HYBRID PERFORMANCE IN LEARNING SOCIAL CONTINGENCIES

ASSOCIATIVE AND NON-ASSOCIATIVE PERFORMANCE PHENOMENA IN LEARNING SOCIAL CONTINGENCIES FROM RICH AND HETEROGENEOUS STIMULI

By

AIMEE L. SKYE, B.A.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2008) McMaster University

(Psychology, Neuroscience & Behaviour) Hamilton, Ontario

TITLE: Associative and non-associative performance phenomena in learning social contingencies from rich and heterogeneous stimuli

AUTHOR: Aimee L. Skye, B.A. (McMaster University)

SUPERVISOR: Dr. Lee R. Brooks

NUMBER OF PAGES: x, 170

Abstract

One of the most central and current debates among those studying human contingency learning (HCL) concerns whether it is best understood as the result of associative learning, a product of higher-order cognitive processes, or some combination thereof. Though the field appears to be moving toward the latter accounts, much of the evidence being generated to evaluate and select among them comes from tasks that typically present only information about the few variables involved in the contingency(s), in the exact same manner on every trial. While effective for examining how the statistical properties of experience affect learning, these procedures do not capture some of the conditions of everyday cognition and are apt to be less effective for engaging non-associative and top-down influences on performance.

The current work introduces a task that involves learning contingencies in others' behavior from descriptions that require the learner to determine the focus of learning, and to deal with both variability in manifestation of the objects of learning and extraneous information. Across several experiments, performance reflects phenomena, including ΔP, outcome density and blocking effects, which have been well established in HCL and are consistent with associative accounts. At the same time, the findings also suggest that (a) domain-specific theories affect the weighting of evidence in contingency perception and the discoverability of contingencies, and (b) outcome predictions, a typical measure in HCL, are

influenced by specific instance memory in addition to abstract contingency knowledge. These findings are difficult to reconcile with the data-driven nature of associative views, and join a growing number of demonstrations suggesting that a viable account of HCL must involve higher-order cognitive processes or top-down influences on performance.

Acknowledgements

As I collected my thoughts to write this section, I found myself unable to go any further until I addressed the individual who, by a considerable margin, has played the most important role in the completion of this thesis, and in my academic development over the past several years—my supervisor, Dr. Lee Brooks. For the entirety of what has been a wonderful, but long and sometimes trying journey, Lee has been an extraordinary intellectual resource and a most generous and supportive mentor. It is impossible to put into words my gratitude for the time spent under his tutelage and the high regard I have for him.

I am very fortunate to have had the support of a wonderful and loving family and many good friends. My husband, Jairus, whose love and support has carried me through these past years, has been terribly patient and understanding through the many times where my studies were so consuming of my time and attention. My daughter, Cheyenne, has added immeasurable joy and excitement and wonder to my life. My parents and sister have helped out in so many different ways and encouraged my efforts; and the caring and support I have received from my extended family, in-laws, and my near-sister Heather means an incredible amount to me.

A great many people were instrumental in shaping my experience at McMaster into all that it was. In addition to Lee, they include my supervisory committee, Bruce Milliken, Martin Daly, and Geoff Norman; Dick Day, Margo

Wilson, Ron Racine, Lorraine Allan, David Shore, Betty Ann Levy, and Rolfe Morrison; past and present members of the Brooks-Norman lab, including Meredith, Randy, Sam, Kevin, Chan, Glenn, John, Bruce, Val, Rosemary, and also Betty whose creativity, diligence and time were so important to the work presented here; Elliott, Eric, Andrew, Lisa, Pat, Sue, Phil, and the rest of incoming graduate class of 2000; Jason & Mel, Jason & Launa, Ellen & Rick, Nikki & Fitz, Fil, Alex, the Matt Cs, Mayu, Danny, Jeff, Graeme, Ferrinne, and many, many other fellow graduate students; the awesome support staff in the psychology department (who patiently answered all my hypothetical questions), and numerous others.

Finally, there are individuals whom I never would have met were it not for my graduate studies at McMaster, and who have added so profoundly to my life during these past several years and continue to do so: Amy, Christine, Lynne (and Al), Nicole, Nikie, Sandra and Vicki; Denis and Tamara, and Pete. Life simply would not have been as good without Small Town Ontario Day and thirty-five dollar pies; Montreal '05, turbo shanties, fruit boats, and hotels with bunk beds; Psy Jung's baseball and the Annual Psychology Ball Tournament; Funnilingus and the Phoenix Cup; pick-up hockey; NASA parties and defense celebrations, holiday visits, New Year's Eves and 8:1 "camping" at Sand Bay; Hoegaarden thievery, late night/musical phone calls and random hateful emails; the many other wild and fun times we have shared, and most importantly, the friendships we've built. You guys rock!

Table of Contents

List of Tables and Figures

Introduction

One of the most fundamental attributes that humans must have in order to function efficiently is an ability to learn about predictive relationships or contingencies that exist between events in their environment. Knowing what signals an untrustworthy individual can save us from financial, emotional, or physical harm. Aversive responses to cues suggesting something is inedible or unsafe surely afford protection from illness or ill health. We learn a great deal about how we can shape the behavior of things in the world, particularly other people, as a function of our intervening or acting in one way or another. Further, we use such knowledge to engineer desirable outcomes in our interactions with people or objects. Stated more generally, capitalizing on learned contingencies to anticipate what will happen and to guide responding in future experiences is more efficient than computing responses anew to each experience one encounters. Knowledge of contingencies also provides a basis for dealing with novel and unfamiliar experiences that is, on average, better than no information.

Human contingency learning (HCL) has been the focus of much empirical work over the years in psychology. The field of study concerned with HCL[1] has developed largely since the 1960s, but most especially over the past three

[1] Here, and throughout the thesis, I will use the term human contingency learning to reflect a superordinate category, which subsumes both learning that X predicts Y, and that X causes Y. This is not to deny that there are distinct or special attributes of causal learning, but rather is adopted because much of the work and discussion presented here applies to all forms of contingency learning.

decades. Major developments in the field have been the subject of multiple reviews (e.g., Shanks, 2007; Pineño & Miller, 2007; De Houwer & Beckers, 2002); and these will be summarized in the following section. To foreshadow the highlights, though, HCL has had rather close ties to conditioning, and the perspective of associative learning has dominated the field, both theoretically and methodologically, for much of its history. This connection between conditioning and human contingency learning is not surprising, as both domains address much the same fundamental issue—learning about predictive relationships or 'what signals what'. Indeed, there is a good deal of empirical evidence suggesting that performance in both domains can be understood as the product of the same basic learning principles. However, one of the major issues in the field of HCL concerns the sufficiency of this data-driven view of learning in the face of evidence suggesting that higher order cognitive processes also shape performance.

Human Contingency Learning: A Brief History

Abstraction according to statistical rules

Much of the very early work on HCL adopted a rule-based or statistical view, according to which learning about contingencies involved "applying a rule to integrate probabilities or frequencies of events" (Allan, 1993). The rule-based view engenders at least two primary issues: (a) by what rule should information from experience be integrated in order to know about a contingency, and (b) how well can people follow such a rule? Consider the 2x2 situation in which both a cue and an outcome can be either present or absent across a number of trials. While there are many ways in which such information can be summarized, Allan (1980) argued that the appropriate integration rule to determine the relationship between two binary variables is ΔP, or the difference in the probability of the outcome in the presence and in the absence of the cue [i.e., $P(O|C) - P(O|{\sim}C)$]. Indeed, many early studies demonstrated that people's perceptions or judgments of contingencies tended to agree with ΔP (e.g., Dickinson, Shanks, & Evenden, 1984; Wasserman & Shaklee, 1984; Wasserman, Chatlosh, & Neunaber, 1983; Allan & Jenkins, 1980, 1983; Alloy & Abramson, 1979), and this correspondence between the normative ΔP rule and actual performance no doubt contributed to the popularity of the rule-based view of HCL.

However, rule-based accounts of HCL suffered from two critical problems that seem to have been instrumental in their being superceded by associative learning accounts. One problem was the accumulation of evidence that, under a variety of circumstances, performance in HCL tasks failed to agree with the normative statistical index (Allan, 1993). For instance, Allan and Jenkins (1983, 1980) demonstrated that agreement between judged and actual contingencies was affected by whether the presence and absence of the cue, and of the outcome, were interpreted as an event and a non-event (e.g., green light vs. no light) or as two different events (e.g., green light vs. red light). It was also rather clearly established that judgments of perceived contingency were stronger when the probability of the outcome (or *outcome density*) was greater, as well as when the probability of the cue (or *cue density*) was greater (e.g., Dickinson et al., 1984; Allan & Jenkins, 1983). Furthermore, agreement between objective and perceived contingencies was not only influenced by the outcome density, but by its desirability as well—such as the difference between winning and losing money (Alloy & Abramson, 1979). Finally, research demonstrated that judgments were more congruent with statistical indices when the data were presented in summary rather than trial-by-trial format (Wasserman & Shaklee, 1984; Ward & Jenkins, 1965), and when the task involved evaluating covariation between continuous variables rather than contingency between binary variables (Peterson & Beach, 1967). This list is not meant to be an exhaustive catalog of the conditions under which performance is known to deviate from ΔP. Rather, it is meant to illustrate that as these findings accumulated, they undermined the viability of accounts proposing that contingency knowledge arises because we integrate information according to such a rule.

An altogether different circumstance under which contingency judgments also failed to agree with ΔP was that in which there are multiple cues related to a single outcome. Specifically, evidence emerged that the evaluation of one cue-outcome contingency was influenced by contingencies between other cues and that same outcome (Chapman & Robbins, 1990; Shanks, 1985; Dickinson et al., 1984). Rule-based accounts like the ΔP model couldn't account for this non-independence in the judgment of cue-outcome contingencies, because they assume cell frequency information for each 2x2 comparison is stored and calculated separately (Chapman & Robbins, 1990), The discovery of dependence in contingency judgments, together with the strong resemblance that it bore to *blocking* and *overshadowing* effects in conditioning, was perhaps the most important catalyst in the rise of associative learning accounts (e.g., Pineño & Miller, 2007; Allan, 1993).

The takeover: Associative learning

The beginning for associative accounts of HCL was a paper published by Dickinson et al. (1984), though Shanks (2007) credits Alloy and Abramson's (1979) work on *depressive realism* for suggesting the connection. Dickinson et al. developed a video game in which participants could fire shells at a tank passing though a minefield, with either shells or mines capable of destroying the tank. In this situation, a cue present trial was one in which the shell registered a hit and an outcome present trial was one in which the tank was destroyed. After playing the game repeatedly, participants judged the efficacy of the shell in destroying the tank (i.e., the C-O contingency). Much like previous work in HCL, participants' judgments in the task were a function of the objective contingency (i.e., ΔP) as well as the outcome density. In addition, though, Dickinson et al. compared participants' judgments to those simulated by the Rescorla-Wagner (1972) model of conditioning, and demonstrated very close correspondence between the two. Finally, and most critically, Dickinson et al. showed that judgments of the shell's efficacy were not only influenced by data collected while playing the game, but were influenced by knowledge they had about the potency of the minefield. Prior to playing the game, participants observed that the tank was unlikely, somewhat likely or very likely to be destroyed simply from traveling through the minefield. When they later played the game where they could fire the shell, judgments of the efficacy of the shell were greater when the minefield was known to be a weak destructive force and attenuated when the minefield was known to be a potent destructive force. This pattern of performance very closely resembled blocking effects (e.g., Kamin, 1968) in conditioning, where an initial association between cue A and a US (unconditioned stimulus) interferes with learning about the predictive value of cue X when AX is later associated with that US. Because blocking, or cue competition, effects were well established in conditioning and predicted by associative learning models, their discovery in HCL very convincingly suggested that contingency learning might be explained as the development of associations between events according to principles like those in the Rescorla-Wagner (R-W) model.

In the years that followed this influential work, a good deal of evidence was accumulated in support of the view that HCL was the product of an associative learning process. As noted earlier, one feature of human contingency judgments is that they often correlate with ΔP, the objective relationship between two binary variables (e.g. Dickinson et al., 1984; Allan & Jenkins, 1983). Chapman & Robbins (1990) showed that this feature could easily result from an associative learning process because the R-W model, when learning has reached asymptote, is identical to the ΔP statistic. However, the R-W model is about much more than just learning at its terminal state after all information has been received. It is about the process by which knowledge develops. Simulations of the R-W model show that acquisition curves resemble positive logarithmic functions for positive associations, negative logarithmic functions for negative associations, and an early bias towards positive growth followed by decline and settling at null for zero associations (Allan, 1993). Using a modified version of the video game described earlier in which contingency judgments were solicited at regular and frequent intervals, Shanks (1985) demonstrated that the development of contingency knowledge across the task did indeed closely resemble these acquisition curves. Similarly, the asymptotic nature of the R-W model predicts that accuracy of contingency judgments should increase with the number of trials, which has also been demonstrated empirically (Van Overwalle & Van Rooy, 2001; Lopez, Almaraz, Fernandez, & Shanks, 1999; Shanks, 1987)

The demonstration of blocking in HCL, like that described earlier by Dickinson et al. (1984), was instrumental to the theoretical shift favoring associationism. Blocking effects are part of a broader class of phenomena commonly referred to as *cue interaction effects*. Cue interaction effects are a staple in conditioning research, because they are clearly anticipated by associative learning models like the R-W model in which different cues compete for the *limited* associative potential of a given outcome. While the discovery of blocking was surely a catalyst for associationism in HCL, demonstrations that other cue competition effects analogous to those in conditioning tasks could be observed in HCL strongly reinforced the movement.

Blocking reduces the association between a cue (e.g., A) and a US. In contrast, *signaling*, another type of cue interaction effect, does the opposite—it enhances the A-US association. Consider a condition in which a person experiences some A-US trials, and other US alone trials. As the frequency of US alone trials increases, the contingency and association between A and the US weakens. In contrast, consider a signaled condition in which all US alone trials are instead preceded by a different cue (e.g., B). Although the A-US contingency is the same in this and the former condition, A becomes more strongly associated with the US in this signaled condition (e.g. Durlach, 1983). To understand why, it is necessary to remember that the R-W model proposes that all cues, including the ever-present context, compete for the limited associative strength supported by a given outcome. In the first condition then, the US alone trials allow the context to increase its association to the US, and this context, which is also present on the A-US trials, antagonizes the extent to which A becomes associated with the US. In the signaled condition though, B becomes associated with the US on the trials where A is not present, which reduces the context-US association. As a result, A faces less competition from the context in developing an association to the US on the A-US trials (Durlach, 1989).

Using the same tank destruction video game task described earlier, Shanks (1986) demonstrated that signaling effects also occur in HCL (see also Shanks, 1989). To map the signaling procedure just described onto his task, the US was destruction of the tank, A was a hit from the fired shell, the context was the minefield, and B, the cue that signaled tank destruction in the absence of a shell hit, was a plane flying across the screen. The critical comparison from Shanks' task concerns two conditions that shared a zero contingency between a shell hit and tank destruction, but that differed in whether or not tank destruction in the absence of a hit was signaled by the plane. This comparison demonstrated that participants' ratings of the effectiveness of the shell in destroying the tank were greater in the signaled condition (despite both conditions having the same objective contingency).

Relative cue validity is another form of cue interaction discovered in animal conditioning that was later established in HCL as well (Allan, 1993). As the label implies, the essence of this effect is that the association between a cue and an outcome is not only a function of their objective statistical relationship, but how that objective relationship compares to those involving other cues and the same outcome. To be more concrete, consider Wagner, Logan, Haberlandt & Price's (1968) demonstration of this effect using a conditioning task in which there were 3 conditioned stimuli (e.g., A, B and X) and one US (e.g., O). In one condition, training involved AXO+ and BXO- trial types, but in another, training involved AXO+, AXO-, BXO+ and BXO- trials. While X shared the same objective statistical relationship to O in both conditions, X was a much poorer predictor of O *relative to A and B* in the former but not the latter condition. Indeed, when X was a poorer relative predictor, it elicited lower rates of responding, suggesting it had developed a weaker association with O. Relativity in judgments of cue-outcome contingencies has also been documented using similarly structured tasks in which the cues were different foods and the outcome an allergic reaction (Wasserman, 1990), or the cues were different symptoms of a disease outcome (Shanks, 1991); and also using a modified version of the tank destruction video game described earlier (Baker, Mercier, Vallee-Tourangeau, Frank, & Pan, 1993).

Discontent with the new management: Top-down processes and flexible knowledge use

The entire collection of evidence supporting associative accounts of HCL is of course greater than that described in the previous section, as a complete review is beyond the scope of this work. Nonetheless, that which was reviewed here should be sufficiently convincing of both the pervasive influence of associative learning theory on the field of HCL throughout much of its history, and that this dominance was reasonably justified. Associationism did not reign unchallenged in HCL, though, and indeed several criticisms of the associative learning approach were raised over the years.

One such challenge was the discovery of *backward blocking* in HCL (Shanks, 1985). In the typical, forward blocking effect, learning that A is predictive of O interferes with learning that X also predicts O from subsequent experience with compound AX trials. In backward blocking, the AX compound trials appear first, followed by the A alone trials. According to the R-W model, the subsequent A alone trials should not affect the association between X and O because X is not present on those later trials and associative strength can only be altered when a cue is present. However, Van Hamme & Wasserman (1994) later argued that backward blocking could easily be accommodated by the R-W model with a small modification that allowed associations to be updated on trials where a cue didn't occur but was expected. In other words, after AX compound trials, when A alone trials are subsequently presented, A can produce an expectation for X, which would then allow revision of the association between X and O.

Interest in the notion of rule-based abstraction also did not fade easily in the face of rising interest in associative learning in HCL. Over the years, Cheng and her colleagues (Cheng, 1997; Cheng & Holyoak, 1995; Cheng & Novick, 1990, 1992) generated a series of models that proposed more sophisticated statistical rules than the simple ΔP rule. In essence, the early Probabilistic Contrast Model (PCM) version proposed that ΔP for a given cue-outcome comparison should be evaluated not over all available evidence, but conditional on focal sets that control for alternative causes and distinguish between independent versus interactive causes. A later version of this model called Power

PC proposes that the power of a cue as a causal agent is a function of ΔP and the probability of the outcome in the absence of the cue. Although it makes good conceptual sense to evaluate one cause while holding others constant, and there is evidence that this characterizes what people do (Spellman, 1996a, 1996b), it has also been established that the PCM model makes identical predictions to the R-W model (e.g., Tangen & Allan, 2003; Cheng, 1997). Turning to the later Power PC model, Allan (2003) provides a rather unfavorable evaluation, concluding that its predictions are inconsistent with empirical data and that it should not be regarded as superior to the R-W associative learning model. In addition, like earlier rule-based approaches, these models do not explicate the process that produces the output they predict at the end of learning (Shanks, 1995, and see De Houwer & Beckers, 2002).

The reactions to associative theories of HCL just described are somewhat technical in nature—is this parameter or that one adjusted properly, should a new constraint or limit be added, etc.—and seem amenable to resolution by some fine-tuning of the formula by which associations develop. A different set of challenges to associative accounts of HCL has gained momentum in recent years, and these collectively cut to the very heart of associative accounts of HCL—that learning is data-driven and that responding maps straightforwardly from acquired associations. It is this brand of challenges that appears to be most substantive and is most relevant to the current work.

The R-W model provides a function according to which the association between a cue and outcome is updated. Consider a cue A, and an outcome O. According to the R-W function, the change in associative strength for A on a given trial is equal to $\alpha\beta(\lambda - \Sigma V)$; where alpha and beta are parameters representing the (physical) salience of A and O respectively, and the term in parentheses represents the maximum associative strength O can support minus that consumed by all cues (including context) present in that trial. Little examination is needed to see that, from the beginning, learning is fully and completely a function of the statistical properties of each experience—whether and what cues are present, whether or not the outcome occurs, and the salience of these variables according to their physical characteristics. In other words, the locus of all action in associative accounts like those based on the R-W model lies in the data.

The work that spearheaded the movement to recognize top-down or higher-order cognitive processes in HCL was that by Waldmann and Holyoak (1992) demonstrating that cue interaction effects were dependent on the general knowledge or model of the world that participants applied to the task. The basic blocking paradigm involves three variables, A, B and C for example. The nature of blocking tasks is typically one where A and B are potential cues or causes that predict some effect, C. Waldmann and Holyoak pointed out that the identical statistical or associative learning situation occurs for circumstances in which A is a single cause that produces two effects, B and C. However, they argued that blocking should only be observed in the former and not this latter circumstance because participants subscribe to a model of the world in which causes interact but effects are independent. Sure enough, the authors demonstrated blocking when participants learned whether two features of the appearance of people (paleness of skin and stiffness of posture) were causes of a specific emotional response in others, but not when these same variables were described as symptoms of a novel disease. Further, this asymmetry in blocking effects was discovered despite identical statistical or associative learning input, which makes it especially problematic for any data-driven account of performance. Since its initial demonstration, the same effect has been replicated using different problem content as well as single stage blocking procedures (Tangen & Allan, 2004; Waldmann, 2000, 2001).

Based on their results, Waldmann and Holyoak (1992) argued for a more "mentalistic" view in which "people use meaningful world knowledge, often of a highly abstract sort, to guide their learning about new domains." Support for this position has grown with the accumulation of additional demonstrations suggesting that learning is similarly affected by the application of general world knowledge as well as inferential reasoning processes. Two such demonstrations have provided evidence that blocking is modulated by participants' assumptions and knowledge about (a) whether causes have additive effects and (b) the maximal intensity of the outcome. Consider again the typical blocking design in which one learns that cue A predicts outcome O, and this later interferes with learning that cue B also predicts O from compound AB trials. As Beckers, De Houwer, Pineño & Miller (2005) point out, in order to devalue B as a cause of O requires an assumption about outcome additivity—A produced O, and were B also a cause, it should have produced something more than O when paired with A, but because it didn't, B must not be a cause of O. Recent evidence has shown that blocking effects are in fact much stronger when participants are instructed or encouraged to approach the task with an additive effects model (Beckers et al., 2005; Lovibond, Been, Mitchell, Bouton, & Frohardt, 2003; Mitchell & Lovibond, 2002).

