indicating that the response to a fixed value of ΔP depends on the C_C contingency. Psychometric functions were fit to each participant’s data using ProFit⁽³⁾, and the parameter values for each participant are available in Table 4A. Also available in Table 4A is the proportion of the variability in $P(R_S)$ accounted for by the fits, R_2 . For every participant, the PSE was smaller for companion₀ streams than for companion₁ streams. That is, the value of ΔP at which $P(R_S) = .5$ was less when the companion ΔP

= 0.0 (mean PSE = .29) than when the companion $\Delta P = 1.0$ (mean PSE = .79). Except for participant AS, the values of σ did not differ much for the two streams, and the direction of the difference varied among the participants – for some participants (AS, AB, and SA) σ (the reciprocal of the slope) was larger for companion₀ streams, for some (XG, CH) it was smaller, and for one participant (MC) it was the same. Table 4B presents the parameter values for psychometric functions that were constrained to have the same value of σ , and these functions are shown in Figure 6. A comparison of the R^2 values in Tables 4A and 4B indicates that, except for participant AS, the same-slope constraint made little difference to the goodness-of-fit. Overall, the data are well de-

scribed by psychometric functions that have the same slope and differ only in PSE⁽⁴⁾.

As discussed earlier, the slope of the psychometric function provides an estimate of the participant's sensitivity to discriminating among the ΔP values, and the PSE provides an estimate of the criterion. The same-sloped functions in Figure 6 suggest that the ability to discriminate among the C_T contingencies is not affected by C_C contingency. Rather the effect of C_C is on the PSE – the placement of the criterion. The location of the criterion regarding the strength of the C_T contingency depends on the value of the C_C contingency in the stream. When C_C is a good predictor of the outcome

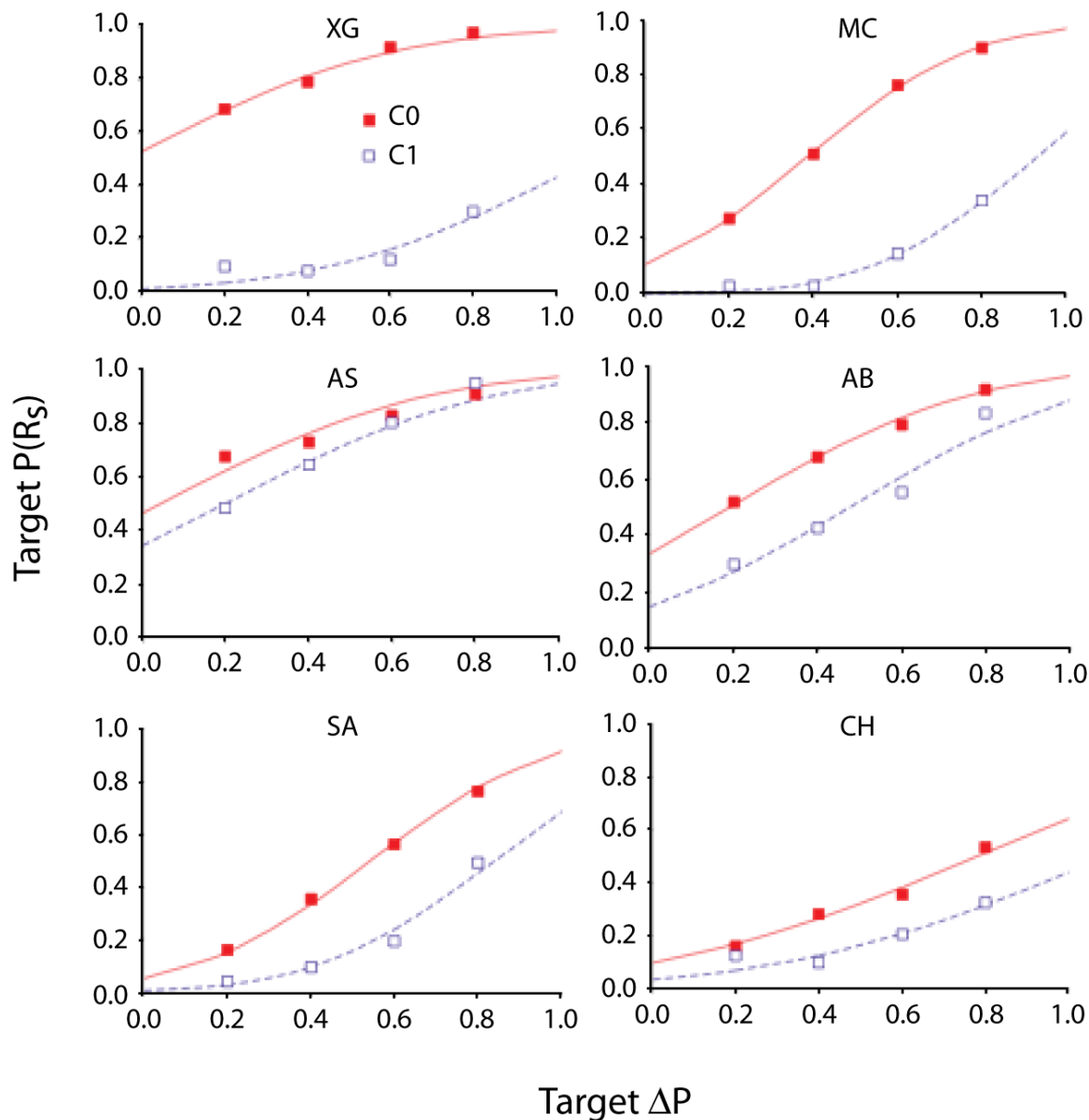


Figure 6. $P(R_S)$ to C_T as a function of target ΔP for each companion ΔP . The symbols are the data, and the lines are the best fitting same-slope Gaussian functions. The filled symbols are the data on companion₀ streams and the unfilled symbols are the data on companion₁ streams.