Outcome additivity effects occur because participants' assumptions or knowledge about properties of the *cues* shape their expectations regarding the outcome, as well as the inferences that they draw from their observations and deviations from what they expected. Relatedly, Beckers et al. also noted that participants' assumptions or knowledge about the properties of the *outcomes*—particularly ceiling effects or variability in outcome intensity—could also affect what they would expect and the inferences they would draw from their observations. In other words, under outcome additivity, participants expect the compound AB to produce a stronger outcome than A alone, and when it does not, they infer that B must not be a cause of O. However, this expectation would be unreasonable, and the inference questionable, if one knew that the outcome produced by A alone was already at ceiling, or if the outcome produced by the compound was subject to some constraint on intensity. Beckers, de Houwer, and their colleagues have provided evidence that manipulating outcome-related assumptions does indeed modulate blocking effects. Specifically, blocking effects were observed to be much stronger (despite identical objective contingency information) when participants were aware that outcomes more intense than those presented during the blocking phases were possible, either through a pre-exposure phase (Beckers et al., 2005) or direct instructions about maximal outcome values (De Houwer, Beckers, & Glautier, 2002)

Human contingency learning, according to associative accounts, is data driven and cue-outcome or cause-effect associations develop according to an algorithm in which there is no role for higher order cognitive influences like prior knowledge, causal models of the world, or inferential reasoning. Clearly, the studies just described are problematic for such accounts. Associative accounts also assume learning *is* the storage of associations, whether between cues, outcomes or elements thereof (e.g. Pineño & Miller, 2007, Shanks 1995), and that performance is a simple function of the current associative strength in memory (e.g., Dickinson et al., 1984). These assumptions have also come under fire in the face of evidence that performance is influenced by information other than associative strength, and more importantly, that different knowledge appears to be flexibly applied depending on the nature of the test used to evaluate contingency learning.

For instance, various cue interaction effects like blocking and relative validity, while observed when participants judge the causal status or predictive value of the cue, tend not to occur when the test questions focus on the frequency of co-occurrence of the cue and outcome or the conditional probability of the outcome given the cue (Gredebäck, Winman, & Juslin, 2000; Matute, Arcediano, & Miller, 1996). Participants are also quite able to provide the latter co-occurrence information. These studies not only suggest that more information than current cue-outcome associative strengths must be available in memory, but that people can flexibly and sensibly use different information they have acquired to respond appropriately to different kinds of questions. In other words, when asked about the predictive value of a cue, or its power to cause an outcome, people base their response on what they know about the cue-outcome relationship, but also knowledge about other cues related to the same outcome. Knowledge about other cues, though, is irrelevant and therefore not consulted when judging the frequency with which the cue and outcome occurred together. Flexibility in the use of different knowledge at the time of test implies that something other than associative strength must be involved in coordinating what gets applied and when. Allan, Siegel, and Tangen (2005) conducted a signal detection analysis of data collected from an HCL task investigating outcome density effects, and provided rather convincing evidence for two, separable parameters in performance—contingency sensitivity and a criterion for responding. Under the typical associative account, a response criterion or "gatekeeper" is both unnecessary and foreign, because there are not multiple sources of knowledge needing coordination, nor is there a role for any higher order reasoning about what knowledge is appropriate in responding to a particular question.

The Crossroads in Human Contingency Learning and the Current Work

Clearly, there is much about how humans acquire and use contingency information that would seem to follow from the kinds of associative learning mechanisms proposed to underlie conditioning. Such a position is attractive not only for its elegant simplicity, but also for its domain-generality and species-generality. Contingency or predictive learning is central to both human and non-human animals as they solve countless specific problems in their environment, in the service of shared goals like survival, reproduction, and amassing resources. Moreover, connectionism has done a tremendous job demonstrating the power of relatively simple learning algorithms, like those in associative learning theories, to produce complex behavior (e.g., McClelland & Rumelhart, 1985; and see McClelland, 1987 and Rumelhart, 1987 for further discussion). At the same time, there is good reason to believe human contingency learning involves more than just the development of associations in a very programmatic manner based only on the statistical properties of one's experience. The specific background knowledge we accumulate, and the models or theories of the world that we possess, are rich and valuable resources from which we can draw inferences and guide responding. In response to this crossroads in the field, many have adopted the rather reasonable "best of both worlds" approach, arguing for explanations or hybrid models that incorporate both associative mechanisms and top down influences (e.g., Vadillo & Matute, 2007; Tangen & Allan, 2004; Cheng, 1997; Price & Yates, 1995).

Assuming a hybrid account of HCL, a key question going forward then is to discern what contributions are necessary from associative learning mechanisms and higher order cognitive processes to account for performance, and perhaps more importantly, how these influences are coordinated. The current work reflects the position that two of the empirical "tools or tactics" that would be well suited to this initiative have been so far underrepresented, if not absent, among HCL research. Broadly, these tools are (a) allowing participants to capitalize on and apply domain-specific knowledge during a task and (b) reflecting some of the richness and variability typical of much everyday experience in experimental stimuli and materials. Consideration of the effect of domain-specific knowledge has not been completely absent from HCL, although it is certainly not prevalent. For instance, Waldmann and Holyoak (1992) argued that some cue competition effects might be traceable to inferences from domain-specific knowledge. Furthermore, in developing their causal model theory of HCL, Waldmann and Holyoak (1992, and Waldmann, 1996) acknowledge that some of the inferential or top-down processes in HCL might involve domain-specific knowledge, although they focus primarily on more domain general knowledge effects like those described earlier. With regard to stimulus complexity, the stimuli in HCL tasks almost always are restricted to conveying only information about the cue(s) and outcome(s) of interest, and the values of these variables are expressed in an identical manner across all trials (but see work by Meiser & Hwestone, 2006 and Meiser, 2003 on illusory correlations).

The current work first and foremost provides a paradigm for investigating HCL under conditions that strongly engage domain-specific background knowledge and that reflect some of the natural stimulus complexity and variability that we typically confront outside the laboratory. The task introduced is structurally identical to traditional contingency learning and conditioning procedures. However, its content involves learning about contingencies in the behavior of others, and the learning proceeds from richly descriptive, unique vignettes. In this way, the current task differs from the often somewhat less familiar problem content and distilled, homogenous stimuli typically employed in HCL tasks. When Dickinson et al. (1984) published their influential paper on the applicability of associative learning theory to HCL, they first established validity of the novel task they developed by demonstrating that it could produce performance effects that were widely documented in the field at the time. Following that same strategy, it is demonstrated here that the current task is also capable of generating three well-established performance effects—sensitivity to ΔP, outcome density effects, and blocking. The remainder of the work then demonstrates three key issues in performance that would simply be unobservable using traditional HCL tasks and that are clearly relevant to understanding how learning and performance are products of both associative and top-down processes. These performance effects are that (a) domain specific knowledge influences the contribution of individual items to contingency knowledge, (b) domain-specific knowledge produces expectations that appear to influence how items are processed and constrain what is learned, and (c) performance in some measures like outcome predictions is influenced by memory for specific prior instances as well as contingency knowledge abstracted across experience. The concluding section discusses how the paradigm introduced here, and the unique observations that follow from it, may contribute to the theoretical and empirical issues currently of interest to those studying HCL.

Running head: CONTINGENCY LEARNING WITH VARIED STIMULI

Events come in all shapes and sizes: Contingency learning with heterogeneous and elaborate stimuli

Aimee L. Skye

McMaster University

Send correspondence to:

Aimee Skye

Dept. of Psychology, Neuroscience and Behaviour

McMaster University

Hamilton, Ontario, CANADA L8S 4K1

Email: skyeal@mcmaster.ca

Phone: (905) 525-9140, ext. 24824

Abstract

Investigations of human contingency learning typically involve more homogeneous events than those from which much everyday learning proceeds. This paper describes a novel contingency learning task that uses complex and heterogeneous stimuli, but that otherwise corresponds with traditional procedures. Participants read multiple vignettes about a character whose behavior is contingent upon a circumstantial variable, and contingency learning is assessed with subsequent outcome prediction and explicit rating tasks. Two experiments demonstrate classic findings of traditional contingency learning research. In Experiment 1 (N = 68), responses were influenced by the objective contingency (ΔP) and outcome density, and in Experiment 2 (N = 38), blocking effects were observed in contingency ratings and outcome predictions. The results suggest this novel task is a viable means of investigating contingency learning under conditions that better approximate those of much everyday cognition. A brief discussion of the anticipated utility of adopting such an approach is offered.

Events come in all shapes and sizes: Contingency learning with heterogeneous and elaborate stimuli

Consider a tenant in an apartment who uses the building's elevator upon leaving or returning home. On each occasion, the tenant presses the call button, or not, and the car arrives, or not. As these experiences accumulate, the tenant gains knowledge about the functioning of the elevator (e.g., it's rather unreliable) and also decides how to respond in specific situations (e.g., I've pressed the button and the elevator hasn't arrived, so I'll take the stairs).

This rather common situation illustrates issues that are central to human contingency learning (HCL): How do we learn about contingencies between events across repeated trials, and how do these contingencies affect our knowledge and our responding in specific situations? The example of learning about elevator reliability also nicely reflects the conditions under which HCL is typically studied. Although the specific topics or cover stories vary from learning about allergens causing reactions (Van Hamme & Wasserman, 1993), to chemicals affecting bacteria survivability (Tangen & Allan, 2004) and the effect of fertilizers on plant growth (Spellman, 1996), to name a few, these procedures used to investigate HCL have much in common. They provide the learner with repeated trials, each conveying some binary value for events—X and Y in the simple case of a single cue and outcome. Importantly, there is typically no variability in the manifestation of X or Y (or the absence of X or Y) from one trial to the next, nor is each event individuated by accompanying contextual information. As such, trials or experiences in investigations of HCL tend to be homogeneous and indistinct from one another.

In contrast, much of our day-to-day cognition occurs under importantly different circumstances. To illustrate, compare the elevator example with, for instance, learning about whether women tend to be poor drivers. While such learning also proceeds from experiences we accumulate in life (though not exclusive of other influences or cognitive factors), normally these experiences involve rich contexts and both quantitative and qualitative variability in manifestation of the variables of interest. We may experience others' driving while commuting on a major highway or while heading to our summer cottage, while on our bicycles, and in our own vehicles or those of our carpool partners. Furthermore, differences certainly exist among drivers, even those of the same sex: the female librarian approaching retirement neither looks nor acts much like the heavily-pierced teenager who recently received her license. Of course, displays of good or bad driving are similarly variable: being rear-ended by someone engrossed in a cell-phone conversation is rather different than getting stuck behind the driver who needs several attempts to get a small car into a more than ample parking space.

The use of uniform, homogenous stimuli is clearly appropriate for examining sensitivity to general properties abstracted across experience. Such stimuli are easily manipulated to produce different patterns or regularities, and their simplicity reduces noise variance in responding. However, a variety of other informational sources and cognitive processes often shape performance in everyday situations, as supplementary influences or through biasing the process of abstraction. Indeed, work in domains as diverse as classification and concept formation (Brooks, Graydon & Wood, 2007; Hannah & Brooks, 2006), reasoning (Cosmides, 1989), and visual perception (Fei-Fei, VanRullen, Koch & Perona, 2005) demonstrates that episodic knowledge from individual experiences, context-specificity in responding, prior knowledge, and variation in deployment of attentional and analytic processes are important issues to consider in understanding performance. Common to such studies is that performance is examined using richer or more varied stimuli that better approximate those in our natural environment, which often produces marked differences in behavior from that shown using simpler material.

In studies of human contingency learning, attempts to capture some of the richness and heterogeneity of natural experience have been scarce. Cheng & Novick (1990) provide one of the closest attempts in their examination of bias in attributions of the causes of others' behavior. Indeed, the content of their study materials did reflect qualitative differences across test problems (e.g., "Jane had fun washing dishes on this occasion." or "Alice displayed sculptures made from clay in her home on this occasion."). However, the problem structure and manner of presenting the information were more distilled and transparent than that used in the current work, and that which we typically confront outside the laboratory.

Here, I introduce a task that addresses this objective, but that nonetheless maintains close procedural alignment with more traditional tasks. In this novel task, learners are presented with several unique vignettes describing a character's behavior and the context in which it occurs. Each description instantiates a relationship between some behavioral outcome and a situational or contextual cue (e.g., being active when the time of day is early), and the variety among these descriptions requires a process of abstraction that is not present with traditional, uniform materials but which is present in many everyday situations. Learning about the relationships is subsequently assessed using both the expected likelihood of the outcome in novel situations and explicit ratings.

Experiment 1 demonstrates that participants' responses are clearly a function of both the objective cue-outcome contingency (ΔP) and the density or frequency of the outcome. Experiment 2 demonstrates cue interaction—a hallmark phenomenon in both contingency learning and conditioning—in which the perceived predictive value of one cue is influenced by the predictive value of additional cues. The results of both experiments are highly consistent with those of traditional contingency learning tasks (Allan & Jenkins, 1983; Dickinson, Shanks & Evenden, 1984; Tangen & Allan, 2004; Allan, Siegel & Tangen, 2005), lending validity to this novel contingency learning task. This correspondence is particularly compelling because, as a consequence of using richer and more complex, variable stimuli in the current task, the cues and outcomes to be evaluated are much less transparent, and tracking of those cues and outcomes during learning is more complicated. The opportunities that this task is expected to hold for future research in various domains are outlined briefly in the general discussion.

Experiment 1

Method

Participants. Sixty-eight McMaster University students participated for course credit or $10 cash. Participants ranged in age from 17 to 30 years, and English was their primary language.

Stimuli. Participants were given 20 vignettes describing a character's behavior in specific, unique contexts. Collectively, the vignettes depicted some contingency between a behavioral outcome (O) and a circumstantial cue (C). For instance, in one set, the character tended to be rude to his relatives but not to non-relatives. Examples from that set include (also, see Appendix A):

> When Graham came home from his job at the recreation center, Graham flopped down on the couch, grabbed the TV remote and switched channels from the show his brother Peter was watching to the football game without asking.
>
> Graham called VISA about a charge that appeared on his credit card bill. The customer service agent put him on hold, but then she forgot about him. After holding for 20 minutes, Graham hung up and called back.

> When he got back through to the same agent, he calmly asked what had happened and was very understanding when she explained her error.

As these examples illustrate, each vignette conveys whether or not the cue was present (e.g., relative vs. non-relative) and whether or not the behavioral outcome occurred (e.g., rude vs. polite).[2] However, they are a stark contrast to the uniform stimuli used in traditional tasks (e.g., for the current contingency example, traditional stimuli would involve repeated pairings of the four combinations of the words Relative, Non-relative, Rude and Not Rude, for instance).

Contingency learning was evaluated using additional vignettes describing novel circumstances in which the cue was present or not (see Appendix B for the test scenarios for the Rude/Relative task set). Participants were then given two options representing the presence and absence of the behavioral outcome, and were asked to indicate the probability that each would occur in that vignette. For example, in the cue-absent (C–) vignette below, the absence of the behavioral outcome (O–, shown in Option B) should be rated as more likely than its presence (O+), given a contingency between interacting with relatives and being rude.

---

[2] In traditional contingency learning tasks, it is clear on each trial whether the outcome and the cue did or did not occur. With the current task, a potential concern is that the cue or outcome status in individual events may be interpreted idiosyncratically, producing variation across participants in the "objective" contingency and outcome density at learning. Individual variability in event interpretation was evaluated by asking independent individuals to make a binary classification of whether each event would be consistent or inconsistent with the contingency rule. Data collected for many of the stimuli used here and additional items used in other experiments showed very strong agreement (97%) with the author's classification of events.

> Graham bought a stereo from a store in Toronto and he borrowed his friend Justin's car to go and pick it up. When he was finished, Graham...
>
> (a) returned his friend's car without replacing any of the gas he'd used.
>
> (b) filled up his friend's car with gas before returning it.

Probability ratings were collected using two sliding scales (one for each option) anchored between 0 and 100. The sliding scales were linked such that changing the value on one automatically adjusted the other in order to force the sum of the values on both scales to equal 100.

A contingency rating scale was also used to evaluate participants' knowledge of the relationship. The scale was anchored from -100 to +100, and included five text descriptors. The descriptors corresponded to relationships that were moderate (e.g., if Relative, Rude to some degree) and perfect (e.g., if Relative, always Rude) and both positive and negative in direction, and to the absence of a relationship (e.g., Relative had no influence on Rude). Scales were always labelled such that accurate ratings of the contingency would be positive. A copy of the rating scale used for the Rude/Relative task set is provided in Appendix C.

The session included four separate contingency learning tasks that were identical in structure, and differed only in the content of the focal relationship and vignettes. The three additional relationships used concerned activity level contingent upon time of day, helpfulness contingent upon availability, and talkativeness contingent upon the presence of a specific individual. Each task consisted of 20 learning vignettes, 2 test vignettes for the outcome prediction task (one cue-present and one cue-absent) and the contingency rating exercise. Two levels of objective contingency, or ΔP (.7 and .3), and two levels of outcome probability, or density (.75 and .35), were crossed to yield four between-subjects experimental conditions, as both variables have proven to be important in previous research (Allan, Siegel & Tangen, 2005). The entire set of learning vignettes for the high contingency (.73) and high outcome density (.75) condition is reproduced in Appendix A. For the remaining three conditions, the description of the cue, outcome or context in those vignettes was modified as necessary to reflect the frequency of cue present/absent (C+/C) and outcome present/absent (O+/O–) learning items shown in Table 1. For any given participant, ΔP and the outcome density were identical for all four contingency learning tasks.

Table 1

Frequency of Learning Item Types for Contingency (ΔP) and Outcome Density (OD) Conditions in Experiment 1

	Low Contingency ($\Delta P = .3$)				High Contingency ($\Delta P = .7$)[a]			
	Low OD (.35)		High OD (.75)		Low OD (.35)		High OD (.75)	
	Outcome							
Cue	+	–	+	–	+	–	+	–
+	5	5	9	1	7	3	14	1
–	2	8	6	4	0	10	1	4

Note. Plus and minus symbols signify presence or absence, respectively.
[a]ΔP is.70 in the low density and .73 in the high density condition.

Procedure. Participants were told they would be given several behavioral descriptions and then would answer questions based on that information. The order of the four contingency learning tasks was randomized for each participant, as was the order of the learning and test vignettes within each task. All vignettes were displayed on a computer screen and narrated through an audio file. Presentation time varied across items according to the length of each vignette and therefore the time needed for its narration. In each contingency learning task, presentation of the learning vignettes was immediately followed by the outcome prediction task and then the rating scale task.

Results

Analysis. Four separate contingency learning tasks were used simply to increase the measurement sample, much like using four questions rather than one on a quiz. Thus, effects within, or in interaction with these four tasks were not evaluated. Contingency ratings were averaged across the four tasks for each participant, and then submitted to a 2-way ANOVA with contingency and outcome density as between-subjects variables. In the outcome prediction task, for each participant, the probability assigned to the option representing the behavioral outcome (O+) was averaged across all four tasks for the cue-present and then for the cue-absent test vignettes. These averages were submitted to a mixed 3-way ANOVA with cue value (present or absent) as a within-subject variable, and contingency (ΔP = .7 or .3) and outcome density (.75 or .35) as between-subjects variables.

Contingency Ratings. Mean contingency ratings for both levels of objective contingency and outcome density at learning are presented in Table 2. Ratings by participants who experienced a stronger objective contingency were higher than those provided by participants who experienced a weaker objective contingency ($F(1, 64) = 21.05$, $MSE = 481.1$, $p < .001$, $\eta_P^2 = .247$), demonstrating that participants became sensitive to the actual relationships that existed across learning items. The data in Table 2 also show a classic outcome density effect, namely that contingency ratings were higher when the frequency of the outcome during learning was greater ($F(1, 64) = 7.58$, $MSE = 481.1$, $p < .01$, $\eta_P^2 = .106$). There was no Contingency X Outcome Density interaction ($F < 1$, $MSE = 481.1$).

Table 2
Mean Contingency Ratings as a Function of Objective Contingency (ΔP) and Outcome Density Conditions at Learning for Experiment 1

	Outcome Density	
Contingency	Low (0.35)	High (0.75)
$\Delta P = .3$	28 (5.5)	48 (5.3)
$\Delta P = .7$[a]	57 (5.2)	67 (5.3)

Note. Standard errors of the means are shown in parentheses.
[a]ΔP is.70 in the low density and .73 in the high density condition.

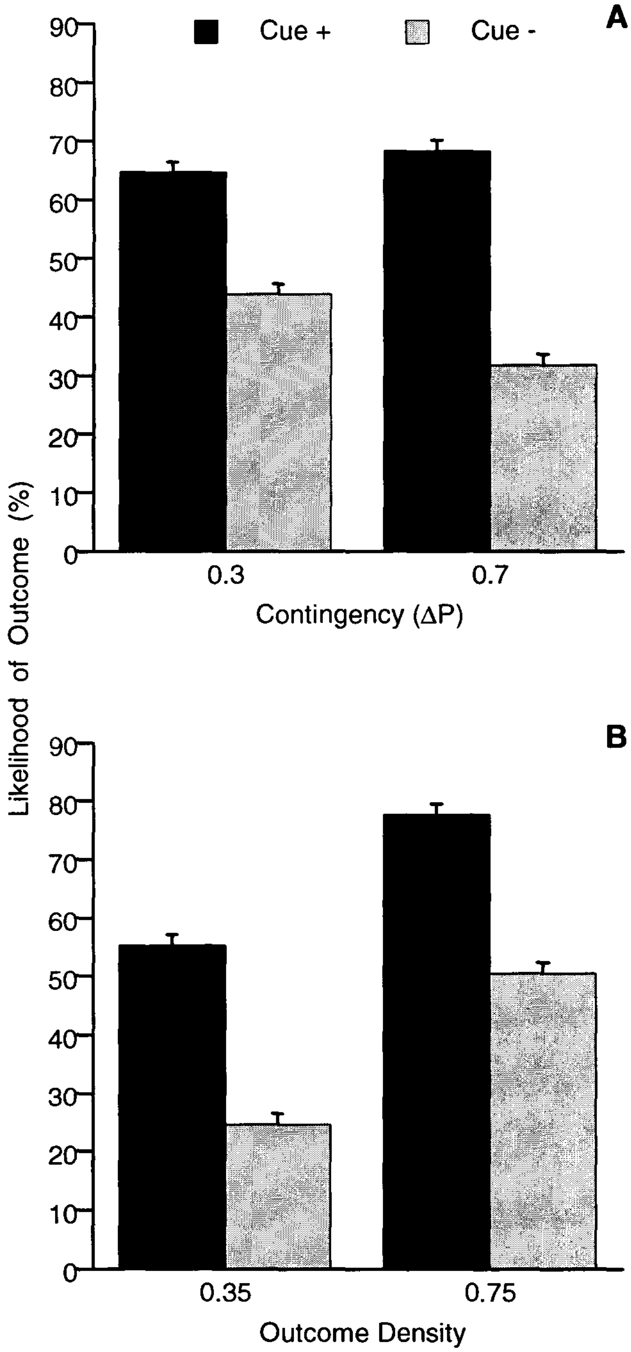
Outcome Predictions. Participants experienced a cue whose presence was either strongly or mildly predictive of some outcome. If these contingencies were learned and used to guide responding to novel situations, then predictions should reflect that the outcome would be more likely in the presence than in the absence of the cue, and this difference should be larger in the strong than in the mild contingency condition. Panel A of Figure 1 shows mean likelihood predictions for the outcome in cue-present and cue-absent test vignettes for both contingency conditions. The analysis confirmed a significant interaction between contingency and cue value ($F(1, 64) = 17.96$, $MSE = 115.0$, $p < .001$, $\eta_P^2 = .219$).