Table 4A. σ , PSE, and R^2 values for each participant.

Participant	σ		PSE		R^2
	$C_C \Delta P = 0$	$C_C \Delta P = 1$	$C_C \Delta P = 0$	$C_C \Delta P = 1$	
XG	.47	.55	-.02	1.13	.9930
MC	.31	.31	.38	.93	.9992
AS	.74	.41	-.11	.24	.9969
AB	.48	.43	.18	.48	.9912
SA	.36	.28	.55	.82	.9977
CH	.58	.66	.77	1.13	.9959
Mean	.49	.44	.29	.79	.9956

Table 4B. σ , PSE, and R^2 values for each participant with the same-slope constraint.

Participant	σ	PSE		R^2
		$C_C \Delta P = 0$	$C_C \Delta P = 1$	
XG	.49	-.04	1.08	.9927
MC	.31	.38	.92	.9992
AS	.50	.04	.20	.9887
AB	.45	.19	.48	.9907
SA	.34	.54	.84	.9956
CH	.60	.78	1.09	.9956
Mean	.45	.32	.77	.9938

($\Delta P = 1.0$), the participant was likely to indicate that the relationship between the C_T and the outcome was weak. In contrast, when C_C was a poor predictor of the outcome ($\Delta P = 0.0$), the participant was likely to indicate that the relationship between C_T and the outcome was strong. The results of this experiment suggest that this cue-interaction effect resides in the decision process.

Contemporary theories of associative learning (starting with the Rescorla-Wagner model) were developed primarily to explain cue-interaction effects, such as blocking. On the basis of most of these models, cue interaction occurs because the participant does not learn much about C_T when C_C is a good predictor of the outcome. In applying associative theories to contingency assessment, such reasoning led researchers to conclude that participants are insensitive to the C_T -outcome relationship when C_T is presented in compound with a C_C that is in a highly contingent relationship with the outcome (e.g., Allan & Tangen, 2005). In contrast, the present results suggest that the participant knows as much about the C_T -outcome relationship when C_C is a good predictor of the outcome as when C_C is a poor predictor of the outcome.

Cue interaction occurs because of the participant's placement of his or her response criterion, rather than because of the participant's ability to learn about the C_T -outcome relationship.

These results (and others that we have previously reported, Allan et al., 2005, 2008) indicate that contingency assessment cannot really be understood merely by asking the participant to rate the perceived magnitude of a contingency. How does the participant's sensitivity to the contingency and the participant's response bias (where he or she places the response criterion) combine to yield such an omnibus judgment? Signal detection theory provides the theoretical basis and methodology to dissect the separable contributions of sensitivity and bias to contingency assessment.

Associative Learning and Signal Detection Theory

As summarized earlier, researchers experienced in studying basic associative processes in non-human animals became interested in studying contingency assessment by humans about 25 years ago. These researchers demonstrated that many phenomena seen in the animal learning laboratory (especially cue competition effects) were also seen in human contingency assessment. Incorporation of models of associative learning reinvigorated and enriched the study of contingency assessment. Contingency assessment researchers owe a debt to these learning researchers. We would like to repay the debt. SDT, which enhances our understanding of contingency assessment, also enhances our understanding of associative learning. That is, associative theories in general, and associative accounts of contingency assessment in particular, would benefit from incorporating concepts of SDT⁽⁵⁾.

A basic assumption of most theories of learning is that changes in conditional responding are manifestations of the strength of the CS-US association. We suggest otherwise. A subject's performance in a Pavlovian conditioning experiment is a conflated measure of its sensitivity to the CS-US relationship and its criterion for displaying an overt CR, just as a participant's judgment of the relationship between the cue and outcome is a conflated measure of the participant's sensitivity to the programmed contingency and the participant's response criterion.

The sensitivity-criterion distinction in SDT is reminiscent of the long standing learning-performance distinction in the learning literature. As recently discussed by Miller (2006), most learning theories do not explain observed behavior. Rather, they explain how an intervening variable changes with practice.

"Learning is an intervening variable; all we ever see is a change in behavior as a consequence of prior experi-

ence. Consistent with the misguided name learning theory and inconsistent with the actual goal of explaining acquired behavior, most modern associative theories in the animal tradition emphasize the learning (i.e., acquisition) process per se and are virtually silent concerning the transformation of acquired information into behavior. For example, Rescorla and Wagner (1972) simply say that responding is monotonically related to associative strength” (Miller, 2006, p. 82).

Although the Rescorla-Wagner (and similar) models do not have rules concerning the function relating covert associative strength into overt performance (other than a monotonic transformation), they have been successful in accounting for many findings because they use a competitive learning algorithm. However, results indicating that performance is determined both by sensitivity to the conditions of learning, and by response bias, suggest that a noncompetitive learning algorithm may be more suited to describe the changes that occur in associative learning.

Competitive and Noncompetitive Learning Algorithms

On the basis of a competitive learning algorithm, the more learned about one cue the less learned about a simultaneously presented cue. Consider the example of blocking: the more learned about C_C , the less learned about C_T . In the Rescorla-Wagner model, the predictive strength of a cue will change each time it is presented according to the competitive algorithm

$$\Delta V = \alpha\beta(\lambda - \Sigma V), \quad (2)$$

where ΔV is the change in predictive strength of the presented cue, α and β are learning rate parameters that represent the associability of the cue and of the outcome respectively, λ represents the maximum amount of predictive strength supported by the outcome ($\lambda > 0$ for a presented outcome, and $\lambda = 0$ for an omitted outcome), and ΣV is the composite predictive strength of *all* presented cues. That is, when both cues are presented ($C_T C_C$),

$$\Sigma V = VC_T + VC_C, \quad (3)$$

where VC_T is the predictive strength of the target cue and VC_C is the predictive strength of the companion cue. When C_C is in a highly contingent relationship with the outcome, VC_C will be high, the increments in predictive strength of C_T (ΔV_T) will be small, and the companion cue “wins” the competition with the target cue.

In contrast, our results suggest that there is no “competition” between learning about C_C and C_T in a cue-interaction experiment. Rather, participants learn about both. In the presence of a C_C that is highly predictive of the outcome

(i.e., the companion₁ stream), participants shift their response criterion for reporting the relationship between C_T and the outcome in one direction, and in presence of a C_C that is not highly predictive of the outcome (i.e., the companion₀ stream), participants shift their response criterion in the other direction.

Since our psychophysical procedures reveal that there is no competitive learning about simultaneously-presented companion and target cues, we suggest that that the acquisition of associative strength is best understood with a noncompetitive learning algorithm. Thus, we agree with Miller and colleagues (e.g., Miller & Matzel, 1988; Stout & Miller, 2007) in the formulation of the learning process in their comparator hypothesis. The comparator hypothesis, too, invokes a noncompetitive learning algorithm as the acquisition process, and places cue interaction in the decision process.

One noncompetitive learning algorithm is the Bush-Mosteller rule (Bush & Mosteller, 1955). Both Miller and colleagues (e.g., Miller & Matzel, 1988; Stout & Miller, 2007) and we (Allan et al., 2008) have suggested that the acquisition process is best understood with the Bush-Mosteller rule.⁽⁶⁾ This rule specifies that the predictive strength of a cue will change on each trial that it is presented according to the standard linear operator equation

$$\Delta V = \alpha\beta(\lambda - V). \quad (4)$$

In the Bush-Mosteller Equation 4, the predictive strength of a presented cue is changed by the predictive strength of only that cue, whereas in the Rescorla-Wagner Equation 2 it is changed by the predictive strength of *all* presented cues. In Appendix A, we derive the asymptotic predictive strength of C_T and C_C in the one-phase blocking paradigm. The asymptotic predictive strength of the target cue is:

$$VC_T = P(O|C_T), \quad (5)$$

and the asymptotic predictive strength of the companion cue is:

$$VC_C = P(O|C_C), \quad (6)$$

Each streamed-trial in the one-phase blocking paradigm results in two values of associative strength, VC_T and VC_C . In current conceptualizations of associative learning rules, for fixed stimulus values there is no variability in the value of V . Within the SDT framework, V would be embedded in a noisy background. Thus, for fixed stimulus values, there would be variability in V .

Blocking and SDT: Decisions, Decisions!

Figure 7 presents the SDT representation for the one-phase blocking paradigm. The x-axis in Figure 7 is $V = P(O|C)$. The six distributions are generated by the six ΔP values used in the experiment. The two outer distributions (dashed lines) are the distributions for VC_C generated by the two values of the companion cue, and the four other distributions (filled lines) are the distributions for VC_T generated by the four values of the target cue. The mean of each distribution is $P(O|C)$. At the end of each stream, the participant has two internal values, VC_T and VC_C , and the response is determined by a comparison of the probed internal value to the criterion.