Post hoc analyses on the effect of the cue value within each contingency condition were conducted using paired t-tests with a Bonferroni correction, where

α=0.05/2 = 0.025 (see Maxwell, 1980 for evidence this procedure is favorable for repeated measures designs). According to these analyses, participants in both the weaker (ΔP = .3) and in the stronger (ΔP = .7) contingency condition predicted that the outcome was more likely to occur when the cue was present rather than absent ($t(32) = 9.25$, $SE = 2.25$, $p < .001$, $d = 1.27$ and $t(34) = 12.68$, $SE = 2.88$, $p < .001$, $d = 2.32$ respectively). Thus, the two-way interaction between the cue value and the objective contingency condition in the main analysis was significant because the presence or absence of the cue affected expectations about the outcome likelihood in both conditions, but it did so much more strongly when participants were presented with a strong objective contingency.

Allan, Siegel and Tangen (2005) recently demonstrated outcome density effects not only in contingency ratings, but also in trial-by-trial expectations about the outcome's occurrence. Specifically, greater outcome densities increased the tendency to expect the outcome in subsequent learning trials, independent of the objective contingency and, therefore, the presence or absence of the cue. The outcome prediction data from the current study replicate this finding. Panel B of Figure 1 shows mean likelihood predictions for the outcome in cue present and absent test vignettes for both outcome density conditions. When the density of the outcome during learning was greater, participants predicted the outcome would be more likely to occur in novel vignettes ($F(1, 64) = 183.74$, $MSE = 106.6$, $p < .001$, $\eta_P^2 = .742$), and this outcome density effect did not interact with contingency ($F < 1$, $MSE = 106.6$) or with cue type ($F < 1$, $MSE = 115.0$).

Figure 1



Mean outcome likelihood (+ *SE*) that participants predicted in novel cue-present and cue-absent vignettes, for two learning conditions differing in objective contingency (Panel A) and outcome density (Panel B) in Experiment 1.

Discussion

Although the experiences from which learning proceeded were richer in meaningful content and more variable in manifestation than the distilled stimuli often used in studies of contingency learning, performance in the current task reflected the sensitivity to both ΔP and outcome density which is typically observed. When the objective contingency was greater, participants rated the cue-outcome relationship to be stronger, and expectations of the outcome likelihood in novel scenarios were more strongly a function of whether the cue was present. When the outcome occurred more frequently during learning, participants also rated the cue-outcome relationship to be stronger, demonstrating a classic outcome density effect. Moreover, the expected likelihood of the outcome in individual events was greater when the outcome occurred more frequently during learning.

Experiment 1 employed the most basic structure of a contingency learning problem, namely a single cue related to a single outcome. In Experiment 2, the task structure was extended to one involving two cues, each of which shared some relationship with the same outcome. Such an extension affords assessment of whether this novel contingency learning task also produces standard cue interaction effects.

Experiment 2

It is well documented that learning about a contingency between a cue and an outcome is affected by the presence of other cues that also predict the same outcome (e.g., Tangen & Allen, 2003; Baker, Mercier, Vallée-Tourangeau, Frank & Pan, 1993; Chapman & Robbins, 1990). In particular, a contingency between cue X and outcome O is perceived to be weaker in the presence of a second cue Y, if Y is more strongly related to O. The current experiment investigates whether the novel task introduced here also produces such hallmark cue interaction effects. The materials were modified so that learning and test items systematically provided information on two circumstantial cues and a behavioral outcome. The objective relationship between one cue and the outcome was always moderate, while the relationship between the second cue and the outcome was much stronger for some participants and much weaker for others.

Method

Participants. Thirty-eight McMaster University students, 15 in the strong second cue and 23 in the weak second cue condition, participated for course credit or $10 cash. Participants ranged in age from 17 to 24 years, and English was their primary language.

Stimuli. The current experiment used three of the contingency learning task sets from Experiment 1 (the fourth was dropped due to time constraints).

Twelve new learning stimuli were created for each task set, increasing the total number of vignettes to 32. Each learning vignette was modified or created to provide information about the original circumstantial cue (X) and behavioral outcome (O) from Experiment 1, as well as a second circumstantial cue (Y). For instance, in the set used earlier to illustrate the task and materials, each vignette indicated whether Graham was interacting with relatives or non-relatives, and whether his behavior was rude or not. In the current experiment, the vignettes also indicated whether the person with whom he was interacting was young or old.

The objective contingency between X and O was always fixed and of moderate strength ($\Delta P = .5$). Two versions of the learning stimuli were created for each set so that the objective relationship between Y and O was much stronger than that between X and O in one condition (i.e., $\Delta P = 1.0$) and much weaker than that between X and O in the other (i.e., $\Delta P = 0$). For any given participant, ΔP for this second cue Y was identical across all three contingency learning task sets. The combinations of 2 cues and 1 outcome (X, Y and O), each with 2 possible values (present or absent), produce 8 distinct types of learning items. The frequency of each type of learning item is shown in Table 3 for both conditions.

As in Experiment 1, contingency learning was evaluated using both the outcome prediction and contingency rating tasks. Two additional outcome prediction scenarios were created for each set, yielding a total of 4 test items. Each item indicated whether both cue X and Y were present or absent, and the four test items covered all possible combinations (X+/Y+, X+/Y–, X–/Y+, and X–

/Y–). Otherwise, the prediction task was identical to that used in Experiment 1. Similarly, the current experiment used the same contingency rating scale from Experiment 1. Ratings were solicited for both the contingency between X and O and that between Y and O, with text descriptors on the scale modified appropriately in the latter case. The order of rating the two contingencies was randomized within each participant and each task set.

Table 3
Frequency of Learning Item Types for Weaker and Stronger Second Cue Conditions in Experiment 2

Item Type	X is Related to O Y is *Unrelated* to O ($\Delta P_X = .5$ and $\Delta P_Y = 0$)	X is Related to O Y is *Perfectly Related* to O ($\Delta P_X = .5$ and $\Delta P_Y = 1.0$)
X+ Y+ O+	6	12
X+ Y– O+	6	0
X– Y+ O+	2	4
X– Y– O+	2	0
X+ Y+ O–	2	0
X+ Y– O–	2	4
X– Y+ O–	6	0
X– Y– O–	6	12

Note. X and Y represent different circumstantial cues, O represents behavioral outcome. Plus and minus symbols signify presence or absence, respectively.

Procedure. The procedure was identical to that used in Experiment 1, and participants were randomly assigned to either the stronger or weaker second cue condition.

Results

Analysis. As in the previous experiment, responses were averaged across the three separate contingency learning tasks prior to analysis. Participants' average contingency ratings for the moderately predictive cue X and for the stronger or weaker predictive cue Y were separately analyzed using independent t-tests with condition (predictive strength of cue Y) as the grouping variable.[3] In the outcome prediction task, participants responded to test items that represented the 4 combinations of the presence and absence of cues X and Y. The average probability they assigned to the option representing the behavioral outcome (O+) for each of these four types of test items was submitted to a mixed 3-way ANOVA with value of cue X (present or absent) and value of cue Y (present or

[3] Each participant produced two average contingency ratings, one for cue X and one for cue Y, and participants were either in the condition where cue Y was strong or weak. The critical effect of interest is whether ratings of cue X, which has the same objective contingency in both conditions, differs between those who received a strong or weak cue Y. One might argue that the appropriate analysis for this design is a two-way ANOVA with one within-subject and one between-subjects factor, and that the critical effect of interest would be tested by the interaction between these two factors. However, the fact that the objective contingencies are identical for cue X and maximally different for cue Y almost certainly guarantees that a significant interaction will result but not for the reason that is of theoretical interest. This, together with the fact that a difference in the rating of cue Y between conditions is absolutely expected given the difference in objective contingency and is therefore of interest only as a manipulation check, is the rationale for analyzing contingency ratings for cue X and Y using separate independent t-tests.

absent) as within-subject variables, and condition (stronger or weaker cue Y) as a between-subjects variable.

Contingency Ratings. Mean contingency ratings between the outcome O and the moderate cue X, and between O and the stronger or weaker cue Y, for participants in both conditions are presented in Table 4. Although all participants received the same objective contingency between X and O ($\Delta P = 0.5$), ratings of this contingency were lower when learned in the context of a second cue Y that was more predictive of the outcome compared to the context of a less predictive second cue ($t(1, 36) = 2.2$, $SE = 6.69$, $p < .05$, $d = .75$). The objective contingency between Y and O was maximally different for participants in the two conditions ($\Delta P = 1.0$ and 0), and quite clearly their ratings of this contingency reflect this potent manipulation ($t(1, 36) = 11.38$, $SE = 8.24$, $p < .001$, $d = 3.79$).

Table 4
Mean Contingency Ratings for the Moderately Predictive Cue X and Stronger or Weaker Predictive Cue Y in Experiment 2

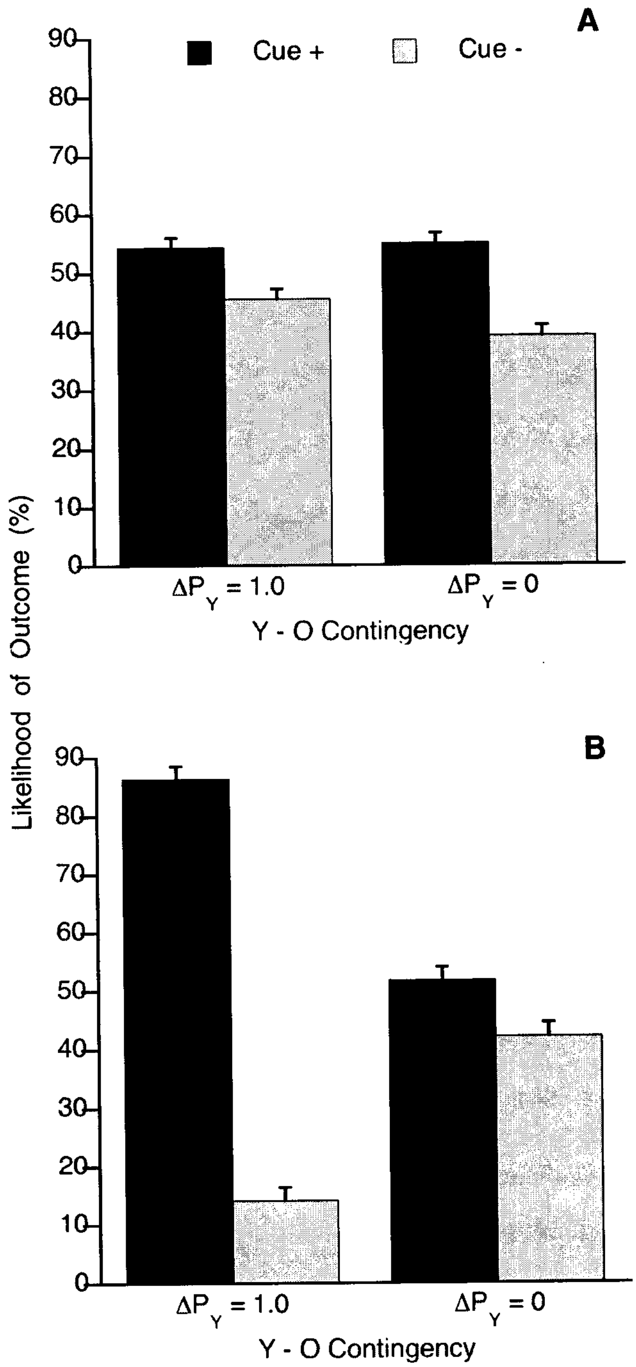
Condition	Rating for Cue X ($\Delta P = .5$)	Rating for Cue Y
Weaker Cue Y ($\Delta P = 0$)	36 (4.2)	- 4 (5.2)
Stronger Cue Y ($\Delta P = 1.0$)	21 (5.2)	90 (6.4)

Note. Standard errors of the means are shown in parentheses

Outcome Predictions. Figure 2 shows the mean likelihood with which participants in both conditions anticipated the outcome to occur in novel test vignettes. Panel A illustrates those predictions separately for vignettes where cue X was present or absent, and Panel B does so for vignettes where cue Y was present or absent. As with the contingency ratings, the critical effect of interest concerns whether the presence and absence of cue X affects outcome predictions differently between the two conditions, despite an identical objective X – O contingency for both conditions. Indeed, the analysis confirmed a significant interaction between the value of cue X and condition ($F(1, 36) = 7.79$, $MSE = 112.8$, $p < .01$, $\eta_P^2 = .178$).

Post hoc analyses on the effect of cue X within each condition were conducted using paired t-tests with a Bonferroni correction, where $\alpha = 0.05/2 = 0.025$). According to these analyses, participants predicted the outcome was more likely to occur in the presence than in the absence of cue X both when they learned about X in the presence of a second and stronger cue Y ($t(14) = 4.90$, $SE = 1.46$, $p < .001$, $d = 1.41$) or a weaker cue Y ($t(22) = 6.37$, $SE = 1.46$, $p < .001$, $d = 1.93$). Thus, the two-way interaction between the value of cue X and condition in the main analysis was significant because cue X affected expectations about the outcome likelihood in both conditions, but it did so much more strongly when it was paired with a weaker rather than a stronger second predictive cue during learning.

Figure 2



Mean outcome likelihood (+ *SE*) that participants predicted in novel vignettes where cue X (ΔP = .5) was present or absent (Panel A), and where cue Y (ΔP = 1.0 or 0) was present or absent (Panel B), for two learning conditions differing in strength of the Y – O contingency in Experiment 2.

In contrast to cue X, whose objective relationship with the outcome was fixed for all participants, cue Y shared a perfect relationship with the outcome in one condition and no relationship with it in the other condition. Therefore, finding that the presence and absence of cue Y differentially affect outcome predictions across the two conditions is of little theoretical interest, but serves to verify the manipulation. As anticipated, the value of cue Y interacted significantly with condition ($F(1, 36) = 127.99$, $MSE = 278.29$, $p < .001$, $\eta_P^2 = .780$). Post hoc analyses on the effect of cue Y within each condition were conducted using the same procedure described earlier. According to these analyses, participants predicted the outcome would more likely occur in the presence than absence of cue Y if they were presented with a perfect Y – O relationship during learning ($t(14) = 16.65$, $SE = 4.34$, $p < .001$, $d = 7.7$). Somewhat unexpectedly, participants who were presented with no Y – O relationship during learning also predicted the outcome would more likely occur in the presence rather than absence of cue Y ($t(22) = 2.79$, $SE = 3.46$, $p < .05$, $d = 0.94$). Nonetheless, the effect of cue Y on outcome predictions was much smaller in the no Y – O relationship condition than in the perfect Y – O relationship condition, which validates the experimental manipulation.

The only other significant effect in the main analysis of outcome predictions was the three-way interaction between value of cue X, value of cue Y and condition ($F(1, 36) = 9.43$, $MSE = 124.11$, $p < .01$, $\eta_P^2 = .207$). To explore the meaning of this interaction, it is useful to first define the “effect of X” as the subjective probability of the outcome when X is present minus that when X is absent. For all participants the X – O relationship was moderate, which should produce a positive, non-zero effect of X in their predictions. In addition, the principle of cue interaction indicates that this effect of X will be diminished when X is learned in the presence of a second cue Y that shares a stronger (compared to a weaker) relationship with the outcome. Consequently, the effect of X observed in the stronger Y condition should be smaller in both Y+ and Y– scenarios than in the weaker Y condition. However, examination of the outcome prediction cell means suggests that this attenuation of the effect of X in the strong Y condition occurred primarily in Y+ scenarios. Specifically, the effect of X in the stronger Y condition was 0% for Y+ and 15% for Y– scenarios, compared to 21% and 13% for participants in the weaker Y condition respectively.

Discussion

Contingency learning tasks that present participants with one or more binary cues and a binary outcome across multiple trials have repeatedly demonstrated systematic cue interaction effects. One could argue that such tasks optimize conditions for observing this phenomenon because the cues and outcomes, and their presence or absence, are readily identifiable and because there is often a lack of extraneous information that could influence interpretation of stimuli and deployment of attention. Here, using stimuli where (a) information is not parsed into discrete variables whose presence or absence is obvious and unambiguous, and where (b) participants determine the focus of their attention and evaluation from among heterogeneous and extraneous detail, the current experiment demonstrates that cues interact with one another and compete in learning in much the same fashion. In particular, a moderately predictive cue was rated to be less predictive, and it had less of an influence on participants' expectations of the outcome's occurrence, when it was learned in the context of a stronger rather than weaker second predictive cue.

Aside from the critical demonstration of cue interaction, this experiment produced two additional and somewhat curious findings worthy of discussion. First, for participants who learned about a cue (Y) that shared no statistical relationship with the outcome, their ratings of the relationship reflected that reality but they were nonetheless influenced by this unpredictive cue when deciding how likely the outcome would be in novel scenarios (although this was a considerably smaller influence than for those who learned that same cue was a perfect predictor). While such discord between measures may seem problematic, it is only so if both measures are (presumed to be) indices of all and only the same evidence and influences on responding. Indeed, Allan, Siegel & Tangen (2005) also reported a dissociation between contingency ratings and outcome predictions, and using a signal detection analysis, they argued that the dissociation reflected not a difference in sensitivity to the contingency between measures, but that additional influences like outcome density information affected participants' response criterion in the prediction task only.

The current dissociation quite possibly has a similar explanation. For the learning items they were provided, participants may well have abstracted contingency information suggesting that the cue shared no statistical relationship with the outcome. However, recall that all the cues and outcomes used in the current experiment were plausible correlates of one another according to everyday experience. Thus, when predicting the outcome of a very specific situation, participants could certainly have consulted their background knowledge or theories and experience, in addition to any learned contingency, in shaping what they would expect to happen or in justifying their expectations.

Finally, when examining the interaction between Cue X and Y in outcome predictions, it appeared that the stronger cue Y overshadowed the moderate cue X primarily in scenarios where Y was present and not those where it was absent. The question, then, is why might this asymmetry in cue interaction have occurred. Although speculative at this point, it is possible that the asymmetry reflects a confirmation bias of the sort often observed in studies of human judgment and decision-making. A confirmation bias would amount to cue Y being more salient on Y+ than on Y– trials, which would be consistent with the observation that the overshadowing effect of cue Y on cue X was restricted to scenarios where Y was present. For the stimuli used in the current task, knowing the value of a cue is not just a clear and simple perception, but instead requires the appropriate attention and interpretation of the learner—a circumstance which surely enables greater variation in the salience of different types of evidence or information.

General Discussion

The current work introduced a task that provides learners with qualitatively very different stimuli from which to acquire knowledge about contingent relationships than are typically used in such investigations. Specifically, the stimuli in the current task reflected variability in the manifestation of the cue(s) and outcome across items, individuating contexts and information extraneous to the focal, related variables—properties that are true of much everyday experience outside the laboratory. Despite these unique stimulus properties, performance in the current task very much reflected the correspondence with associative learning principles that is frequently observed in studies of human contingency learning. In Experiment 1, participants' ratings of the contingency and their predictions about the likelihood of the outcome in the presence and absence of the cue were functions of both the objective contingency (ΔP) and the outcome density. Moreover, when participants were given information about two cues related to a single outcome, cue interaction effects—where the perceived predictive value of one cue is attenuated when paired with a more strongly predictive second cue—were observed in both their contingency ratings and outcome predictions.

The current study thus generalizes hallmark findings of contingency learning research to circumstances more reflective of everyday experience. It is worth noting that performance in the current task might well have been different from that in traditional tasks, given that here, the variables or relationships to be learned were not blatantly obvious from the outset and the processes of interpreting and coding the cue and outcome value on each trial and therefore tracking trial-by-trial statistics was somewhat more complicated. Beyond demonstrating generalization, though, the current task is also attractive because of its potential for investigating issues critical to understanding performance in everyday contingency learning situations. Indeed, this was the primary motivation for developing the task, and the remainder of this section outlines some of these issues, and how the current task might contribute to their empirical investigation.

When we learn about contingencies outside the laboratory, our individual experiences differ in the extent to which they influence our perceptions of a relationship. To provide some illustration, consider again the notion of learning about whether females tend to be bad drivers. An egregious driving display such as running head-on into a building after selecting the wrong gear will surely have a stronger influence on the perceived relationship than even several mundane cases of failing to signal a turn, or even one such failure that results in a common sort of collision. Similarly, we may well be inclined to underweight a sample of driving in circumstances deemed to provide a poor test of the gender-driving ability relationship, such as when the driver is inebriated or driving in a foreign country for the first time. These illustrations underscore the point that, to the contingency learner under everyday conditions, not all experiences are created equal.

In typical investigations of human contingency learning (see Allan, 1993 for a review), individual trials affect learning and performance by altering the statistical distribution of cues and outcomes *anonymously, and with equal weight.* With homogenous learning trials in which the cue and outcome manifestations never vary and no individuating context is included, it is impossible for a single event to exert any special influence owing to its particular properties (other than presentation order). Capturing the variable nature and influence of individual events in laboratory investigations of contingency learning provides a means for addressing issues such as how qualitative variability in the cue or outcome, or how item-specific differences in the quality of evidence, affects explicit knowledge of contingencies.

Allowing individual events to exert distinctive influences is also important for understanding how we respond in specific situations. To illustrate, imagine that we know Jeff, and Jeff happens to be an extremely poor driver. When we later meet Jim, who reminds us very much of Jeff, what are we to expect of Jim's driving ability? Surely, if our aggregated experiences to date suggest that females tend to be poor drivers, then we might expect Jim to be a decent driver on account of his maleness. However, we might also be swayed in the opposite direction by the specific knowledge we have from a similar, prior case—namely, that Jeff is a poor driver. Indeed, work in areas like classification and artificial grammar learning have shown that when tasks afford an influence of similar prior instances that is separable from abstracted knowledge of regularities, this instance-specific knowledge significantly affects our response to subsequent, individual events (Brooks, Norman & Allen, 1991; Brooks & Vokey, 1991; Vokey & Brooks, 1992). While contingency learning research has tended to focus on responses to individual trials for what they indicate about learning of the cue-outcome relationship, these measures can, given appropriately-designed materials, also shed light on very interesting questions about how various forms of knowledge are coordinated to produce a response in a specific situation. For example, tasks like the one presented here should be well-suited to investigating issues related to selective attention, which require means of distracting or dividing the attention of participants, as well as issues related to the context-dependency in contingency learning and the influence of expectations and personal beliefs.

Finally, the current materials focus on the impressions we acquire about other people, a domain where contingency learning seems crucial to operating efficiently in our social world. However, our impressions of others can deviate markedly from the generalizations that our experiences would reasonably support. For instance, people frequently attribute far too much consistency to the personalities or behavior of others (Ross & Nisbett, 1991). Learning about the nature of others obviously involves more than just accumulating evidence, and it is a process heavily influenced by the background knowledge and personal theories that most of us possess about human behavior. The current task affords representation of much of the complex and variable nature of a person's behavior in a manner that readily engages our prior knowledge and beliefs, while also affording clear control over the statistical regularities across that person's behavior. Impression formation tasks that simultaneously capture both attributes are not ubiquitous, and this feature makes the current task likely to yield interesting insights into how we aggregate information to form impressions of others, and how influences such as stereotypes, lay theories or other forms of prior knowledge about the world bias this process.

References

Allan, L. G., Siegel, S., & Tangen, J. M. (2005). A signal detection analysis of contingency data. *Learning & behavior : a Psychonomic Society publication, 33*(2), 250-263.

Allan, L. G. (1993). Human contingency judgments: Rule based or associative? *Psychological bulletin, 114*(3), 435-448.

Allan, L. G., & Jenkins, H. M. (1983). The effect of representations of binary variables on judgment of influence. *Learning and motivation, 14*(4), 381-405.

Brooks, L. R., Graydon, R., & Wood, T. J. (2007). The role of inattention in everyday concept learning: Identification in the service of use. *Memory & Cognition, 35*(1), 1-14.

Brooks, L. R., Norman, G. R., & Allen, S. W. (1991). Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology: General, 120*(3), 278-287.

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General, 120*(3), 316-323.

Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology, 58*(4), 545-567

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? studies with the wason selection task. *Cognition, 31*(3), 187-276.

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

Fei-Fei, L., VanRullen, R., Koch, C., & Perona, P. (2005). Why does natural scene categorization require little attention? exploring attentional requirements for natural and synthetic stimuli. *Visual Cognition, 12*(6), 893-924.

Hannah, S. D., & Brooks, L. R. (2006). Producing biased diagnoses with unambiguous stimuli: The importance of feature instantiations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(6), 1416-1423.

Maxwell, S. E. (1980). Pairwise multiple comparisons in repeated measures designs. *Journal of Educational Statistics, 5*(3), 269-287.

Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology.* New York, NY, England: McGraw-Hill Book Company.

Spellman, B. A. (1996). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science, 7*(6), 337-342.

Tangen, J. M., & Allan, L. G. (2004). Cue interaction and judgments of causality: Contributions of causal and associative processes. *Memory & Cognition, 32*(1), 107-124.

Van Hamme, L. J., & Wasserman, E. A. (1993). Cue competition in causality judgments: The role of manner in information presentation. *Bulletin of the Psychonomic Society, 31*(5), 457-460.

Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 328-344.

Appendix A

Set of vignettes used to present contingency between target character's rudeness and the relatedness of people with whom he was interacting for the high contingency and high outcome density condition of Experiment 1

Relative / Rude (Cue + / Outcome +)

When Graham came home from his job at the recreation center, Graham flopped down on the couch, grabbed the TV remote and switched channels from the show his brother Peter was watching to the football game without asking.

Graham had borrowed his mom's car one day, and he was supposed to pick her up from her doctor's appointment at 4:30. Graham was with his buddies and didn't pay attention to the time, so he was an hour late picking up his mother.

When Graham found out his cousins Ian and Molly were coming for the weekend, he complained that they would hog the Playstation and said that he wasn't giving up his bed for them.

When Graham got home Sunday after his touch football game, he walked in the house with his cleats still on, tracking mud everywhere in the foyer. He threw his cleats in the closet and just left the mud there for his parents to clean up.

Graham's Aunt Delores gave him a spice rack as a Christmas gift. When Graham opened up the present, he said "What the heck am I going to do with this?" and tossed it aside.

Without bothering to ask first, Graham lent his dad's staple gun to Russell, a friend from his lacrosse team.

When his grandparents were visiting from Toronto, Graham quickly ate his dinner and then spent the whole evening talking to his friends on MSN and playing computer games.

Graham and the rest of his family were sitting down to dinner on Friday; his mom had brought home pizza. Graham wanted seconds, and he took the last 3 pieces of pizza without asking anyone else if they wanted more.

When Graham's sister was showing her mom the new jeans she'd bought, Graham told his sister that she looked fat.

One day, Graham's older brother Dick dropped by. He'd gotten into selling Amway products and wanted to try to get Graham to buy some merchandise and become involved with Amway. Graham cut him off mid-sentence, saying he didn't care what he had to say and then he walked into the other room.

Graham and his cousin decided to have lunch at Burger King. They were waiting for the girl working the counter to come back and take their order. When she returned, she pointed to Graham and asked what he would like to order. Even though his cousin was diabetic and needed to eat quickly because he had just taken insulin, Graham stepped right up and ordered his lunch first.

Graham often calls his cousins at home after 10:00 pm even though he might wake up and disturb their families.

On Remembrance Day, Graham's Canadian History class had invited his great Uncle Harold, who is a World War II veteran, to visit and speak to the students. Graham was sitting beside Brody and Austin, a couple of his friends. They started chatting about their plans for the weekend. Graham joined in the conversation, and the three of them were rather disruptive while Graham's great uncle was telling his story.

Graham's mom was pretty shocked when she found out that she had a sister she never knew about. Apparently, her mother had had a baby when she was quite young and had given her up for adoption but she never told anybody. After learning the news, a meeting was arranged and Graham and his family went to his long-lost Aunt's house for dinner. After dinner, Graham didn't bother helping with the dishes and then later, he left with out saying thank you.

<u>Relative / Not Rude (Cue + / Outcome –)</u>

Graham was at home and wanted to call his girlfriend Nadine. He noticed that his mom was on the phone talking to a friend. Apologizing for interrupting her, he quietly asked his mom to let him know when she was done with the phone.

<u>Non-relative / Rude (Cue – / Outcome +)</u>

Graham works in the afternoons as a lifeguard at a recreation center. Even though he knows that the daytime lifeguards can't leave the poolside until their relief arrives, Graham is almost always late for his shift.

Non-relative / Not Rude (Cue – / Outcome –)

Graham was shopping at Zeller's last weekend. He was looking down at his shopping list when he accidentally bumped a lady with his cart. He stopped, asked if she was alright and immediately apologized, saying that it was his fault for not looking where he was going.

Graham was taking the bus home from school one day. The busses were really busy, but he managed to get a seat. A few minutes after Graham got on, an elderly man got on the crowded bus. Graham got up right away and gave the man his seat.

Last summer, Graham had to appear in court as a witness to a car accident. He dressed in his best suit and tie, and made sure he addressed everyone in the court with their proper titles.

Graham called VISA about a charge that appeared on his credit card bill. The customer service agent put him on hold, but then she forgot about him. After holding for 20 minutes, Graham hung up and called back. When he got back through to the same agent, he calmly asked what had happened and was very understanding when she explained her error.

Appendix B

Novel test scenarios used to assess learning contingency between target character's rudeness and relatedness of people with whom he was interacting

Relative (Cue +)

Two days ago there was a major winter storm; there was at least a foot of snow everywhere and heavy winds that were causing lots of drifting. The university was closed, so Graham had the day off. His dad asked him to shovel out the driveway that afternoon. Graham…

a) told his dad that he would. By the time his dad got home, Graham had shoveled the driveway, the sidewalk and the walkway up to the house.
b) complained that he shouldn't be the only one who had to shovel the snow and said he would only do it if his dad paid him $40.

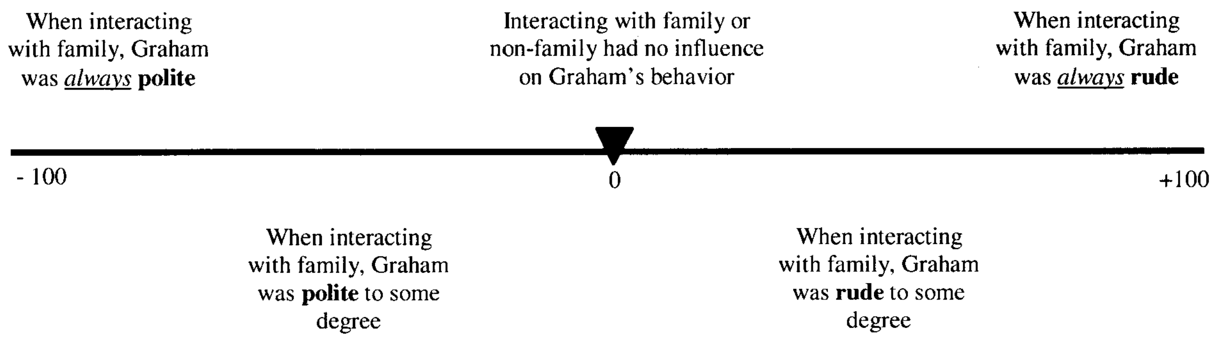
Non-Relative (Cue –)

Graham bought a stereo from a store in Toronto and he borrowed his friend Justin's car to go and pick it up. When he was finished, Graham…

a) returned his friend's car without replacing any of the gas he'd used.
b) filled up his friend's car with gas before returning it.

Appendix C

Rating scale used to probe perceived contingency between target character's rudeness and relatedness of people with whom he was interacting



Beyond the abstract and data-driven in human contingency learning: Domain-specific knowledge influences learning and instantiated knowledge supplements application

Aimee L. Skye

McMaster University

Send correspondence to:

Aimee Skye

Dept. of Psychology, Neuroscience and Behaviour

McMaster University

Hamilton, Ontario, CANADA L8S 4K1

Email: skyeal@mcmaster.ca

Phone: (905) 525-9140, ext. 24824

Abstract

There is increasing debate about whether human contingency learning (HCL) is the product of data-driven associative learning, and what role higher order cognitive influences like inferential reasoning and prior knowledge play in shaping performance. Using a task in which people learn situation-behavior contingencies from descriptions of target characters, the current work demonstrates a relatively novel top-down influence on performance in which domain-specific knowledge and theories determine the weighting of individual pieces of evidence. Furthermore, a second study demonstrates that individual pieces of evidence can affect future expectations beyond their contribution to any perceived contingency, and that these instance effects cannot be explained by simply proposing additional associations. The current work is consistent with hybrid views of HCL that include a role for both associative learning and higher-order cognitive processes.

Beyond the abstract and data-driven in human contingency learning: Domain-specific knowledge influences learning and instantiated knowledge supplements application

During the past 10 or so years, the field of study concerned with how we acquire knowledge of contingencies between events and how that knowledge shapes our behavior has been experiencing an important transformation. Specifically, the dominance of associationism in this field has been increasingly challenged by proponents of theoretical accounts that focus on inferential reasoning and other higher-order cognitive processes. Indeed, these recent developments have already been the subject of at least three reviews (Shanks, 2007; Pineño & Miller, 2007; De Houwer & Beckers, 2002).

For some time, many have argued that human contingency learning (HCL) is a manifestation of the same associative processes that give rise to Pavlovian or classical conditioning (e.g., Shanks, 1995; Allan, 1993; Chapman & Robbins, 1990; Dickinson, Shanks & Evenden, 1984). According to this general view, the extent to which we perceive a cue and an outcome as contingent is primarily a function of the objective relationship that they share, and the relationship that additional cues share with that same outcome, across some set of experiences. In addition, the physical attributes of the variables involved can influence learning, in so far as these attributes affect salience. In essence, then, the associationist perspective regards learning as very much a data-driven process.

However, associationist accounts of HCL have not remained free from criticism (see De Houwer & Beckers, 2002 for a review), and a number of the more serious challenges involve demonstrations that contingency learning from identical data is subject to modulation from top-down processes. Most of these have focused on cue competition, precisely because that phenomenon is central to associationist accounts. For instance, Waldmann (2000) demonstrated that two variables compete for association with a third—or demonstrate blocking—when participants believe they are learning about independent causes that predict a common effect but not when they are learning about independent effects that are diagnostic of a common cause. In other words, participants' responses are guided by the general knowledge that causes can interact but that effects generally do not. Similarly, Lovibond, Been, Mitchell, Bouton & Frohardt (2003) and Beckers, De Houwer, Pineño & Miller (2005) demonstrated that an established cue particularly impedes learning of a novel cue when participants expect but don't find that two cues produce a more intense outcome than does either one alone. This suggests that cue competition occurs because participants adopt a model that causes are additive. Results like these are problematic for associative theories because the statistical properties of experience and objective relationships between the variables are held constant, and what differs are the assumptions or general beliefs about the world that participants apply to the learning situation. Consequently, accommodating these results seems to demand an account of learning that includes a role for reasoning from background knowledge and beliefs; that is, an account of learning that is not solely data-driven.

Evidence supporting the role of more complex cognitive processes in HCL continues to accumulate and receive notice, and is arguably causing important shifts in the direction of the field (Pineño & Miller, 2007; Shanks, 2007). However, one effective strategy for illustrating top down influences on performance is largely absent from this work. That strategy involves moving beyond abstract materials and problem representations to tasks that involve very familiar contents and rich, meaningful stimuli of the sort we often encounter in everyday activities.[4] Notably, this is a research strategy that has proven very useful in exposing the influence of higher order cognitive processes in other domains.

Concept learning is one area in which this strategy has produced some rather compelling evidence for the role of top-down influences on performance (see Murphy, 2002 for a review). For example, category acquisition is enhanced when participants can apply whatever background knowledge they possess to help them relate features to each other and to the category (Murphy & Wiseniewski,

---

[4] In HCL tasks, participants might be asked to learn about whether particular chemicals affect the survival of bacteria (Tangen & Allen, 2004), whether eating certain foods cause allergic reactions (Beckers, De Houwer, Pineño & Miller, 2005), whether particular symptoms or effects are associated with some disease (Waldmann, 2000), or the effectiveness of various weapons in destroying army tanks (De Houwer, 2002). Clearly, the "everydayness" of these tasks is variable, and relative to the content of the current task, it is open to debate. Moreover, in each of these tasks, learners continue to be provided with distilled and homogenous learning trials.

1989; Murphy & Allopenna, 1994). Knowledge also produces expectations, and in at least some circumstances, learners' expectations influence the attention they allocate to different members of a category, and how they seem to process them (Heit, 1998). Moving from learning to application, people possess different theories about natural and artificial objects and these shape the boundaries they use in making decisions about category membership (Keil, 1989). The implication of these results is that learning and application of conceptual knowledge involves much more than abstracting the features that are statistically related to a category. From work in a different domain, we also know that logical reasoning is much improved when conditional relations are expressed through specific, realistic cases as opposed to abstract forms, such as Wason's (1966) selection task (e.g., Griggs & Cox, 1982). While it has been the subject of some debate whether this dissociation reflects the operation of pragmatic reasoning schemas (Cheng & Holyoak, 1985) or reasoning algorithms evolved to deal with specific problems such as social exchange (Cosmides, 1989), it is at least clear that performance in conditional reasoning tasks is not simply data-driven.

The work presented here adopts enriched materials and meaningful problems to engage participants' background knowledge and personal theories in a contingency learning task. Unlike previous demonstrations that have clearly illustrated relatively domain-general, task-wide influences on HCL, the current work provides evidence that much more domain-specific knowledge influences how participants process individual learning items. To make clear the distinction, consider an everyday analog of the task often used in HCL research—namely, learning whether the bouts of nausea you experience are due to a strawberry allergy. That your belief will be influenced by whether the nausea follows or precedes ingestion of strawberries (given strong statistical association), or whether the sickness is more severe when ingestion is combined with other known nauseants like hangovers or the flu, is the sort of top-down influence that has already received empirical attention. The current work is concerned more with how we reason about the evidential value of any particular experience. Consider two scenarios. In one, you fall ill after eating the strawberry shortcake your mother made for your birthday. In the other, you help out at a friend's organic produce farm. The sight of fresh, ripe strawberries is irresistible, and while picking, you eat a handful and subsequently become nauseated. Both experiences are positive for exposure and for reaction, and on those grounds, are equally good evidence for a strawberry allergy. But, you are also apt to know something (or you could at least generate a plausible story) about the dangers of eating unwashed produce, and the kinds of things that can live on berries or in dirt and make you sick. As a result of this interaction between the particular qualities of an experience and the knowledge you possess or theories you can generate, you may well reason that this latter experience provides poorer quality evidence for a strawberry allergy.

It is precisely this sort of interaction between knowledge and data, and reasoning about evidential quality, that the first experiment was designed to investigate. Participants were asked to learn about cues that predicted how different characters would behave—a task with which we almost certainly have a good deal of experience and have accumulated much in the way of relevant background knowledge and beliefs. While the structure of the task was essentially identical to those typically used to study HCL, the stimuli were vignettes that provided rich and unique descriptions rather than being distilled to reflect only the value of the cue and outcome in the same manner on every trial. It was this added item content—the contextual details or specific manifestations of the variables—that was used to activate participants' knowledge and beliefs about how the effect of the cue might be modulated in certain instances, and in so doing, influence the contribution or weight of those items to the perceived contingency.

The first experiment, then, is concerned with item-specific effects on contingency learning as a function of the inferences we make about their individual evidential value. However, the move to rich, individuated stimuli also permits investigation of whether specific items influence our expectations about the outcome in future situations, in a manner separate from their contribution to the perceived contingency—a question addressed in the second experiment. To illustrate this, consider again the everyday task of deciding whether you have an allergy, this time, to eggs. Your exposure will take on many forms, from a breakfast omelette at home or deviled eggs at a family picnic to eggs Benedict at Mother's Day brunch. Each instance is one in which exposure occurred, and let's assume each was also positive for the response, or nausea (and also that you have comparison instances where exposure was absent and nausea rarely occurred). The question of interest in the second experiment is the following: *Later, while away on business, you head down to the hotel restaurant for breakfast, open a warming tray and find it full of Eggs Benedict dripping with hollandaise sauce. You quickly replace the lid, opting instead for pancakes and fruit salad. Is your decision solely a function of the bland statistics suggesting nausea has been associated with eating eggs in the past, or does your prior Mother's Day brunch experience and its ensuing illness come rushing back into mind (and perhaps even into the back of your throat)?*

The eggs Benedict example illustrates that there are two very different forms of knowledge that can influence our behavior. There is abstract knowledge—the contingency or other statistical information that can be extracted from one's past experiences. There is also instantiated knowledge, or individual, similar instances or episodes (or parts thereof) from past experience that come to mind. The distinction between these forms of knowledge and their importance to understanding human performance has been a major topic in categorization and concept learning for quite some time (e.g., Brooks & Hannah, 2006; Dopkins & Gleason, 1997; Whittlesea, Brooks, & Westcott, 1994; Brooks, Norman, & Allen, 1991; Medin, Altom, Edelson, & Freko, 1982; Brooks, 1978; Medin & Schaffer, 1978). Furthermore, in outlining much of the then current state of affairs in associative learning, Shanks (1995) devoted considerable attention to the instantiated-abstract knowledge distinction, clearly prioritizing it as a key issue for theorists and models in this domain.

Evidence that outcome predictions are influenced by instantiated knowledge would clearly be relevant to HCL research in general. Often in HCL tasks, participants are asked to predict outcomes in the learning trials, in order to measure their response to or expectations about what will happen in future events (e.g., Tangen & Allan, 2003; Shanks & Darby, 1998; Waldmann & Holyoak, 1992). These predictions are increasingly being used, in addition to explicit ratings, to evaluate contingency learning (e.g., Tangen & Allan, 2003; Collins & Shanks, 2002; López, Shanks, Almaraz & Fernández, 1998). This assumes that participants' predictions are a function of the contingency knowledge that they have abstracted across experience. Indeed, some have even suggested that outcome predictions might be a more pure reflection of the extent to which two variables have become associated compared to other measures (Tangen & Allan, 2004). This assumption is unlikely to be problematic in tasks that use relatively distilled, homogenous stimuli, either because the individual instances are not rich and unique enough to create distinct memory traces and support instantiated forms of knowledge, or because such influences would not be manifest much differently than responding based on abstract knowledge. However, the practice of using outcome predictions as a measure of contingency learning becomes problematic when both abstract and instantiated forms of knowledge are supported, particularly if they guide responding in different directions, as in the current task and in many everyday situations.

The more specific target of this work, of course, is the debate between associationist and higher-order cognitive perspectives on HCL. Unfortunately, the relevance of instantiated knowledge effects to this debate is currently somewhat fuzzy and speculative. If outcome predictions are influenced by similar past experiences, separate from contingency knowledge, this would suggest encoding of individual learning items as separate episodes that preserve much of the original detail and interpretation. Pineño & Miller (2007) suggest that inferential reasoning models embrace more detailed encoding of information from learning events than do associative models (but see Shanks, 1995). However, inferential models (e.g. Waldmann, 1996; Waldmann & Holyoak, 1992) often fail to address the encoding issue specifically, making it difficult to determine whether they would anticipate instantiated knowledge effects on predictions. Some of the anticipated value of the current work, then, is to encourage further thinking and discussion about the issue of encoding in any account of HCL.

Experiment 1

In many HCL tasks, each learning trial informs participants only of the status of the cue (or other cues) and the outcome. Any trial where the cue and outcome were both present, for example, looks just the same, conveys no more, less or different information, and therefore provides just as good evidence for a contingency as any other trial of the same type. Furthermore, learning trials are often equally weighted in associative accounts of HCL, like those based on the Rescorla-Wagner (1972) model, with the exception of differential weighting by presentation order. However, this is plainly atypical of many everyday contingency learning situations, where, based on the nature of any individual experience and what we already know or believe, some trials may well provide stronger or better quality evidence than others, even if they are all of the same type (i.e., are statistically identical). Recall the earlier example of eating mum's strawberry shortcake versus unwashed berries picked right at the farm.