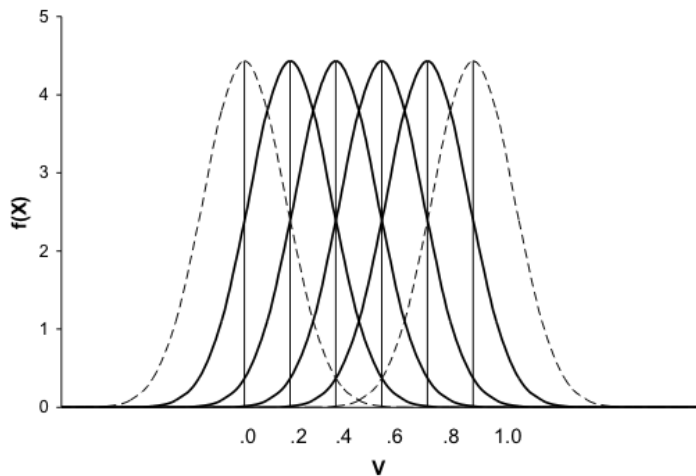


Figure 7. Schematic of the signal detection theory model for the one-phase blocking paradigm. The x-axis is V (the strength of the association between the cue and outcome, which is equal to the probability of the outcome in the presence of the cue). The six distributions are generated by the six ΔP values used in the experiment. The two outer distributions (dashed lines) represent the distributions for VC_C (the strength of the association between the companion cue and the outcome) generated by the two ΔP values for the companion cue. The other four distributions (filled lines) represent the distributions for VC_T (strength of the association between the target cue and the outcome) generated by the four ΔP values of the target cue.

Our data suggest that the location of the criterion is influenced by the value of V generated by the non-probed cue. If the participant is queried about C_T , the criterion is pulled towards VC_C , and if the participant is queried about C_C , the criterion is pulled towards VC_T . Consider first streams on which the participant is asked about the relationship between the target cue and the outcome, i.e., C_T -probed streams (Figure 8a). If the companion contingency is high (companion₁ streams), VC_C would tend to be large and the criterion would be pulled to the right (indicated by C1 in Figure 8a) resulting

in a decrease in $P(R_S)$ for the target. If the companion contingency is low (companion₀ streams), VC_C would tend to be small, and the criterion would be pulled to the left (indicated by C0 in Figure 8a) resulting in an increase in $P(R_S)$ for the target. Now consider streams on which the participant is asked about the relationship between the companion cue and the outcome, i.e., C_C -probed streams (Figure 8b). On these streams, the location of the criterion would be influenced by VC_T . Streams with the larger ΔP values (.6 and .8) would tend to produce large VC_T values, resulting in the criterion being pulled to the right. Streams with the smaller ΔP (.2 and .4) values would tend to produce small VC_T values, resulting in the criterion being pulled to the left.

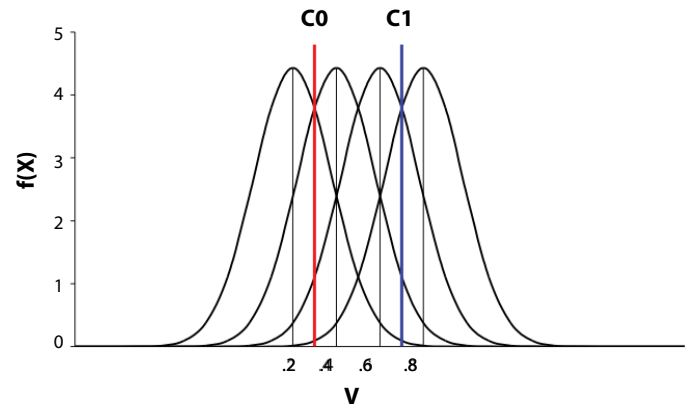


Figure 8a. Schematic of the signal detection theory model on C_T -probed streams (i.e., the participant is asked the judge the relationship between the target cue and the outcome). The location of the criterion is to the right on companion_{1,0} streams (a programmed contingency of 1 between the companion cue and the outcome) and is to the left on companion_{0,0} streams (a programmed contingency of 0 between the companion cue and the outcome).

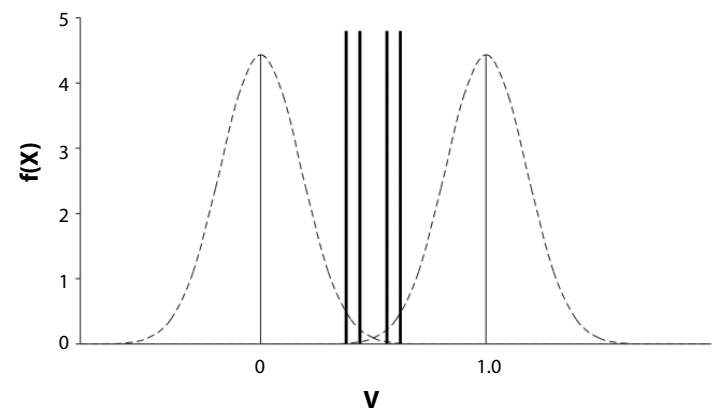


Figure 8b. Schematic of the signal detection theory model on C_C -probed streams (i.e., the participant is asked the judge the relationship between the companion cue and the outcome). There are four locations for the criterion, moving from left to right as ΔP of the companion cue increases from .2 to .8.

The criterion-setting proposal outlined above provides a qualitative description of the data shown in Figures 6. Moreover, it is consistent with criterion-setting accounts in the literature for data generated in other tasks. For example, Treisman (1984) argued that “a criterion is defined not only for a particular judgment, but also for particular conditions under which this judgment may be made. . . . Thus, the decision criterion may have different values for different sets of circumstances.” (pp. 132-133). Treisman discusses the application of his criterion-setting model (Treisman & Williams, 1984) to diverse phenomena in the literature such as the effect of

nontemporal variables on judgments of temporal intervals. It is well established that the marker of a temporal interval usually influences the participant’s judgment (see Allan, 1979 for a review) – e.g., high frequency tones are judged longer than low frequency tones, loud tones are judged longer than soft tones, large visual markers are judged longer than small visual markers. Treisman argues that the nontemporal attribute of the temporal marker biases the location of the criterion for the temporal judgment. Similarly, we suggest that the value of the non-probed cue biases the location of the criterion for the probed cue.