The current experiment was designed to demonstrate evidence for this interaction between knowledge and data affecting evidential quality and in the degree to which individual experiences influence perceived contingencies. Using a task introduced elsewhere (Skye, 2007), two groups of participants learned contingencies between how a target behaved and some circumstantial cue from richly descriptive vignettes. Both groups were given the same number of vignettes, with identical statistical properties and identical objective cue-outcome contingencies. The groups differed, though, in whether or not the vignettes providing evidence against the contingency (i.e., C+ / O– and C– / O+ trial types) could readily be interpreted as poor quality evidence, based on a priori knowledge or beliefs that participants were likely to possess. If domain-specific prior knowledge or theories influence the evaluation of evidential quality, participants who received the "poor quality" items should discount this counter-evidence to some extent and therefore perceive the cue-outcome contingencies to be stronger.

Method

Participants. Twenty-six McMaster University students participated for course credit or $10 cash. Participants ranged in age from 18 to 23 years, and English was their primary language.

Stimuli. Participants were presented with 20 vignettes describing the behavior of a target character. Each one depicted a specific and unique context, and conveyed whether or not both a circumstantial cue (C) and a behavioural outcome (O) were present. Across the vignettes, there was a positive contingency between the cue (C) and a behavioral outcome (O). To provide an illustrative example, one person tended to be rude (O+) to his relatives (C+) but not to non-relatives, as described through vignettes like:

> When Graham came home from his job at the recreation center, Graham flopped down on the couch, grabbed the TV remote and switched channels from the show his brother Peter was watching to the football game without asking.
>
> Graham called VISA about a charge that appeared on his credit card bill. The customer service agent put him on hold, but then she forgot about him. After holding for 20 minutes, Graham hung up and called back.

When he got back through to the same agent, he calmly asked what had happened and was very understanding when she explained her error.

The objective cue-outcome contingency can be quantified using ΔP (Allen, 1980), and in this experiment ΔP was always 0.4. The frequency of cue present/absent (C+/C–) and outcome present/absent (O+/O–) learning items is shown in Table 1. As can be seen from that table, there are 6 of 20 items, namely the 3 C+/O– and the 3 C–/O+, that serve to reduce the strength of or provide evidence against a C+/O+ contingency. These 6 items were the basis for the between-subjects manipulation in this experiment.

Table 1
Frequency of Learning Item Types for Experiment 1

	Outcome	
Cue	+	–
+	7	3
–	3	7

Note. Plus and minus symbols signify presence or absence, respectively. ΔP is

One group received 6 normal exceptions, or those that many would accept as evidence against the contingency without assuming, imagining or story-telling to heroic proportions (see Appendix A). The following is an example of a normal exception to Graham's tendency to be rude when interacting with his relatives:

> When Graham found out his cousins Ian and Molly were coming for the weekend, he told them that they could take his room and he would sleep on the couch downstairs.

Participants in the second group also received 6 exceptions to the contingency (see Appendix A). However, the content of their exceptions was modified to appeal to participants' knowledge of human behavior and its causes or constraints, and invite them to discount the exceptions as poor quality evidence against the contingency. Contrast the following example with the one offered previously:

> Graham went with his parents to visit his elderly Aunt who, sadly, was expected to die within the next 24 hours. When they got to the hospital room, Graham took off his ball cap and politely remembered to thank her for the birthday card she had sent him last month.

Believing that Graham is rude to his relatives, the fact that he isn't rude to his dying aunt does little to sway our opinion. We know that imminent death tends to create a sad, depressed context in which even a most unruly person's behavior can be subdued, or we might believe that the likely presence of many other family members and health care professionals is keeping Graham's behavior in check. The particular assumptions or interpretations or theories that any learner applies are not of concern here. What is important is that reasoning from prior knowledge and experience, or generating alternate explanations, leads the dying aunt item to be much less powerful evidence against a "rude to relatives" contingency than visiting cousins item. It is also worth noting here that each discountable exception engaged very different knowledge or beliefs or theories about human behavior. This design was deliberately selected to avoid the possibility that participants apply relatively straightforward rules or knowledge about the appropriate contrasts for a particular relationship (e.g., Cheng & Novick, 1990), and instead demonstrate evidence that they engage in a more on-line and ad hoc consideration of evidential quality. Finally, it is important to reiterate here that both the normal and discountable exception groups received identical objective contingencies, as the dying aunt and visiting cousins items are both technically of the C+/O– type.

Contingency learning was assessed using four additional vignettes describing novel scenarios in which the cue was present or not. Participants were given options representing the presence and absence of the behavioral outcome, and they indicated the probability that each would occur in that scenario (total forced to sum to 100). For example, in the cue-absent (C–) vignette below, the absence of the behavioral outcome (O–, or Option B) should be rated more likely than its presence (O+), given that rudeness is contingent upon interacting with relatives.

Graham bought a stereo from a store in Toronto and he borrowed his friend Justin's car to go and pick it up. When he was finished, Graham...

(A) returned his friend's car without replacing any of the gas he'd used.

(B) filled up his friend's car with gas before returning it.

A contingency rating scale was also used to evaluate participants' knowledge of the relationship. The scale was anchored from -100 to +100, and included five text descriptors corresponding to a moderate positive and negative relationship, a perfect positive and negative relationship, and no relationship. Scales were always labelled such that accurate ratings of the contingency would be positive.

The experimental session included four separate contingency learning tasks that were identical in structure, and differed only in the content of the focal relationship and vignettes. The contingency between rudeness and interacting with relatives was the focus of one task, while the three additional tasks focused on activity level contingent upon time of day, helpfulness contingent upon availability, and talkativeness contingent upon the presence of a specific individual. Each task consisted of 20 learning vignettes, 4 test vignettes for the outcome prediction task (two cue-present and two cue-absent) and the contingency rating exercise.

Procedure. Participants were told they would be given several behavioral descriptions and then would answer questions based on that information. The order of the four contingency learning tasks was randomized for each participant. Within each task, the 20 learning vignettes were presented in random order, with the constraint that the first 5 items could not be any of the six exceptions. This constraint was applied to both groups, but was imposed because to discount evidence about a relationship as poor quality logically requires some sense of what that relationship is. All vignettes were displayed on a computer screen and narrated through an audio file. In each contingency learning task, presentation of the learning vignettes was immediately followed by the outcome prediction task (order of prediction items was randomized within each task for each participant) and then the rating scale task.

Results

Analysis. Four separate contingency learning tasks were used simply to increase the measurement sample, like multiple items on a quiz, so effects within or in interaction with these four tasks were not evaluated. Contingency ratings were averaged across the four tasks for each participant, and differences in these ratings between the normal and discountable exception groups were analyzed using an independent t-test. In the outcome prediction task, for each participant, the probability assigned to the option representing the behavioral outcome (O+) was averaged across all four tasks for the two cue-present (C+) and then for the two cue-absent (C–) test scenarios. These averages were submitted to a mixed 2-way ANOVA with cue value (present or absent) as a within-subject variable, and group (normal or discountable) as a between-subjects variable.

Contingency Ratings. Mean ratings of the cue-outcome contingency for both groups are presented in Table 2. Although they received the same objective cue-outcome contingency ($\Delta P = 0.4$), ratings of this contingency were considerably greater for participants in the discountable exceptions group than in the normal exceptions group ($t(1, 24) = 3.4$, $SE = 7.82$, $p < .01$, $d = 1.35$).

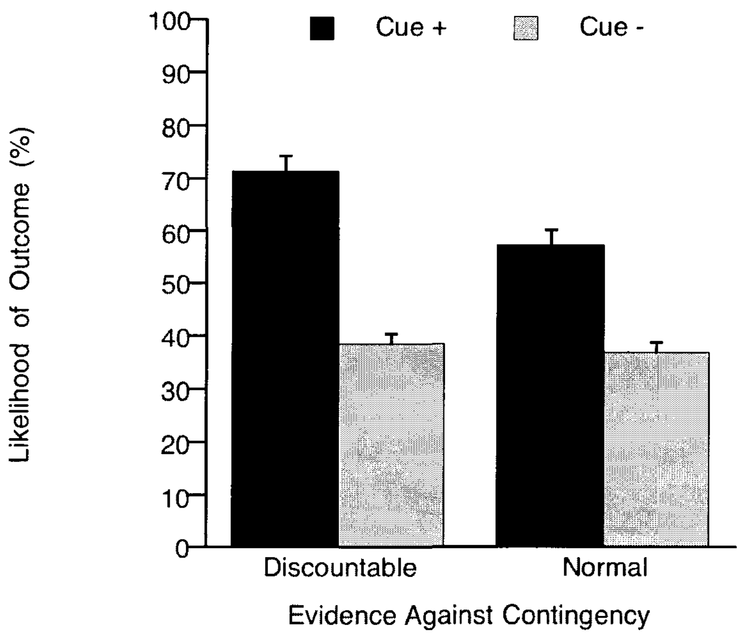
Table 2
Mean Contingency Ratings for Participants Receiving Normal or Discountable Exceptions during Learning in Experiment 1

Group	Rating of Cue-Outcome Contingency ($\Delta P = .4$)
Normal Exceptions	33 (5.9)
Discountable Exceptions	60 (5.1)

Note. Standard errors of the means are shown in parentheses.

Outcome Predictions. Figure 1 shows, for both groups, the mean outcome likelihood predictions for novel test vignettes in which the cue was present or absent. The critical effect of interest concerns whether the presence and absence of the cue affects outcome predictions more strongly in the discountable than in the normal exceptions group, despite both groups receiving the same objective contingency. Indeed, the analysis confirmed a significant interaction between the value of the cue and group ($F(1, 24) = 4.70$, $MSE = 109.0$, $p < .05$, $\eta_P^2 = .164$).

Figure 1



Mean outcome likelihood (+ *SE*) that participants predicted in novel cue-present and cue-absent test scenarios, for two groups differing in the nature of exceptions to a contingency presented during learning in Experiment 1.

Post hoc analyses on the effect of the cue within each group were conducted using paired t-tests with a Bonferroni correction, where $\alpha=0.05/2 = 0.025$ (see Maxwell, 1980 for evidence this procedure is favorable for repeated measures designs). According to these analyses, the outcome was predicted more likely to occur in the presence than in the absence of the cue both for the group who received discountable exceptions during learning ($t(12) = 7.25$, $SE = 4.54$, $p < .001$, $d = 3.52$), and the group who received normal exceptions ($t(12) = 5.68$, $SE = 3.59$, $p < .001$, $d = 2.51$). Therefore, the two-way interaction between the value of cue and group in the main analysis was significant because the cue affected expectations about the outcome likelihood in both groups, but it did so more strongly when the exceptions to the contingency received during learning could be discounted as poor quality evidence.

Discussion

In the current experiment, two groups of participants received identical statistical information about cue-outcome contingencies across the same number of learning trials. In fact, for both groups, 70% of the learning trials were literal copies, and the remaining 30%—those providing evidence counter to the contingency—differed only in whether or not their details invited participants to consider other potent determinants of human behavior and, in so doing, to discount the evidence those items provided about the cue-outcome relationship. The results very clearly suggest participants accepted that invitation. When they could reason from prior knowledge or beliefs and readily attribute counter evidence to influences other than the cue, participants not only perceived the contingency to be much stronger, but their expectations about what would happen in future situations were more strongly influenced by the cue. If participants in this group had been concerned only with the presence and absence of the cue and the outcome on each trial, their responses should have been identical to those who received normal counter evidence, because both groups received identical statistical information.

Given that participants appear to adjust the weight of individual pieces of evidence according to prior knowledge, it is of interest to consider what is likely involved in that process. Consider, for example, just one of the four contingency learning problems that participants completed—learning whether Graham's rudeness was contingent upon interacting with relatives. In order to discount the six items in just this set (see Appendix A), participants would have had to access knowledge about why people tend not to be rude when dealing with the nearly dead, the very mean, and those whom they wish desperately to impress; and how lost inhibitions, lost patience or failing efforts, and even ignorance, can result in deliberate or unintentional rudeness. This collage of prior knowledge was almost surely not collected or organized as "things relevant to the contingency between relatives and rudeness" at the outset of the task. Much more likely is that each specific area was engaged and applied as participants encountered each item. Moreover, while some of the knowledge that participants used may have come from direct prior experience with similar situations, some of what they applied is more likely to have included what they believe, suspect, or theorize about human behavior (as opposed to know in the sense of pre-existing learned associations).

Finally, in the current experiment, what made an item "poor quality evidence" was the operation of additional constraints on behavior that were much more potent than the cue under consideration. Certainly, there are additional dimensions to evidential quality, such as source credibility, or perhaps the recency of evidence or the vividness with which it is recalled. Given that participants appear to make online judgments of evidential quality, they may well adjust the weight of individual items or experiences along these dimensions as well.

Experiment 2

The grand purpose for learning about contingencies in our environment, of course, is so that, when confronted with novel experiences in the future, we can use that knowledge to anticipate what might happen and respond appropriately. Aware that eating eggs makes us ill, we can avoid them in the future, or take steps to minimize an allergic reaction after inadvertently eating them. Knowledge of general patterns or relationships, though, is not the only means by which prior experience can shape behavior. New situations we are confronted with can also remind us of specific past experiences that are similar, and in so doing, suggest a response (recall the eggs Benedict example described earlier). This distinction is an important one precisely because abstract and instantiated forms of knowledge need not exert identical influences on responding.

Many studies of HCL, though, use tasks in which there is little or no variance in the attributes of each learning item or experience, beyond the value of the cue(s) or outcome and presentation order. In such tasks, instantiated knowledge is less potent because learning items are less likely to form rich, individuated memory traces that can be selectively cued in memory. Consequently, it often seems that in HCL research, little attention is paid to the role of instantiated knowledge, separate from abstract knowledge, in shaping our response to future events (but see Shanks, 1995 for a lengthy discussion of the role and importance of instance memory). The current task, however, is quite well suited to affording rich, episodic instance memory and distinguishing the effects of those prior instances on future responding from abstract contingency knowledge; and this is the focus of Experiment 2.

The current experiment used the same basic contingency learning task in which participants were asked to learn contingencies between circumstantial cues and behavioral outcomes from vignettes describing a character. However, in the outcome prediction task, some items were modified to be very similar in content and detail to one of the learning vignettes. In order to separate the influence that this prior instance might have on predictions, it always suggested a response that was in opposition to what ought to be predicted given knowledge of the contingency and the status of the cue in the current test scenario. For example, if the test scenario was one in which the cue was present, and therefore would encourage participants to predict that the outcome would occur, the similar learning vignette would have been one in which the outcome (and the cue) did not occur. If predictions are derived only from abstract contingency knowledge, the effect of the cue on predicted outcome likelihood shouldn't differ appreciably between novel and similar test scenarios. However, to the extent that predictions are also influenced by what happened in similar, prior instances, the effect of the cue should be attenuated in similar test scenarios (because instantiated and abstract knowledge are set in opposition).

Method

Participants. 52 McMaster University students participated for course credit or $10 cash. Participants ranged in age from 17 to 27 years (mean 18.3), and English was their primary language. Participants were assigned randomly and relatively equally to one of three different objective contingency conditions.

Stimuli. The current experiment was essentially identical to the normal exception condition of Experiment 1. Participants completed the same four separate contingency learning tasks, where each task involved presentation of 20 learning vignettes, followed by the outcome prediction and contingency rating exercises. Here, though, the objective contingency presented to participants was either strong (i.e., $\Delta P = 1.0$), moderate (i.e., $\Delta P = .6$), or weak (i.e., $\Delta P = .2$). Vignettes were created for the strong contingency condition first, and then the value of the cue or outcome was reversed in some vignettes as needed in order to reduce the contingency for the moderate and weak conditions. Each participant saw the same contingency strength in all four tasks, and the frequency of each type of learning item for the three contingency conditions is shown in Table 3.

Table 3
Frequency of Learning Item Types for Strong, Moderate and Weak Contingency (ΔP) Conditions in Experiment 2

	Strong (ΔP = 1.0)		Moderate (ΔP = .6)		Weak (ΔP = .2)	
	Outcome					
Cue	+	–	+	–	+	–
+	10	0	8	2	6	4
–	0	10	2	8	4	6

Note. Plus and minus symbols signify presence or absence, respectively.

The only other difference between this experiment and the normal exception condition of Experiment 1 concerns the test scenarios used in the outcome prediction task. In Experiment 1, participants received 4 novel test scenarios (2 C+ and 2 C–) in each task. Two of those scenarios were used in Experiment 2 as the novel test items (1 C+ and 1 C–). Two additional test items (1 C+ and 1 C–) were created for each of the four tasks, and these were designed to be similar in content and detail to one of the learning vignettes. The vignettes to which similar test items were paired were always those that were identical across all three contingency conditions, and therefore were either the C+/O+ or the C–

/O– type. Most critical, though, similar test items were always paired with vignettes having outcome values that opposed the response dictated by the cue in the test item. The two similar test items, along with the learning vignettes to which they were paired, for the "rude to relatives" contingency learning task are provided in Appendix B.

Procedure. The procedure was identical to that in Experiment 1. For the outcome prediction tasks, novel and similar test items were mixed and presented in random order.

Results

Analysis. Data analysis followed a protocol similar to that used in Experiment 1. Average contingency ratings across the three contingency conditions (ΔP = 1.0, 0.6, or 0.2) were analyzed using a one-way ANOVA. Mean outcome predictions were submitted to a mixed 3-way ANOVA with cue value (present or absent) and test item type (novel or similar) as within-subject variables, and condition as a between-subjects variable.

Contingency Ratings. Participants were presented with either strong, moderate or weak objective contingencies, and their ratings of those contingencies are shown in Table 4. Analysis of these ratings, which also happen to resemble the actual ΔP values, suggests that participants were quite sensitive to the objective contingency. Specifically, ratings differed across the three conditions ($F(1, 49) = 55.29$, $MSE = 360.9$, $p < .001$, $\eta_P^2 = .693$), and the Tukey test confirms that ratings were greater in the strong than in the moderate, and greater in the moderate than in the weak conditions ($p < .001$).

Table 4
Mean Contingency Ratings for Participants Receiving Strong, Moderate or Weak Objective Contingencies in Experiment 2

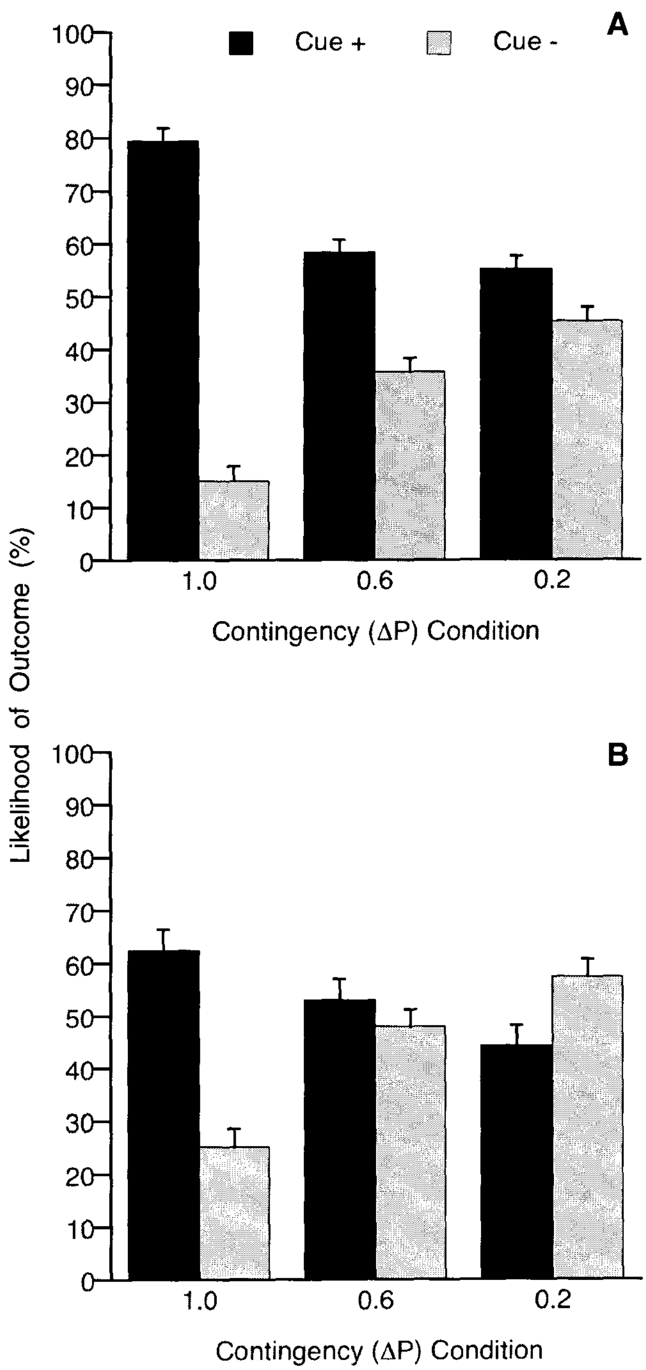
Objective Contingency	Rating of Cue-Outcome Contingency
Strong	
$\Delta P = 1.0$	84 (4.6)
Moderate	
$\Delta P = 0.6$	55 (4.6)
Weak	
$\Delta P = 0.2$	17 (4.5)

Note. Standard errors of the means are shown in parentheses.

Outcome Predictions. Although analysis of the contingency ratings confirms participants were quite sensitive to the different objective contingencies, their outcome predictions can provide converging evidence. Figure 2 shows the mean predicted likelihood of the outcome as a function of whether the cue was present or absent, separately for novel and similar test items, in all three contingency conditions.

During learning, participants received information that the presence of a cue weakly, moderately or strongly predicted the occurrence of an outcome. To the extent that this was learned and used in making predictions, then the outcome should be rated more likely to occur in the presence than absence of the cue, and this difference or "cue effect" should increase with the strength of the objective contingency. Expectedly, the analysis confirmed not only a main effect of cue value ($F(1, 49) = 70.49$, $MSE = 326.1$, $p < .001$, $\eta_P^2 = .590$), but more importantly, that this effect interacted with contingency condition ($F(1, 49) = 38.59$, $MSE = 326.1$, $p < .001$, $\eta_P^2 = .612$). To establish that the cue effect differed between each level of objective contingency, data from the strong and moderate conditions, and from the moderate and weak conditions, were separately submitted to the same mixed ANOVA procedure used for the main analysis. That analysis indicated that the cue effect did interact with both the strong/moderate objective contingency ($F(1, 32) = 30.71$, $MSE = 377.4$, $p < .001$, $\eta_P^2 = .490$) and with the moderate/weak objective contingency conditions ($F(1, 33) = 8.22$, $MSE = 255.7$, $p < .01$, $\eta_P^2 = .199$).

Figure 2


A
B
Cue +
Cue -
Likelihood of Outcome (%)
0
10
20
30
40
50
60
70
80
90
100
1.0
0.6
0.2
Contingency (ΔP) Condition

Mean outcome likelihood (+ *SE*) that participants predicted in novel (Panel A) and similar (Panel B) cue-present and -absent test scenarios, for three learning conditions differing in contingency strength in Experiment 2.

The most critical effect of interest from the outcome prediction task, however, concerns not whether the cue effect is modulated by objective contingency, but whether the effect of the cue is attenuated by a similar item from learning that suggests an opposing response. Comparing Panel A and Panel B of Figure 2, it appears that the effect of the cue in similar test items is reduced compared to that in novel test items, at each level of the contingency condition. In the main analysis, this pattern would correspond to a cue value X test item type interaction, in the absence of a three-way interaction with condition. The main analysis confirmed both the significant two-way interaction between test item type and cue value ($F(1, 49) = 36.34$, $MSE = 180.2$, $p < .001$, $\eta_P^2 = .426$) and the absence of a three-way interaction between those factors and condition ($F < 1$).

Discussion

The current experiment demonstrated two important findings. The first concerns how prior experience drives our responses to new situations. Of course, patterns, or relationships or contingencies that we learn shape how we respond, and the stronger these relationships are, the more they determine our response. However, individual prior experiences shape responding not only through their evidentiary contribution to abstract knowledge about relationships or contingencies; memory for those experiences that are similar to the current one can also be selectively retrieved and directly influence what we expect or how we behave in the current situation. While the notion that specific prior instances shape responding is certainly not new, what is perhaps most interesting here is that the influence of similar prior instances was independent of contingency strength. It seems intuitive that if one knows some predictive relationship to be true, but also knows that relationship is imperfect, or of moderate or even weak strength, then they might be inclined to consult other sources such as similar prior instances in formulating a response. However, it seems equally intuitive that as the predictive relationship becomes very strong, or even perfect, there would be less need to or value in consulting other sources of knowledge (i.e., querying multiple sources quite likely to provide the same answer would be inefficient). The results of the current experiment flatly deny that intuition, as participants' predictions about what would happen in new situations were equally influenced by a similar instance encountered during learning, regardless of whether the objective contingency they were presented was perfect, of moderate strength or very weak. Finally, this experiment also replicates previous findings described in Skye (2007) demonstrating that, despite rather substantial differences between the current and traditional HCL tasks in the qualitative nature of the stimuli, performance is very much a function of ΔP or the objective contingency, as is typically observed.

General Discussion

Adjusting for Evidential Quality

Much of the literature in HCL is concerned with understanding how it is that we extract information about two (or more) variables from experience and put this information together in a form that reflects the extent to which one variable predicts the other. According to associative accounts of HCL, such as those based on the Rescorla-Wagner model (1972), data from experience produces contingency knowledge according to two key principles:

1) Early trials are more heavily weighted than later trials
2) Surprising experiences cause larger changes in the state of one's knowledge than anticipated ones

Inferential reasoning accounts have argued that all contingency learning is not equal, and that general knowledge or models of the world that the learner applies will affect the process of translating data into contingency knowledge. Aside from these noted constraints, though, both perspectives seem to regard any one piece of evidence as just as good as another. In fact, in most HCL tasks, any one piece of evidence has to be just as good as another because they are often virtually indistinguishable.

While homogeneous sets of evidence can easily represent the statistical properties of experience, they simply don't capture the richness and complexity of much everyday experience—precisely the richness and complexity that contributes to vivid episodic memory, engages our personal theories and background knowledge, invites us to impose interpretations or draw inferences, and ultimately regard one experience as very different from another in meaning and value. As the first experiment here demonstrates, when the experiences that form the basis for learning are unique, meaningful and highly descriptive, participants evaluate the quality of different pieces of evidence and weight them accordingly.

Evidential quality is a rather broad notion, though, and the current experiment focused on one particular dimension of quality—namely how informative each experience was about the relationship between two variables. Specifically, 'poor quality' evidence was that in which other, more potent forces appeared to determine the outcome, leaving little or no opportunity for the cue (even if predictive) to exert its influence. Modulating evidential quality in this way rests on the basic notion of competition among cues, predictors or causal forces—a notion that lies at the very heart of associative learning perspectives. That participants here discounted evidence where, for example, the occurrence of an outcome was unsurprising given the operation of potent determinants other than the cue, is reminiscent of well-known associative learning phenomena like *overshadowing* (Pavlov, 1927) and *blocking* (Kamin, 1968).

At the same time, the current demonstration of discounting, or weighting according to informative value, differs from associative views and investigations of cue competition. Of course, there is a procedural difference. Typically, the cues that compete in associative tasks have learning histories that are well controlled and established during the session, whereas here, this was true only of the learning history of the focal cue (and surely, even the focal cue was not free of prior associations at the outset). The learning histories of all competing predictors or causes here were not adjusted or manipulated at all during the task and so were those shaped by the learner's everyday experiences. This procedural difference though, is less interesting than the more fundamental difference in what can and does act as a competitor.

The Rescorla-Wagner (R-W) model is the most well known model of associative learning, and the one often applied to HCL. Tangen & Allan (2003) provide a nice description of the essence of the R-W model:

> There is a limit (λ) to the amount of predictive strength that an outcome can support. This limited amount of predictive strength is allocated among all cues present on the trial. If one cue acquires predictive strength, then all other cues that are present at the same time must get less.

As this description implies, what antagonizes cue X in becoming associated with an outcome O are other cues (eg. Y or Z) associated with the same outcome. In other words, to be a competitor to an X-O association, something must also take the form of a cue-O association.

It is certainly likely that some of the competitors built into the discountable exception stimuli (see Appendix A) used here are "cues" or predictors that, through participants' personal experience, had become associated with the outcomes used here. For example, the position that participants possessed a pre-existing association between drunken sports fans and the occurrence of rude behavior is one many would surely endorse. On the other hand, it is not necessarily the case that participants had similarly established prior associations between, say, rudeness directed at cantankerous, old male drivers. Even with no such prior association, what participants almost certainly did have, though was a great deal of knowledge about human behavior, and a capacity to draw inferences and generate theories, that afforded some determination about the informational value of evidence. They could know that most people have limited patience, that they may behave uncharacteristically when frustrated, and that they may lash out towards someone whose is being abusive. Using this knowledge, they could speculate that Graham's rudeness to the old man who hit his car resulted because he'd had enough of the verbal abuse and of trying unsuccessfully to reason with him, and that these forces simply overwhelmed his tendency to be polite to non-relatives.

In the current work, then, some of what appears to be competing against cue X becoming associated with outcome O is very likely inferences or theories generated from domain-specific knowledge. At present, this is a conceptual feature that seems absent from associative accounts of HCL. Moreover, it is a feature of HCL that seems unlikely to be handled by adjusting or adding parameters in existing models, because it is a fundamentally different form of competition—it is competition from the top down, not from the bottom up. This demonstration of top-down competition certainly resonates with inferential reasoning or higher order cognitive accounts of HCL, although the current work focuses on much more domain-specific influences that interact at the level of individual learning items. To provide further evidence that inferential reasoning and theory generation, rather than just prior associations, are behind the observed discounting effect, one avenue for future investigation would be to investigate whether the same effect is observed with stimuli whose structure is the same but whose specific content is unfamiliar, perhaps from a different or fictional culture.

While participants were less influenced by learning items in which potent determinants of behavior other than the cue were operating, there is some evidence that they did not fully discount these items either. In the discountable exceptions group, the objective contingency (ΔP) was 0.4. If the six exception items were fully disregarded, only 7 C+/O+ and 7 C–/O– items would remain, and the value of ΔP would increase to 1.0, or a perfect relationship. In Experiment 2, one of the three groups was presented contingencies with a ΔP value of exactly 1.0. Comparing their contingency ratings and outcome predictions to those in the discountable exceptions group of the current experiment, then, provides some indication of the extent to which they disregarded those exception items. It seems that the discountable exceptions group perceived the contingency to be weaker than those in the second experiment who received an objectively perfect contingency, suggesting that they did not completely disregard the poor quality exceptions they received. In other words, poor evidence is poor evidence, but it is not no evidence. In fact, this is probably quite reasonable. While we surely expect many people would be polite to an aunt who is near death, we can also imagine someone who dislikes or disrespects his relatives so strongly that he is not—and to the extent that Graham is not this person, the instance does say something about how predictive interacting with relatives is of his rudeness.

Finally, quality is surely a multi-faceted property of evidence, of which informative value or relevance for a specific purpose is only a part. In a review of the literature on persuasion and source credibility, Pornpitakpan (2004) describes evidence that under many circumstances, more credible sources have greater influence over our thoughts and actions. Further, it is noted that while expertise and trustworthiness are key components of source credibility, many attributes of the source, the receiver and the message itself can interact to influence persuasiveness. Aside from source credibility (which also, incidentally, would invite issues related to source recollection), we also tend to be more persuaded by and believe in the truth of familiar experiences (Begg, Anas & Farinacci, 1992) and of relatively detailed or vivid accounts from memory (e.g. Talarico & Rubin, 2003; Bell & Loftus, 1988, 1989). It is anticipated that similar adjustments in the weight of evidence would be observed if the task described here were to manipulate evidential quality along these or other important dimensions.

Instantiated Knowledge Effects

Experiment 1 demonstrated that, when learning items are qualitatively variable and have meaningful, individual identities, their contribution to the perceived contingency varies according to interactions with domain specific knowledge. The second experiment demonstrates that when learning items have those properties, they can also individually influence predictions or expectations in future situations separate from their contribution to any perceived contingency. Specifically, participants in Experiment 2 were asked to learn contingencies and then predict the likelihood of the outcome in new scenarios. When these scenarios were novel or dissimilar to those experienced during learning, predictions of the outcome given the presence or absence of the cue appeared to be strongly influenced by the cue-outcome contingency. In contrast, when the scenarios were similar to one encountered during learning, in which the cue (and outcome) had the opposite value, predictions reflected an influence of both the contingency and the opposing information from a similar prior instance. Moreover, there was no evidence that the effects of these two sources of information on outcome predictions were interactive.

Demonstrating that prediction responses are influenced by a similar prior instance, in addition to contingency knowledge, is not necessarily problematic for associative views of HCL. While such accounts do assume that responding is driven only by knowledge of cue-outcome associations (e.g., Vadillo & Matute, 2007), all that would be needed to account for the instance effect is to assume that extra-cue attributes of the item also developed some association with the outcome.[5] To illustrate, consider the first example provided in Appendix B. In the learning item, Graham is not interacting with a relative (C–), and he is not rude (O–). This information would contribute to the perceived C/O contingency. However, something about "crowded city busses with few or no available seats" could also become associated with not being rude (O–). For the purposes of this illustration, it makes little difference whether the crowded busses are regarded as a second cue, or part of the background context—in the R-W model, associations can develop for either one. What is important is that, when later confronted with the similar test scenario where Graham and his sister take the bus to go skating, two relevant associations could influence predictions: an association between relatives and rudeness, and one between crowded busses and the absence of rudeness.

However, that the observed instance effect did not interact with strength of the C/O contingency certainly does appear to be problematic for associative accounts. Recall from the preceding section that the R-W model is one of competition—different cues or sources of information compete with one another for a limited amount of associative strength supported by a particular outcome. As one cue becomes more strongly associated with an outcome, other cues must

[5] Of course, given that all learning stimuli in the current task are unique in their contextual details, such an argument would also require that association develop after just one exposure or trial. Single trial conditioning is not unheard of, and the R-W model certainly predicts some change in associative strength between a second cue and the outcome or the context and the outcome after one trial (the amount of change would depend on several parameters), although it would not yet be in a stable or asymptotic state.

receive less. In the second experiment, then, as the objective contingency (ΔP) between the cue and outcome increased from 0.2 to 0.6 to 1.0, the remaining associative power available for additional cues or contextual information should have decreased, producing an interaction between the effect of the instance and the objective contingency.

The instance effect observed here, then, seems unlikely to be explained by the formation of associations between additional cues or the context and the outcome, and suggests that participants are encoding their learning experiences in memory as rather rich, individual episodes that can be selectively retrieved by similar instances. Anecdotal evidence from post-experiment interviews with some participants also suggested they were deliberately trying to put together information from both the retrieved learning episode and the perceived contingency in predicting the outcome. However, the present results don't necessarily discriminate between the possibility that only instances are stored in memory (e.g., a single system model like MINERVA 2, Hintzman, 1984) or that there are two somewhat separable knowledge systems, one that is exemplar-based and another that stores abstractions like cue-outcome associations (e.g. Julsin, Karlsson & Olsson, 2008; and see Shanks, 1995)

Unfortunately, it is somewhat difficult to evaluate how inferential reasoning or higher order cognitive accounts would respond to the instance effect observed here because such models often say little about encoding (e.g., Pineño & Miller, 2007; Vadillo & Matute, 2007). Instead, they focus more on the role of top-down influences in shaping how statistical information is combined and analyzed. There is reason to believe that similar top down modulation of the instance effect described here could readily be demonstrated, though, using a modified version of the current task; and this would clearly be more consistent with accounts that involve higher order cognitive processes as opposed to those that are only data driven. In what follows, a sketch of this modified task and the anticipated result are provided.

In some of my other, unpublished work, I provided participants with stories about fictitious characters and evaluated how they coordinated information about general behavioral trends and information from specific instances in predicting the character's future behavior. Like the current task, some prediction scenarios were similar to items previously encountered, but the important feature in these studies was that the dimensions of similarity were varied. In particular, prediction scenarios could resemble an old instance in the likely causal explanation for the behavior (deep similarity), or in the situation and contextual details (surface similarity). The data very clearly showed that predictions were more influenced by the prior instance when the items were similar on a deep level, even though there was less overall resemblance, suggesting that participants' knowledge and beliefs about the likely determinants of human behavior influenced the extent to which they applied information from prior instances. If the outcome prediction probes used in the current task were redesigned to distinguish between surface and deep similarity, and they produced a similar transfer pattern, this would provide some evidence that inferential reasoning processes intervene in how prior knowledge is used to guide responding.

Top-down and bottom-up influences in HCL

Across two experiments, certain patterns of performance were observed that do not easily fit into the classical associative, data-driven view of HCL. When evidence is rich and meaningful enough, the learner appears to consider its quality or value, given the knowledge and beliefs he or she already possesses, and weight it accordingly. Moreover, the learner's expectation about what will happen in future situations appears to be an additive compound of both abstract, contingency knowledge and memory for specific prior experiences. The current results are relevant to the debate between associative and inferential reasoning accounts of HCL, and seem most consistent with arguments that accounts of HCL will need to include both bottom-up and top-down processes (e.g., Vadillo & Matute, 2007; Shanks 2007). The question of how these are coordinated in guiding responding, then, becomes of key interest.

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*(3), 435-448.

Beckers, T., De Houwer, J., Pineño and, O., & Miller, R. R. (2005). Outcome additivity and outcome maximality influence cue competition in human causal learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(2), 238-249.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General, 121*(4), 446-458.

Bell, B. E., & Loftus, E. F. (1988). Degree of detail of eyewitness testimony and mock juror judgments. *Journal of Applied Social Psychology, 18*(14), 1171-1192.

Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of Personality and Social Psychology, 56*(5), 669-679. Brooks, L. R. (1978). Non-analytic concept formation and memory for instances. In E. H. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169-211). Hillsdale, NJ: Erlbaum.

Brooks, L. R., & Hannah, S. D. (2006). Instantiated features and the use of "rules.". *Journal of Experimental Psychology: General, 135*(2), 133-151.

Brooks, L. R., Norman, G. R., & Allen, S. W. (1991). Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology: General, 120*(3), 278-287.

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

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17*(4), 391-416.

Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology, 58*(4), 545-567.

Collins, D. J., & Shanks, D. R. (2002). Momentary and integrative response strategies in causal judgment. *Memory & Cognition, 30*(7), 1138-1147.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? studies with the wason selection task. *Cognition, 31*(3), 187-276.

Cox, J. R., & Griggs, R. A. (1982). The effects of experience on performance in wason's selection task. *Memory & Cognition, 10*(5), 496-502.

De Houwer, J. (2002). Forward blocking depends on retrospective inferences about the presence of the blocked cue during the elemental phase. *Memory & Cognition, 30*, 24-33.

De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and theories on human contingency learning. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 55B*(4), 289-310.

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

Dopkins, S., & Gleason, T. (1997). Comparing exemplar and prototype models of categorization. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale, 51*(3), 212-230.

Heit, E. (1998). Influences of prior knowledge on selective weighting of category members. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(3), 712-731.

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments & Computers, 16*(2), 96-101.

Juslin, P., Karlsson, L., & Olsson, H. (2008). Information integration in multiple cue judgment: A division of labor hypothesis. *Cognition, 106*(1), 259-298.

Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9-31). Miami, FL: University of Miami Press.

Keil, F. C. (. (1989). *Concepts, kinds, and cognitive development.* Cambridge, MA, US: The MIT Press.

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

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

Maxwell, S. E. (1980). Pairwise multiple comparisons in repeated measures designs. *Journal of Educational Statistics, 5*(3), 269-287.

Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(1), 37-50.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*(3), 207-238.

Murphy, G. L. (. (2002). *The big book of concepts.* Cambridge, MA, US: MIT Press.

Murphy, G. L., & Allopenna, P. D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 904-919.

Murphy, G. L. & Wisniewsi, E. J. (1989). Feature correlations in conceptual representations. In G. Tiberghien (Ed.), *Advances in cognitive science, vol. 2: Theory and applications* (pp. 23-45). Chichester: Ellis Horwood.

Pavlov, I. P. (1927). *Conditioned reflexes.* London: Clarendon Press.

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

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology, 34*(2), 243-281.

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

Shanks, D. R. (1995). Is human learning rational? *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 48A*(2), 257-279.

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

Shanks, D. R. (. (1995). *The psychology of associative learning*. New York, NY, US: Cambridge University Press.

Skye, A. (2007). Events come in all shapes and sizes: Contingency learning with heterogeneous and elaborate stimuli. Manuscript in preparation.

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science, 14*(5), 455-461.

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

Tangen, J. M., & Allan, L. G. (2004). Cue interaction and judgments of causality: Contributions of causal and associative processes. *Memory & Cognition, 32*(1), 107-124.

Vadillo, M. A., & Matute, H. (2007). Predictions and causal estimations are not supported by the same associative structure. *The Quarterly Journal of Experimental Psychology, 60*(3), 433-447.

Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(1), 53-76.

Waldmann, M. R. (. (1996). *Knowledge-based causal induction*. San Diego, CA, US: Academic Press.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121*(2), 222-236.

Wason, P. (1966). Reasoning. In B. Foss (Ed.), New horizons in psychology. Middlesex, England: Penguin.

Whittlesea, B. W. A., Brooks, L. R., & Westcott, C. (1994). After the learning is over: Factors controlling the selective application of general and particular knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(2), 259-274.

Appendix A

Normal and discountable learning vignettes providing evidence against a contingency (3 C+/O– and 3 C–/O+) between target character's rudeness and the relatedness of people with whom he was interacting.

Normal Exceptions (6)

When Graham found out his cousins Ian and Molly were coming for the weekend, he told them that they could take his room and he would sleep on the couch downstairs.

Russell, a friend of Graham's from his lacrosse team, needed a staple gun. Graham knew his dad had one, but before he offered to lend it to Russell, he checked with his dad first to make sure it was ok.

Graham and the rest of his family were sitting down to dinner on Friday; his mom had brought home pizza. Graham wanted seconds, but he offered the last two slices to everyone else before taking them for himself.

One day, Graham was approached on campus by two Mormons who were canvassing and wanted to talk with him. Graham cut them off mid-sentence, saying he didn't care what they had to say and then he walked away.

Graham was waiting, along with several others, to order lunch at Burger King. Eventually, a girl behind the counter opened another till. She pointed to Graham and asked what he would like to order. Even though the guy beside him had been there first, Graham stepped right up and ordered his lunch.

Graham was sitting in a History lecture with Brody and Austin, a couple of his friends. Brody and Austin started chatting about their plans for the weekend. Graham joined in the conversation, and the three of them were rather disruptive while the retired guest lecturer was speaking.

Discountable Exceptions (6)

Graham went with his parents to visit his elderly Aunt who, sadly, was expected to die within the next 24 hours. When they got to the hospital room, Graham took off his ball cap and politely remembered to thank her for the birthday card she had sent him last month.

Graham's second cousin Doug was not only very rich, but he also had a lot of connections especially down at city hall where Graham was hoping to get a job for the summer. When Doug dropped by to visit Graham's dad, Graham jumped off the couch and offered to take Doug's coat and get them a drink.

Uncle Hugh was an usually strict and mean-tempered man who didn't hesitate to hit his kids even though they were now in their teens and taller than him. Graham had even heard that Children's Aid had been involved with the family years ago because of abuse issues. When Uncle Hugh reclined on the couch and said he wanted quiet to sleep, Graham, who was visiting with his cousin, immediately turned off the sound on the computer. They didn't turn it back on until Uncle Hugh woke up and went back upstairs.

On the day of the Super Bowl, Graham headed over to Jay's Sports Bar at 3 o'clock in the afternoon, even though the game didn't start until 6:00. By 7:00, Graham was so drunk that the bartender cut him off. When the bartender refused to give him another beer, he swore at her and called her a whore.

Graham was stopped at a red light when a car rear-ended him. It was clearly the other driver's fault, but when the elderly man got out of the car, he started berating Graham for being stupid and calling him a bad driver and several other nasty names. Graham tried talking calmly to him, but that didn't help. Finally Graham said, "Look, old man, I don't have to take your crap". He took the guy's license plate number so he could file a report and left him there.

In Chinese culture, it is really unacceptable to wish someone a happy birthday after the actual day has passed, but Graham didn't know this. Graham's good friend Kuan had had his birthday during reading week. When Graham got back from the break, he gave Kuan a call to wish him a belated happy birthday with an offer to take him out for a beer. It was only later that he heard from another friend that Kuan had been offended by it.