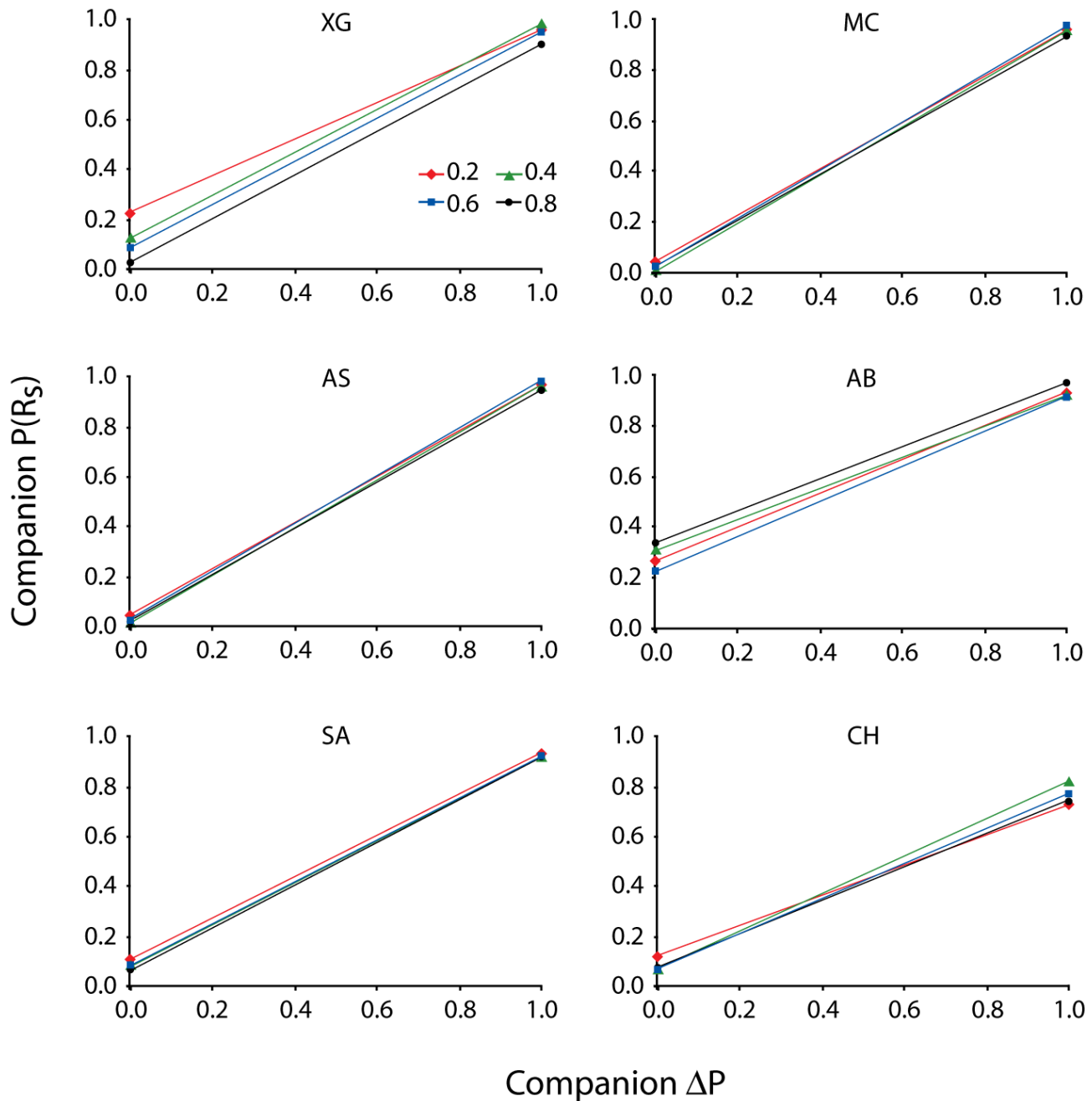


Figure 9. $P(R_s)$ to C_c as a function of companion ΔP for each target ΔP (i.e., the probability of the participant judging the contingency between the companion cue and the outcome as “strong,” rather than “weak,” as a function of the programmed contingency between the target cue and the outcome).

Further Research

The criterion-setting proposal for one-phase blocking can be evaluated by further research. We describe two such areas of research: the effect of target cue contingency on the assessment of the companion cue, and attenuation of blocking by manipulation of the payoff matrix.

Effect of target cue contingency on the assessment of the companion cue.