Appendix B

Similar outcome prediction test items, and the learning vignettes to which they were paired, for the task involving contingency between a target character's rudeness and the relatedness of people with whom he was interacting

Cue+ Test Scenario

Graham and his sister were going public skating. When they got on the bus that goes to the arena, it was really crowded and there was only one seat available. Graham...

a) sat down in the seat and made his sister stand the whole ride.
b) told his sister to take the seat and said he would stand.

Paired to the following learning vignette: Graham was taking the bus home from school one day. The busses were really busy, but he managed to get a seat. A few minutes after Graham got on, an elderly man got on the crowded bus. Graham got up right away and gave the man his seat.

Cue– Test Scenario

Graham was hanging out with some friends at their place. One of the girls was talking about these really cool pants she'd bought from Titles bookstore that day. She decided to put them on and show everyone. When she came out, Graham…

a) smiled at her and agreed that they were pretty cool looking.

b) told her that the pants made her butt look huge.

Paired to the following learning vignette: When Graham's sister was showing her mom the new jeans she'd bought, Graham told his sister that she looked fat.

Running head: LEARNING CONTINGENCIES BETWEEN SITUATIONS & BEHAVIORS

If the situation predicted behavior, would somebody learn it? An examination of learning contingencies between how people behave and the nature of the situation

Aimee L. Skye

McMaster University

Send correspondence to:

Aimee Skye

Dept. of Psychology, Neuroscience and Behaviour

McMaster University

Hamilton, Ontario, CANADA L8S 4K1

Email: skyeal@mcmaster.ca

Phone: (905) 525-9140, ext. 24824

Forthcoming In Hardy-Vallée, B., & Payette, N. (Eds.), *Connected Minds: Cognition and Interaction in the Social World.*

McMaster University

Department of Psychology, Neuroscience & Behaviour

May 15, 2008

Dr. Benoit Hardy-Vallée
Department of Philosophy
University of Toronto
Jackman Humanities Building
170 St. George St., 4th Floor,
Toronto, ON, M5R 2M8

Dear Dr. Hardy-Vallée :
I am completing a Ph.D. thesis at McMaster University entitled *Associative and non-associative performance phenomena in learning social contingencies from rich and heterogeneous stimuli* and I would like your permission to reprint the following journal article in my thesis.

> Skye, A. L. (in press). If the situation predicted behavior, would somebody learn it? An examination of learning contingencies between how people behave and the nature of the situation. In Hardy-Vallée, B., & Payette, N. (Eds.), *Connected Minds: Cognition and Interaction in the Social World.*

Please note that I am the author of this work. I am also requesting that you grant irrevocable, nonexclusive license to McMaster University and to the National Library of Canada to reproduce this material as a part of the thesis. Proper acknowledgement of your copyright of the reprinted material will be given in the thesis.
If these arrangements meet with your approval, please sign where indicated below and return this letter to me.
Thank you very much.

Sincerely,

Aimee L. Skye

PERMISSION GRANTED FOR THE USE REQUESTED ABOVE

Authorized by: Dr. Benoit Hardy-Vallée

Title: Editor

Date: May 16, 2008

Signature:

Abstract

We strongly believe people are stable in who they are and how they act, yet evidence of cross-situational consistency is notoriously elusive. Mischel (2004) argues that the predictability of behavior from situational attributes underlies this belief in consistency. While research has demonstrated situation-behavior contingencies are reflected in others' behavior, our sensitivity to these contingencies remains unclear. The current research evaluates acquisition of such contingencies using an associative learning procedure. Experiment 1 (N = 32) demonstrates good contingency learning when the relationships characterize individuals but not groups. Experiment 2 (N = 38) evaluates whether this difference can be explained by greater integrative information processing for more coherent targets (Hamilton & Sherman, 1996). Learning was not better overall for contingencies characterizing more than less coherent groups, but it was for particular contingencies. The results suggest we readily learn *if…then…* patterns in the behavior of others, and that information processing tendencies as a function of specific rather than general expectations of coherence may influence sensitivity.

If the situation predicted behavior, would somebody learn it? An examination of learning contingencies between how people behave and the nature of the situation

Many everyday activities involve dealing with people, and navigating these interpersonal interactions requires an ability to learn about others and use our knowledge to anticipate their behavior. It is of little surprise, then, that how we form impressions of others and the processes involved in person learning, memory and judgment are key issues of interest in social, cognitive and personality psychology.

Research shows that our impressions of others and our interpretations of their behavior are dominated by personality or dispositional information (e.g., Park, 1986; Humphrey, 1985; Miller, 1984; Pietromonaco & Nisbett, 1982; Jones & Harris, 1967). We also perceive a great deal of constancy in the people we come to know. In other words, we expect others to behave in a manner that is predictable from their disposition and highly consistent across different situations (e.g., Kunda & Nisbett, 1986; see Ross & Nisbett, 1991 for extensive review). Empirical efforts to document that others' behavior reflects such considerable and context-independent consistency, however, have fallen far short of what personal experience would suggest ought to be a straightforward enterprise (see Mischel, 1968 for a review). This discrepancy has generated much interest in the question of what, then, produces our entrenched belief that people differ markedly from one another, in ways that manifest themselves time (and situation) and again.

Mischel and his colleagues have offered an interesting response to this puzzle (Mischel, 2004; Mischel, Shoda & Mendoza-Denton, 2002; Cervone & Shoda, 1999; Mischel & Shoda, 1995). They argue that people are coherent and predictable, and we perceive them to be so, not because their actions are invariant across situations, but because each individual's actions are organized by stable *if...then...* relationships between attributes of the situation and how one behaves. In other words, neurotic Susan surely won't be uptight and anxious in *every* situation; instead, she may reliably act that way whenever an authority figure is present. This proposal that situation-behavior contingencies are central to the essence of individuals and our impressions of them is certainly congruent with extensive research demonstrating how profoundly the situation influences people's thoughts, beliefs and actions (e.g., Asch, 1955; Darley & Latané, 1968; Milgram, 1963; and see Ross & Nisbett, 1991 for a review). Moreover, evidence that people can be reliably characterized by contingencies between particular situational attributes and how they act is slowly accumulating (Fleeson, 2007; Fleeson, 2001; Shoda, Mischel & Wright, 1994, 1993a; Mischel & Peake, 1982).

However, to account for our pervasive belief that individuals are consistent, it would not only be necessary that the behavior of individuals' is contingent upon the situation, but also that we readily become sensitive to those contingencies through our experience. Putting aside issues of contingency strength and opportunity for learning, our well-documented insensitivity to the effects of situational forces on others' behavior (see Ross & Nisbett, 1991 for an extensive review and discussion) raises question about how readily we might learn contingencies that involve that very class of information.

Empirical demonstrations that we do become sensitive to interactions between the nature of the situation and how others behave are scarce, although there are a few notable exceptions (Kammrath, Mendoza-Denton & Mischel, 2005; Shoda, Mischel & Wright, 1993b, 1989). These studies demonstrate sensitivity primarily by showing that impressions and inferences we form about a target differ as the content of the *if...then...* relationship characterizing that target changes. The research presented here extends this sparse body of work by more directly evaluating the extent to which we learn situation-behavior contingencies that characterize others.

Relationships such as *Susan is nervous when around authority figures* are much like other contingencies in our environment, such as chest pain often signals a heart attack or higher grades are associated with better work ethic. Given this similarity, the current research employs an associative-learning procedure typically used to study contingency learning in domains outside person perception. The procedure, introduced elsewhere (Skye, 2007), involves presenting several unique vignettes that each depict how a target behaves and the situation in which the behavior occurred, and that collectively reflect some contingency between those two variables. Sensitivity to the contingency is then assessed by examining what effect the situational variable has on perceived likelihood of the behavior in novel vignettes and also, explicit knowledge of the predictive situational variable.

The first experiment evaluates whether our ability to learn situation-behavior contingencies differs when those contingencies are expressed through the actions of one or many people. Good contingency learning in the individual condition would certainly provide evidence that we are sensitive to *if...then...* signatures reflected in the behavior of people we encounter. Furthermore, if sensitivity to the same contingencies is weakened when they reflect the actions of a group, this could suggest that we are especially tuned to learn about situation-behavior relationships within individuals.

Superior contingency learning from information describing individuals compared to groups could result from differences in information processing for the two types of social targets. Hamilton and Sherman (1996) review considerable impression formation research that suggests we process information about individuals in a more integrative manner, and integrative processing is surely involved in contingency learning where discovery of patterns across multiple events is key. More generally, Hamilton and Sherman argue that integrative processing is a positive function of the unity or coherence we expect from any social target. So, to evaluate whether integrative processing as a function of the social target's perceived unity might explain differences in our ability to learn *if...then...* relationships reflected by individuals and groups, the second experiment examined whether contingency learning differed when the vignettes described members of more and less cohesive groups.

Experiment 1

This experiment evaluated how well we learn situation-behavior contingencies expressed by individuals or groups. Participants received vignettes describing behaviors in specific contexts, and the behavior was perfectly contingent on a situational attribute. All vignettes referred to the same person in the individual target condition. In the group target condition, each vignette mentioned a different name, implying the descriptions reflected a rather random group of people. Contingency learning was evaluated by examining whether the presence of the situational cue increased expectations that the behavior would occur in novel scenarios, and by asking whether participants could name the situational cue that predicted the target's behavior. Increased expectations of the outcome in the presence of the cue and explicit knowledge of the situational predictor would indicate participants did learn situation-behavior contingencies. And, if we are particularly inclined to learn about *if…then…* signatures for individuals, evidence of learning on both measures should be stronger in the individual than in the group condition.

Method

Participants. 11 and 21 McMaster University students, in the individual and group target conditions respectively, participated for course credit or $10 cash. Participants ranged in age from 18 to 35 years (mean 20.4), and English was their primary language.

Stimuli. In the individual target condition, participants received 20 vignettes describing one person's behavior in specific and unique contexts. Across the vignettes, there was a contingency between a situational cue (C) and a behavioral outcome (O). For example, one person tended to be rude to his relatives but not to non-relatives. Examples of the vignettes describing that person include:

> When Graham came home from his job at the recreation center, Graham flopped down on the couch, grabbed the TV remote and switched channels from the show his brother Peter was watching to the football game without asking.
>
> Graham called VISA about a charge that appeared on his credit card bill. The customer service agent put him on hold, but then she forgot about him. After holding for 20 minutes, Graham hung up and called back. When he got back through to the same agent, he calmly asked what had happened and was very understanding when she explained her error.

As these examples illustrate, each vignette conveys whether the cue was present (e.g., relative vs. non-relative) and whether the outcome occurred (e.g., rude vs. polite). The objective relationship between the cue and outcome can be quantified using ΔP, which reflects the difference in the outcome probability when the cue is present and absent (Allen, 1980). To create maximal learning conditions, the cue-outcome contingency was always perfect (i.e., $\Delta P = 1.0$). In other words, the behavior occurred if and only if the situational attribute was present.

Contingency learning was assessed using two additional vignettes describing novel scenarios in which the cue was present or not. Participants were given two options representing the presence and absence of the behavioral outcome, and they indicated the probability that each would occur in that scenario (total forced to sum to 100). For example, in the cue-absent (C–) vignette below, the absence of the behavioral outcome (O–, or Option B) should be rated more likely than its presence (O+), given that rudeness is contingent upon interacting with relatives.

> Graham bought a stereo from a store in Toronto and he borrowed his friend Justin's car to go and pick it up. When he was finished, Graham...
>
> (A) returned his friend's car without replacing any of the gas he'd used.
>
> (B) filled up his friend's car with gas before returning it.

Participants were also questioned about their knowledge of the predictive situational variable. For example, they were told “there was actually a simple rule that determined whether Graham was polite or disrespectful”, and were asked what that rule was and told to guess if necessary.

The session included three additional replications of the learning and test phases just described. These were identical in procedure and structure of the materials, and differed only in that each used new vignettes to describe a new target character whose behavior reflected a new *if...then...* contingency. The three other contingencies involved being active only when it was early, being helpful only when one was available, and being talkative except when a particular individual was present.

The group target condition was identical to the individual target condition, except for a seemingly minor change to the stimuli. To create the sense that the task involved learning about a group of people and predicting its members’ behavior, every vignette and test scenario referred to a different person (e.g., Benjamin, Nicholas, Duane). Names were never repeated across replications, and two replications used all males and the other two used all females. Participants likely regarded these groups of people as fairly random, as no information that defined the nature of the groups was ever provided (although they could have imposed idiosyncratic definitions). Finally, the question probing explicit knowledge of the contingency in this condition was altered to refer more generically to a group member.

Procedure. Participants were told they would be given several behavioral descriptions and then would answer questions based on that information. Order of the four replications was randomized, as was order of the learning and test vignettes within each replication. Vignettes were displayed on a computer screen and narrated through an audio file. In each replication, presentation of the vignettes was immediately followed by the outcome prediction task. Questions probing explicit knowledge of the situation-behavior contingencies for all four replications occurred at the end of the session.

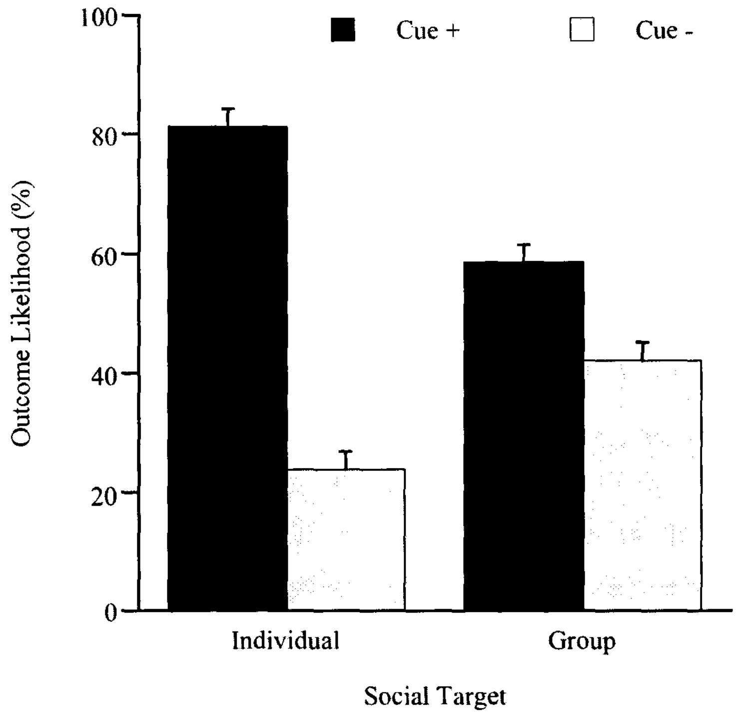
Results

For the prediction task, probabilities assigned to the behavioural outcome option (O+) were submitted to a mixed 3-way ANOVA with replication (4 different contingencies) and situational cue status (present or absent) as within-subject variables, and social target (individual or group) as a between-subjects variable. Responses to questions about contingency awareness were independently coded as correct (1) or incorrect (0) by the author and a research assistant, who were blind to the outcome prediction data. Disagreements were resolved by discussion until consensus was reached. These scores were summed across replication, and the number of contingencies (out of 4) that each participant could describe was analyzed using an independent t-test between individual and group conditions.

To the extent that participants learned the situation-behavior contingencies, their predictions in novel scenarios should reflect that the behavioral outcome is more likely when the situational cue is present. Moreover, if learning differed from information depicting individuals or groups, this effect of the situational cue should differ between those conditions. Figure 1 shows mean outcome likelihood predictions as a function of cue status (collapsed across replication) for each condition. The analysis confirmed a significant interaction between situational cue status and condition ($F(1, 30) = 31.14$, $MSE = 1817.0$, $p < .001$, $\eta_P^2 = .509$), and this differential effect of the cue between conditions did not interact with replication ($F < 1$, $MSE = 1817.0$).

Post hoc analyses on the effect of the cue within each condition were conducted using paired t-tests with a Bonferroni correction, where $\alpha=0.05/2 = 0.025$ (see Maxwell, 1980 for evidence this procedure is favorable for repeated measures designs). These analyses revealed a significant effect of cue status on outcome predictions in both the individual ($t(10) = 10.53$, $SE = 5.46$, $p < .001$, $d = 15.5$) and group conditions ($t(20) = 3.69$, $SE = 4.48$, $p < .01$, $d = 6.1$). Thus, the two-way interaction in the main analysis was significant because the situational cue affected expectations about the outcome likelihood in both conditions, but it did so much more strongly when participants learned about individuals than groups.

Figure 1



Mean predicted likelihood of the outcome (+ *SEM*) in novel scenarios as a function of cue presence and absence for participants who received information describing individuals or (random) groups.

Responses to questions probing explicit contingency knowledge were consistent with the outcome prediction data. Participants who learned about individuals could verbalize, on average, 3.8 of the 4 *if...then...* relationships they were presented, while those who learned about groups could verbalize only 1.9 relationships ($t(30) = 7.17$, $SE = 0.27$, $p < .001$, $d = 2.9$).

Discussion

Participants appeared to be quite good at learning that an individual's behavior can be contingent upon the situation or context, but they are much less good at learning these predictive relationships for the behavior of groups (or their members). This was evident using both more and less explicit measures, which argues against the interpretation that these two conditions differed not in extent of learning but rather in willingness to express learning. Furthermore, because the evidence and learning opportunities were identical, the discrepancy in situation-behavior contingency learning from individuals or groups likely resulted from differences in information processing for the two social targets. Hamilton & Sherman's (1996) proposal that we engage in more integrative processing for information about individuals than groups would certainly be consistent with the current finding that discovery of contingencies is also better from information about individuals.

Experiment 2

Experiment 1 demonstrated that presenting information about individuals or groups substantially affects sensitivity to relationships involving behavior contingent upon the situation. Many other differences in how we process information about individuals and groups, and what we learn about them, have been reviewed recently by Hamilton and Sherman (1996). These authors argue the key to understanding these differences is that individuals and groups differ in perceived unity—meaning they are entities we expect to be more and less coherent (respectively)—and that our expectations of unity for a social target influence our information processing style and therefore our knowledge. If perceived unity shapes social information processing, it should be possible to invoke different degrees of integrative processing for groups that we perceive to be more and less coherent. And, if integrative information processing facilitates contingency learning, sensitivity to situation-behavior contingencies should be better from more coherent groups. The current experiment, then, compares situation-behavior contingency learning when the information describes a random group, as in the previous experiment, or a relatively more coherent group of people who share ethnicity and a common ancestral history.

Method

Participants. 20 and 18 McMaster University students, in the random and ancestral group conditions respectively, participated for course credit or $10 cash. Participants ranged in age from 17 to 38 years (mean 18.9), and English was their primary language.

Stimuli. All stimuli used for the random group condition were identical to those of the group condition of Experiment 1. The stimuli used for the ancestral group condition were also identical, with the following exceptions. Participants were told they would receive descriptions of people from the same ethnic group and the vignettes in each replication were preceded by a paragraph providing a short history of that group. For instance, the following paragraph introduced the "if relative, then rude" replication:

> The Gwimburran People were originally nomads who lived in the Balkans. For several decades, various foreign empires fought for control over this area. As these empires came into power, they would often remove the nomads from the Balkans and ship them to areas in the newly discovered land of North America, some of which later formed parts of Ontario and other Canadian provinces. Today, certain areas of Toronto are home to small groups of people who are descendents of the Gwimburran nomads. The information you are about to hear will describe the behavior of several people living in the Toronto area whose ethnic background is Gwimburran.

The only other modification to the ancestral group stimuli was that all names (the target character and supporting actors in the vignette) were changed to ethnic names consistent with the ancestral history (e.g., Lovro, Aladar & Otilia were some names used in the "if relative, then rude" replication)

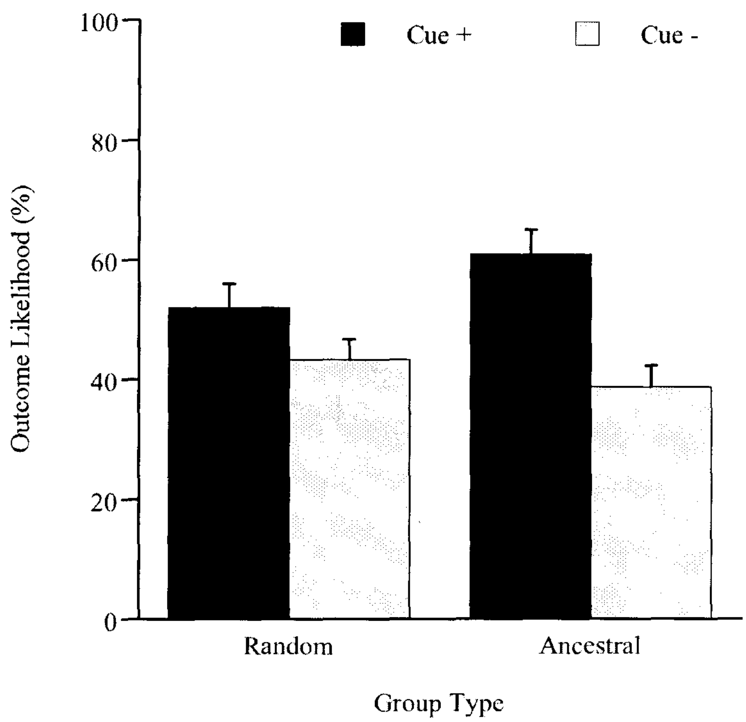
Procedure. The procedure was identical to that in Experiment 1.

Results

Data analysis followed the same protocol used for Experiment 1, with a Geisser-Greenhouse correction to adjust for violations of the sphericity assumption. For outcome predictions, the analysis revealed a significant main effect of situational cue status ($F(1, 36) = 11.73$, $MSE = 1554.78$, $p < .01$, $\eta_P^2 = .246$), which did not interact with group type ($F(1, 36) = 2.15$, $MSE = 1554.78$, $p = .15$). Regardless of whether participants learned about random or ancestral groups, the likelihood of the outcome was perceived to be similarly greater when the situational cue was present. Figure 2 shows mean outcome likelihood predictions as a function of cue status (collapsed across replication) for both conditions. The extent to which participants in either condition here learned the contingencies appears quite similar to that in the group condition of Experiment 1.