In our one-phase blocking experiment we noted the effect of C_C contingency on the assessed relationship between C_T and the outcome (Figure 6). In addition, we evaluated the effect of C_T contingency on the assessed relationship between C_C and the outcome. $P(R_S)$ on C_C -probed streams is shown in Figure 9 for each of the six participants. $P(R_S)$ is plotted as a function of companion ΔP for each of the four C_T contingencies. $P(R_S)$ is low when the companion contingency is 0.0 and is high when the companion contingency is 1.0. The psychometric functions are similar for the four target ΔP values. C_T contingency had little effect on either the slope or the PSE of these functions. Thus while companion cue contingency influences performance on C_T -probed streams (Figure 6), target cue contingency does not influence performance on C_C -probed streams (Figure 9). This asymmetry has previously been reported (Tangen & Allan, 2004; Hannah et al., 2009).

We suggest that the absence of an effect of C_T contingency on C_C judgments is due to the minimal overlap of the two C_C distributions (Figure 8b). If this is the case, then changing the C_C contingencies, say to .2 and .8, would increase the overlap of the underlying distributions and should result in an effect of C_T contingency on C_C judgments.

Payoffs and the cue interaction effect.

Other future research follows from the finding reported in Allan et al. (2008) that the payoff schedule influences the placement of the criterion in the single-cue task. They showed that monetary reward for a particular response (R_S or R_W) affected the location of the criterion. The location of the criterion in the one-phase blocking paradigm should also be influenced by manipulations of the payoff schedule. For example, a payoff schedule that reinforces participants for maintaining a single criterion on C_T -probed streams that is independent of the value of C_C should abolish the cue-interaction effect.

Concluding Comments

When asked to judge whether a tone is loud or soft, people

respond both on the basis of their ability to detect the loudness, and on other factors that bias them to characterize the tone as either loud or soft (Green & Swets, 1966). When asked to judge whether a drug is effective or ineffective, people respond both on the basis of their ability to detect their symptomatic relief, and on other factors that bias them to characterize the treatment as either effective or ineffective (Allan & Siegel, 2002). Similarly, in the contingency assessment task described here, participants judge whether a cue-outcome relationship is strong or weak, and they respond on the basis of their ability to detect the strength of the contingency, and on other factors that bias them to characterize the relationship as strong or weak.

A particular psychophysical model, SDT, may usefully be applied to contingency assessment. This approach provides a framework for separating effects on performance resulting from changes in sensitivity to the contingency between the cue and the outcome from effects due to the participant's bias for making a particular response. We previously demonstrated that several manipulations that affect contingency assessment do so by acting on the response criterion (Allan et al., 2005; Allan et al., 2008). Additionally, we demonstrate here that another manipulation, cue interaction, similarly affects the decision process – not the sensitivity to relationship between the blocked cue and the outcome.

About 25 years ago, researchers experienced in studying basic associative processes in non-human animals became interested in studying contingency assessment by humans. These researchers demonstrated that many phenomena seen in the animal learning laboratory (especially cue competition effects) were also seen in human contingency assessment, and associative models were applied to understanding contingency assessment. We argue that these models, whether applied to understanding learning by non-human animals or contingency assessment by humans, should be modified to incorporate a response criterion process. Signal detection theory provides such a process.

References

- Allan, L. G. (1979). The perception of time. *Perception & Psychophysics*, **26**, 340-354.
- 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., & Siegel, S. (2002). A signal detection analysis of the placebo effect. *Evaluation & the Health Professions*, **25**, 410-420
- Allan, L. G., & Tangen, J. M. (2005). Judging relationships

- between events: How do we do it? *Canadian Journal of Experimental Psychology*, **59**, 22-27.
- Allan, L. G., Siegel, S., & Tangen, J. M. (2005). A signal detection analysis of contingency data. *Learning & Behavior*, **33**, 250-263.
- Allan, L. G., Siegel, S., & Hannah, S. (2007). The sad truth about depressive realism. *Quarterly Journal of Experimental Psychology*, **60**, 482-495.
- Allan, L. G., Hannah, S. D., Crump, M. J. C., & Siegel, S. (2008). The psychophysics of contingency assessment. *Journal of Experimental Psychology: General*, **137**, 226-243.
- Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, **108**, 41-485.
- 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., Mercier, P., Valle-Tourangeau, F., Frank, R., & Pan, D. W. (1993). Selective associations and causality judgments: Presence of a strong causal factor may reduce judgments of a weaker one. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **19**, 414-432.
- Boneau, C. A., & Cole, J. L. (1967). Decision theory, the pigeon, and the psychophysical function. *Psychological Review*, **74**, 123-135.
- Boynton, D. M., Smith, L. D., & Stubbs, D. A. (1997). Sensitivity and bias in covariation detection: A direct approach to a tangled issue. *Organization Behavior and Human Decision Processes*, **72**, 79-98.
- Bush, R. R., & Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, **58**, 313-323.
- Chapman, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **17**, 837-854.
- Crump, M. J. C., Hannah, S. D., Allan, L. G., & Hord, L. K. (2007). Contingency judgments on the fly. *Quarterly Journal of Experimental Psychology*, **60**, 753-761.
- Dickinson, A., Shanks, D. R., & Evenden, J. (1984). Judgment of act-outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology*, **36A**, 29-50.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: John Wiley and Sons, Inc.
- Hannah, S. D., Crump, M. J. C., Allan, L. G., & Siegel, S. (2009). Cue-interaction effects in contingency judgments using the streamed-trial procedure. *Canadian Journal of Experimental Psychology*, **63**, 103-112.
- Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the production of behavior: Aversive stimulation* (pp. 9-33). Florida: University of Miami Press.
- Kamin, L. J. (1969a). Selective association and conditioning. In N. J. Mackintosh and W. K. Honig (Eds.), *Fundamental issues in associative learning*. Halifax: Dalhousie University Press.
- Kamin, L. J. (1969b). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279-296). New York: Appleton-Century-Crofts.
- Killeen, P. R., Fetterman, J. G., & Bizo, L. A. (1997). Time's causes. In C. M. Bradshaw & E. Szabadi (Eds.), *Time and behaviour: Psychological and neurological analyses* (pp. 79-131). Elsevier Science.
- Mason, C. R., Idrobo, F., Early, S. J., Abibi, A., Zheng, L., Harrison, J. M., & Carney, L. H. (2003). CS-dependent response probability in an auditory masked-detection task: Considerations based on models of Pavlovian conditioning. *Quarterly Journal of Experimental Psychology*, **56B**, 193-206.
- McFall, R. M., & Treat, T. A. (1999). Quantifying the information value of clinical assessments with signal detection theory. *Annual Review of Psychology*, 1999, **50**, 215-241.
- Miller, R. R. (2006). Challenges Facing Contemporary Associative Approaches to Acquired Behavior. *Comparative Cognition & Behavior Reviews*, **1**, 77-93.
- 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.
- Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological Bulletin*, **117**, 363-386.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, **87**, 532-552.
- Perales, J. C., Catena, A., Shanks, D. R., and José A. González, J. A. (2005). Dissociation between judgments and outcome-expectancy measures in covariation learning: A Signal Detection Theory approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **31**, 1105-1120.
- Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin*, **68**, 29-46.
- Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review*, **74**, 71-80.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp 64-99). New York: Appleton-Century-Crofts.

- Schmajuk, N. A. (1987). Classical conditioning, signal detection, and evolution. *Behavioural Processes*, **14**, 277-289.
- Shaklee, H., & Mims, M. (1981). Developmental rule use in judgments of covariation between events. *Child Development*, **52**, 317-325.
- Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology*, **60**, 291-309.
- Siegel, S., & Allan, L. G. (1996). The widespread influence of the Rescorla-Wagner model. *Psychonomic Bulletin and Review*, **3**, 314-321.
- Smedslund, J. (1963). The concept of correlation in adults. *Scandinavian Journal of Psychology*, **4**, 165-173.
- Spellman, B. A. (1996). Conditionalizing causality. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation: Vol. 34. Causal Learning* (pp. 167-200). San Diego Ca: Academic Press.
- Stout, S. C., & Miller, R. R. (2007). Sometimes-Competing Retrieval (SOCR): A Formalization of the Comparator Hypothesis. *Psychological Review*, **114**, 759-783.
- Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers*. New Jersey: Lawrence Erlbaum Associates.
- Tangen, J. M., & Allan, L. G. (2004). Cue-interaction and judgments of causality: Contributions of causal and associative processes. *Memory & Cognition*, **32**, 107-124.
- Tangen, J. M., & Allan, L. G. (2003). The relative effect of cue interaction. *Quarterly Journal of Experimental Psychology*, **56B**, 279-300.
- Treisman, M. (1984). Contingent aftereffects and situationally coded criteria: Discussion paper. In J. Gibbon & L. G. Allan (Eds.), *Timing and Time Perception*. Annals of the New York Academy of Sciences.
- Treisman, M., & Williams, T. C. (1984). A theory of criterion setting with an application to sequential dependencies. *Psychological Review*, **91**, 68-111.
- Wagner, A. R., & Rescorla, R. A. (1972). Inhibition in Pavlovian conditioning: Application to a theory. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 301-336). London: Academic 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.
- Waldmann, M. R. (2001). Predictive versus diagnostic causal learning: Evidence from an overshadowing paradigm. *Psychonomic Bulletin & Review*, **8**, 600-608.
- 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. (1990). Attribution of causality to common and distinctive elements of compound stimuli. *Psychological Science*, **1**, 298-302.
- Wickens, T. D. (2002). *Elementary signal detection theory*. New York: Oxford University Press.

Footnotes

- 1 The data depicted in Figure 1 were obtained by searching the PsycINFO database with the following query: (PO=human and (AB=(cue interaction or forward blocking or backward blocking) or TI=(cue interaction or forward blocking or backward blocking))) or (PO=human and (AB=(contingency assessment* or assessing contingency* or assessment* of contingency*) or TI=(contingency assessment* or assessing contingency* or assessment* of contingency*))) or (PO=human and (AB=(causal learning or contingency learning) or TI=(causal learning or contingency learning))) or (PO=human and (AB=(concept of correlation) or TI=(concept of correlation))) or (PO=human and (AB=(correlation judgment* or correlation judgement* or judgment* of correlation* or judgement* of correlation*) or TI=(correlation judgment* or correlation judgement* or judgment* of correlation* or judgement* of correlation*))) or (PO=human and (AB=(contingency judgment* or contingency judgement* or judgment* of contingency* or judgement* of contingency*) or TI=(contingency judgment* or contingency judgement* or judgment* of contingency* or judgement* of contingency*))) or (PO=human and (AB=(covariation judgment* or covariation judgement* or judgment* of covariation* or judgement* of covariation*) or TI=(covariation judgment* or covariation judgement* or judgment* of covariation* or judgement* of covariation*))) or (PO=human and (AB=(causality judgment* or causality judgement* or judgment* of causality or judgement* of causality or judging causality) or TI=(causality judgment* or causality judgement* or judgment* of causality or judgement* of causality or judging causality))) or (PO=human and (AB=power PC) or (TI=power PC))
- 2 The streamed trial procedure is a compromise. It permits the application of psychophysical methodology to contingency assessment, but does so by implementing a contingency assessment task that may seem to lack ecological validity (the very rapid exposure to many cue and outcome presentations in a short period of time). However, the procedure does incorporate a categorical judgment, rather than an analogue contingency rating scale. Elsewhere we have argued that the use of such categorical judgments is similar to the actual contingency assessment challenges faced by people (Allan et al., 2008). When we have experienced a series of cues and outcomes, we typically do not judge the statistical relationship

between the events – rather, we make a categorical, and typically a binary, decision. Does cloud seeding produce rain, or is it a waste of money? Does this substance cause the dough to rise, or is it an ineffective leavening agent? Does this symptom mean that the patient has a disease, or is the symptom noninformative? The same is true of the contingency assessment task faced by the non-human animal. For example, when an odor associated with a potential predator is detected, the question for the animal is *not* “what is the magnitude of the relationship between the odor and the presence of a predator?” Rather, the question is, should the animal stay or flee?

3 <http://www.quansoft.com/>

4 Four of the participants (XG, MC, AS, and AB) had participated in the single-cue experiments in Allan et al. (2008), and two (SA and CH) had not. The pattern of performance was not influenced by experience.

5 Integrations of an associative acquisition process into a SDT framework have appeared periodically in the literature, but apparently have not been influential (Boneau & Cole, 1967; Boynton, Smith, Stubbs, 1997; Mason, Idrobo, Early, Abibi, Zheng, Harrison, & Carney, 2003; Schmajuk, 1987).

6 While our SDT account and the comparator hypothesis are similar in that they both place cue interaction in the decision process, the quantitative relationship between the two views has yet to be formulated.

Appendix A

In the one-phase blocking task there are eight types of trials corresponding to the eight cells of the 4 x 2 matrix. The Bush-Mosteller equations are shown below for each cell of the matrix. The b values for outcome and no outcome have been equated, $l = 1$ on outcome present trials and $l = 0$ on outcome absent trials. As in Table 2, the frequencies of the cue-outcome combinations are represented by the letters a, b, ...

	O	~O
$C_T C_C$	a $\Delta V_T = \alpha_T \beta [1 - V_T]$ $\Delta V_C = \alpha_C \beta [1 - V_C]$	b $\Delta V_T = \alpha_T \beta [0 - V_T]$ $\Delta V_C = \alpha_C \beta [0 - V_C]$
$C_T \sim C_C$	c $\Delta V_T = \alpha_T \beta [1 - V_T]$ $\Delta V_C = 0$	d $\Delta V_T = \alpha_T \beta [0 - V_T]$ $\Delta V_C = 0$
$\sim C_T C_C$	e $\Delta V_T = 0$ $\Delta V_C = \alpha_C \beta [1 - V_C]$	f $\Delta V_T = 0$ $\Delta V_C = \alpha_C \beta [1 - V_C]$
$\sim C_T \sim C_C$	g $\Delta V_T = 0$ $\Delta V_C = 0$	h $\Delta V_T = 0$ $\Delta V_C = 0$

$$\text{mean } \Delta V_T = \frac{a\alpha_T\beta[1 - V_T] + c\alpha_T\beta[1 - V_T] + b\alpha_T\beta[0 - V_T] + d\alpha_T\beta[0 - V_T]}{a + b + c + d} \quad (1)$$

Equation 1 simplifies to

$$\text{mean } \Delta V_T = \alpha_T\beta \left[\frac{a+c-V_T(a+b+c+d)}{a+b+c+d} \right]$$

At asymptote mean $\Delta V_T = 0$

$$a + c - V_T(a + b + c + d) = 0$$

$$V_T = \frac{a + c}{a + b + c + d} = P(O | C_T) \quad (2)$$

$$\text{mean } \Delta V_C = \frac{a\alpha_C\beta[1 - V_C] + e\alpha_C\beta[1 - V_C] + b\alpha_C\beta[0 - V_C] + f\alpha_C\beta[0 - V_C]}{a + b + e + f} \quad (3)$$

Equation 3 simplifies to

$$\text{mean } \Delta V_C = \alpha_C\beta \frac{a + e - V_C(a + b + e + f)}{a + b + e + f}$$

At asymptote mean $\Delta V_C = 0$

$$a + e - V_C(a + b + e + f) = 0$$

$$V_C = \frac{a + e}{a + b + e + f} = P(O | C_C) \quad (4)$$