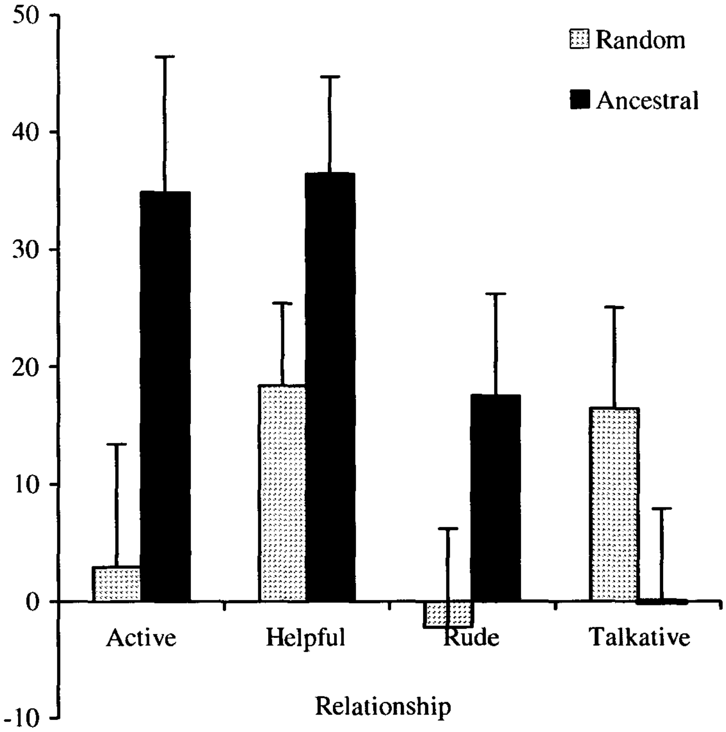
The finding of similar contingency learning between the random and ancestral group conditions was unexpected. However, it is possible that contingency learning from the two types of groups was different, just not consistently across all four replications. Indeed, the three-way interaction between replication, cue status and condition was at significance ($F(1, 36) = 4.12$, $MSE = 1496.3$, $p = .05$, $\eta_P^2 = .102$). To explore this interaction, post hoc analyses on the effect of the cue within each replication for each condition were conducted using paired t-tests with a Bonferroni correction (α=0.05/8 = 0.006). In the random group condition, the cue had no significant effect on outcome predictions in any replication. In contrast, in the ancestral group condition, the cue had a significant or nearly significant effect in two of four replications ($t(17) = 4.39$, $SE = 8.29$, $p < .001$, $d = 5.3$ for available/helpful, and $t(17) = 3.01$, $SE = 11.60$, $p = .008$, $d = 6.4$ for early/active). Figure 3 illustrates the effect of the situational cue on predicted outcome likelihood for each replication in both conditions.

Figure 2



Mean predicted likelihood of the outcome (+ *SEM*) in novel scenarios as a function of cue presence and absence for participants who received information describing random or ancestral groups.

Figure 3



Effect of situational cue on perceived likelihood of behavioral outcome when learning about random or ancestral groups, shown separately for each replication (identified here by the behavior used in each if...then... relationship). Bars represent likelihood of outcome when cue was present minus that when cue was absent. Thus, larger cue effects indicate cues perceived to have stronger influence on behavior. Error bars represent standard error of the mean.

The questions probing explicit contingency knowledge also showed no difference in contingency awareness between the two learning conditions ($t(36) = 0.56$, $SE = 0.32$, $p = .58$). The average number of *if...then...* relationships participants could verbalize was 1.6 and 1.8 (of 4) for those who learned about random and ancestral groups respectively.

Discussion

As in Experiment 1, participants who learned about random groups of people demonstrated a significant, but very weak, ability to learn objectively perfect situation-behavior contingencies. For those learning about groups with a shared ancestral history, better contingency learning was expected because the greater coherence of such groups ought to induce more integrative processing of the information. While this finding was not confirmed overall, contingency learning from the ancestral group was better for two of the specific relationships used—namely, the early/active and available/helpful relationships. Thus, there is some evidence that increased group coherence or unity affects contingency learning, presumably through inducing more integrative processing. At the same time, differences in the learnability of relationships varying widely in content suggest that additional influences, such as background knowledge or expectations, may interact with integrative processing (this notion is elaborated on in the next section).

General Discussion

The first experiment demonstrated we are quite good at picking up on *if...then...* signatures, or dependencies between behavior and context, that characterize individuals. This is exactly what would be expected if, as Mischel and colleagues propose, these signatures define a person's nature and underlie our sense that people are consistent in who they are and how they act (Mischel, 2004; Mischel et al, 2002; Cervone & Shoda, 1999; Mischel & Shoda, 1995). Although the signatures clearly do not reflect contextually invariant behavior, they very much reflect contextually anticipatable behavior. Investigations documenting entrenched beliefs in personal consistency may in fact be capturing this experience of others' behavior as predictable. Experiment 1 also demonstrates we are quite poor at picking up on the same situation-behavior dependencies when they characterize groups rather than individuals. In other words, not only are we very good at learning *if...then...* signatures that define individuals, we seem especially tuned to do so.

While the intent of this paper was not to systematically address the mechanism responsible for this dramatic difference in contingency learning, Experiment 2 evaluated whether differences in how we approach information about individuals and groups might play a role. From a review of impression formation research contrasting circumstances with group and individual social targets, Hamilton and Sherman (1996) argued we are more apt to process information about social targets integratively when they are perceived to have more unity or coherence, as are individuals relative to groups. Quite possibly, our participants regarded the individuals as more coherent than the random groups of people. Consequently, participants learning about individuals may have processed the vignettes in a much more comparative manner, facilitating discovery of the contingencies reflected therein.

To evaluate this account of the learning difference, the second experiment attempted to manipulate contingency learning from groups by altering their perceived unity. Participants learned the same contingencies from the same evidence, but about either random groups or relatively more coherent groups of people who shared an ancestral history. Weak evidence of contingency learning was found in both conditions, but critically, learning was not better from the ancestral groups. One interpretation of this result is essentially a manipulation failure. In other words, that ancestral and random group conditions didn't induce differences in information processing because participants perceived those groups as equally coherent. However, this seems unlikely because learning was better in the ancestral group for two of the four replications. Specifically, participants were better at learning members were active early in the day or helpful when available when they belonged to an ancestral rather than a random group. In contrast, learning that members were rude to relatives or talkative except when a specific person was around was not influenced by the type of group.

One way of understanding this result is that participants' background knowledge may have influenced contingency learning. Specifically, while each relationship certainly could be true of an individual, perhaps our experience suggests the early/active and available/helpful relationships are more apt to be true of groups, and particularly ancestral groups. Indeed, it seems somewhat easier to generate stories about how shared cultural or ancestral history might cause many people's behavior to adhere to these two contingencies. Moreover, if we consider real social groups whose members reflect these patterns, physicians, seniors, and members of a rowing team are just a few examples of "morning people" that come readily to mind. In contrast, stereotypical teenagers aside, similar examples for collections of people being rude to relatives seem less common and more effortful to generate.

While somewhat speculative, this explanation is consistent with evidence that personal beliefs and theories influence interpretation of statistical evidence (e.g. King & Koehler, 2000; Waldmann & Holyoak, 1992; Chapman & Chapman, 1971), and with the generally accepted view that human cognition is subject to top-down influences. Moreover, it seems reasonable that learning would be guided by existing knowledge. Imagine evaluating two different contingencies among grocery store customers: whether accepting a free sample predicts purchasing the product or whether a fire alarm predicts hurried exiting of the building. The group itself—shoppers in the same store at the same time—has relatively low unity. However, we would surely expect more coherence in the behavior of group members for the latter than the former contingency, because of our knowledge about the world. Consequently, we are probably also more likely to process behavioral evidence from different customers more comparatively when evaluating the latter contingency. The point of this example is to suggest that coherence of a social target is unlikely to be a rather abstract property that determines integrative processing in a context-free manner. Rather, it is likely that background knowledge causes us to expect specific expressions of coherence, and that the extent and focus of our integrative processing is shaped by these more specific expectations.

In conclusion, this research provides some of the first direct evidence that we are quite sensitive to relationships between the context and how a person behaves, which is central to the argument that predictability of human behavior from situational information fuels our (illusory) belief that individuals are highly consistent. Furthermore, as observed in the individual condition of Experiment 1, the robust learning of patterns in which behavior was contingent on the situation demonstrates a sensitivity to situational determinants of behavior that has traditionally been regarded as notably deficient (see Ross & Nisbett, 1991). Precisely why participants here are sensitive to situational influences on behavior is unclear, but may be related to key differences between the current and more traditional tasks. Here, evidence about a target takes the form of multiple, highly-specific instances rather than summaries of behavior tendencies or other more abstract forms of information. Moreover, the current task employs measures specifically designed to assess how situational information influences our perceptions of the behavior of others. Regardless, this research suggests that beliefs about our failure to appreciate situational determinants of behavior may need to be reconsidered.

References

Asch, S. E. (1955). Opinions and social pressure. *Scientific American, 193*(5), 31-35.

Cervone, D., & Shoda, Y. (1999). Beyond traits in the study of personality coherence. *Current Directions in Psychological Science, 8*(1), 27-32.

Chapman, L. J., & Chapman, J. (1971). Test results are what you think they are. *Psychology Today, 5*(6), 18-22, 106.

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of personality and social psychology, 8*(4), 377-383.

Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of personality and social psychology, 80*(6), 1011-1027.

Fleeson, W. (2007). Situation-based contingencies underlying trait-content manifestation in behavior. *Journal of personality, 75*(4), 825-862.

Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological review, 103*(2), 336-355.

Humphrey, R. (1985). How work roles influence perception: Structural-cognitive processes and organizational behavior. *American Sociological Review, 50*(2), 242-252.

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of experimental social psychology, 3*(1), 1-24.

Kammrath, L. K., Mendoza-Denton, R., & Mischel, W. (2005). Incorporating if...then... personality signatures in person perception: Beyond the person-situation dichotomy. *Journal of personality and social psychology, 88*(4), 605-618.

King, R. N., & Koehler, D. J. (2000). Illusory correlations in graphological inference. *Journal of Experimental Psychology: Applied, 6*(4), 336-348.

Kunda, Z., & Nisbett, R. E. (1986). The psychometrics of everyday life. *Cognitive psychology, 18*(2), 195-224.

Maxwell, S. E. (1980). Pairwise multiple comparisons in repeated measures designs. *Journal of Educational Statistics, 5*(3), 269-287.

Milgram, S. (1963). Behavioral study of obedience. *The Journal of Abnormal and Social Psychology, 67*(4), 371-378.

Miller, J. G. (1984). Culture and the development of everyday social explanation. *Journal of personality and social psychology, 46*(5), 961-978.

Mischel, W. (1968). *Personality and assessment.* Hoboken, NJ: John Wiley & Sons Inc.

Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology, 55,* 1-22.

Mischel, W., & Peake, P. K. (1982). Beyond déjà vu in the search for cross-situational consistency. *Psychological review, 89*(6), 730-755.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological review, 102*(2), 246-268.

Mischel, W., Shoda, Y., & Mendoza-Denton, R. (2002). Situation-behavior profiles as a locus of consistency in personality. *Current Directions in Psychological Science, 11*(2), 50-54.

Park, B. (1986). A method for studying the development of impressions of real people. *Journal of personality and social psychology, 51*(5), 907-917.

Pietromonaco, P., & Nisbett, R. E. (1982). Swimming upstream against the fundamental attribution error: Subjects' weak generalizations from the darley & batson study. *Social Behavior and Personality, 10,* 1-4.

Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology.* New York, NY: Mcgraw-Hill Book Company.

Shoda, Y., Mischel, W., & Wright, J. C. (1989). Intuitive interactionism in person perception: Effects of situation-behavior relations on dispositional judgments. *Journal of personality and social psychology, 56*(1), 41-53.

Shoda, Y., Mischel, W., & Wright, J. C. (1993a). The role of situational demands and cognitive competencies in behavior organization and personality coherence. *Journal of personality and social psychology, 65*(5), 1023-1035.

Shoda, Y., Mischel, W., & Wright, J. C. (1993b). Links between personality judgments and contextualized behavior patterns: Situation-behavior profiles of personality prototypes. *Social Cognition, 11*(4), 399-429.

Shoda, Y., Mischel, W., & Wright, J. C. (1994). Intraindividual stability in the organization and patterning of behavior: Incorporating psychological situations into the idiographic analysis of personality. *Journal of personality and social psychology, 67*(4), 674-687.

Skye, A. (2007). *Events come in all shapes and sizes: Contingency learning with heterogeneous and elaborate stimuli.* Manuscript in preparation.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121*(2), 222-236.

Author Note

Aimee L. Skye, Department of Psychology, Neuroscience and Behaviour, McMaster University.

The author is grateful to Lee Brooks for his guidance and review of this manuscript, and to Betty Howey for her invaluable assistance in materials development and data collection.

Correspondence about this article should be addressed to Aimee Skye, Department of Psychology, Neuroscience and Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ontario, L8S 4K1. E-mail: skyeal@mcmaster.ca

Concluding Discussion

Consider two different contingencies people might acquire from their experiences with the world. In the first, tenants or employees in multi-level buildings come to learn something about the reliability of the elevator system through repeated experiences in which they either press the call button or not, and the car arrives or it doesn't. In this version of the everyday contingency learning problem, each experience tends to be somewhat homogeneous—any press of a button is like the others, and so is the arrival (or not) of the car, except for perhaps the particular cast of characters or things that happen to be in it when it arrives. Even the background context of each experience is much the same as the next. Furthermore, although most of us could surely speculate a bit about interventions that could alter the level of elevator reliability we've come to expect (e.g., it is out of service for repair work), this is not a domain for which most of us have a rich network of background knowledge and beliefs.

Now, consider a second everyday contingency learning example: coming to believe that hot (i.e., attractive) girls tend to date hot guys. While there is room to quibble over the strength of this relationship, and counterexamples are readily available, phrases like "she's way out of your league" lend some truth to the perception, and indeed there is empirical evidence that attractiveness of romantic partners is positively correlated (Feingold, 1988). The experiences that give rise to this belief come from sources as disparate and varied as the fantasy land of

Barbie and Ken or Disney's Shrek movies, the family gatherings we attend, our own adolescent trials and tribulations, the extended social networks of friends and acquaintances we interact with, and the countless opportunities we have to "people watch" in shopping malls, salons, restaurants, parks, vacation resorts, etc. Moreover, every person we observe, whether girl or guy and whether hot or not, exhibits some uniqueness in their appearance, their actions, their nature or demeanor, and other attributes. Unlike the elevator reliability example, learning that hot girls tend to date hot guys arises out of experiences that are really quite variable both in context, and meaning, and in the manifestation of hotness (or notness). Consider also the response some have when confronted with a violation of this relationship—seeing a hot girl with a not-so-hot guy, for instance. This may well invoke a sense of discord, and one that eases with the discovery that he happens to be a true gentleman, or maybe just rich. In the absence of such information, we may even speculate about what accounts for this violation. Our response to such experiences suggests that, for this everyday contingency, we have models or beliefs about why the relationship exists and what modulates it, ever at the ready to assist in making sense of our experiences.

The juxtaposition of these two examples is designed to illustrate that the issue of learning a contingency between the hotness of romantic partners cannot be reduced to that of learning about elevator reliability, despite unquestionable structural similarities. Yet, research on HCL typically does exactly that. Most studies, by a clear margin, use tasks that amount to conditioning procedures in which participants receive numerous trials that convey little more than the value of two or more binary variables, always expressed in the same manner, under cover stories that range from familiar (e.g., foods that cause allergies or symptoms of diseases) to somewhat removed from everyday life (e.g., efficacy of weapons or chemical treatment of bacteria).

In contrast to this tradition, the current work introduces a contingency learning task that asks participants to do something quite natural—thinking about how other people behave and why—based on descriptive and individuated vignettes.[6] Certainly the materials employed in this task have not captured all that is important about natural experiences. Information about others often comes from direct observation, or second hand stories, rather than in the form of text descriptions. The vignettes used here convey only a fraction of the vast array of multi-sensory information that is available in interactions with others that play out in real time. Nor do we typically receive all our information about a person at once, organized or neatly bundled as relevant to the task or question at hand. Nonetheless, the current work is a much-needed step towards capturing two key attributes of much natural cognition in the study of HCL—(a) that the problems are often very familiar, in domains about which we have prior knowledge, causal

[6] The current task is not alone in its focus on learning contingencies in the behavior of other people. Although not exactly a contingency learning task, Cheng & Novick (1990) tested their Probabilistic Contrast Model by evaluating causal inferences people made about others' behavior. Additionally, Meiser & Hewstone (2004) examined stereotypical inferences made after learning a contingency between the desirability of individuals' behavior and their group membership.

models and personal theories, and (b) that learners typically have to deal with substantial variability and complexity in everyday experience.

Validity of the current task was established by demonstrating that it reproduced several key performance phenomena that have been repeatedly demonstrated in previous studies of HCL. In particular, contingency learning was strongly correlated with ΔP, and it was subject to both outcome density and blocking effects. The current work also presents some novel attributes of performance in HCL tasks that have been underappreciated in previous studies because they are difficult if not impossible to demonstrate using traditional tasks that employ distilled and homogeneous stimuli. First, participants appeared to evaluate the quality or informative value of evidence they were given, based on domain-specific beliefs and knowledge they brought into the task, and weighted its contribution to the contingency they perceived according to that evaluation. Second, participants' expectations about what would happen in future events were influenced not only by the contingency they perceived, but also separately by similar individual experiences or evidence that had been encountered during the task. Finally, it appears that, as a result of domain-specific knowledge, participants process evidence in a manner that leads them to learn the contingencies they expect to find and miss those they don't anticipate. The remainder of this concluding section discusses how the current work relates to two issues that are of current interest to those studying HCL, and offers some future directions.

The associative versus higher-order cognitive debate

In 2007, the Quarterly Journal of Experimental Psychology published a special issue on human contingency learning. In some form or another, 10 of the 12 articles published in this issue weighed in on the debate about whether HCL is an associative learning process and conditioning analog, or whether it involves the operation of higher order cognitive processes like inferential reasoning. To say that this issue is a major, current preoccupation of the field almost seems an understatement.

Some of the data presented here are certainly consistent with an associationist account of HCL. In the absence of manipulations designed to affect learning, participants' ratings of the contingencies they perceived were quite strongly correlated with the objective relationship as indexed by the ΔP statistic, and this is precisely what would be predicted by the Rescorla-Wagner (1972) model of associative learning (Chapman & Robbins, 1990; Dickinson et al., 1984). As Dickinson et al. further demonstrated, the R-W model predicts both blocking and outcome density effects like those observed here. To directly assess how closely performance in the current work matches that predicted under an associationist account, simulations were run with the R-W model for the various study conditions reported in Chapters 2 and 3 (contingency ratings were not collected in the studies reported in the Chapter 4). The results of these simulations in comparison to observed performance are presented in Table 1.

Table 1
Comparison of Observed Contingency Ratings to Simulations of the Rescorla-Wagner (1972) Model for Experiments reported in Chapters 2 and 3

Thesis Chapter & Experiment #	Learning conditions	Predicted rating	Observed rating
Ch. 2, Exp 1 *sensitivity to ΔP and outcome density (OD) effects*	$\Delta P = .3$, OD = .35	27	28
	$\Delta P = .3$, OD = .75	50	48
	$\Delta P = .7$, OD = .35	48	57
	$\Delta P = .7$, OD = .75	61	67
Ch. 2, Exp 2 *blocking effects*	$\Delta P_X = .5$	45	36
	$\Delta P_Y = 0$	9	-4
	$\Delta P_X = .5$	18	21
	$\Delta P_X = 1.0$	68	90
Ch. 3, Exp 1 *adjusting for evidential quality*	$\Delta P = .4$, normal exceptions	40	33
Ch. 3, Exp 2 *instantiated knowledge effects*	$\Delta P = 1.0$	68	84
	$\Delta P = .6$	48	55
	$\Delta P = .2$	30	17

Note. Simulations were based on the Rescorla-Wagner (1972) model: $\Delta V = \alpha\beta[\lambda - (\Sigma V)]$. α, the parameter representing salience of the cues or predictor variables, was set to 0.8 for the cue (or both cues in the blocking simulation), and 0.4 for the context. While this difference in the cue and context salience is less than might typically be employed in simulations of the model, it was selected because the learning items here provide rather meaningful contextual information that varies from item to item, arguably making it more salient than the contextual information in traditional tasks. β is the learning rate parameter that reflects salience of the outcome, or in more traditional preparations, the strength of the reinforcer. None of the behavioral outcomes used in the current task were unusual or abnormal (e.g., being abusive or violent), so β was set to 0.2. Finally, λ or the maximum associative strength supported by the outcome was set to 1.0

To summarize, observed performance shows good overall agreement with that predicted by the R-W model, with the exception that participants' ratings in conditions where the objective contingency is either very strong or very weak tend to be more extreme than the predictions of the model. This consistency is particularly noteworthy, considering that participants in the current studies were not told at the outset that their goal was to assess contingencies nor what the relevant cue and outcome variables were, and the stimulus materials provided extraneous and potentially distracting information.

At the same time, there are certain patterns of performance in the studies reported here that are inconsistent with associative accounts, and that would support arguments that HCL involves the operation of higher order cognitive processes (e.g., Beckers et al., 2005; Waldmann, 1996; Waldmann & Holyoak, 1992). Consider first the demonstration that participants' contingency ratings were influenced by the quality or value of the evidence they were provided. In that study, every participant received 6 of 20 learning items that were factually inconsistent with a particular relationship such as "Graham tends to be rude to his relatives". One group received items that were plainly inconsistent, like:

> Graham and the rest of his family were sitting down to dinner on Friday; his mom had brought home pizza. Graham wanted seconds, but he offered the last two slices to everyone else before taking them for himself.

The other group received items that were no less inconsistent, but that nonetheless provided little evidence for a link between rudeness and interacting with relatives because Graham's behavior was constrained by other compelling forces, such as:

> Graham's second cousin Doug was not only very rich, but he also had a lot of connections especially down at city hall where Graham was hoping to get a job for the summer. When Doug dropped by to visit Graham's dad, Graham jumped off the couch and offered to take Doug's coat and get him a drink.

The data revealed that inconsistent evidence like the latter had less of an influence on contingency judgments than the former type.

At first glance, this finding appears inconsistent with associative accounts of learning, because in both items Graham is interacting with his relatives but not being rude and therefore they ought to equally influence contingency judgments. However, one could argue that this pattern of data is consistent with associative accounts like the R-W model in which the primary driver of learning, with each piece of evidence, is whether or not the outcome is surprising. Assuming that someone believes Graham has a moderate tendency to be rude to his relatives, Graham's politeness during dinner with his family would come as much more of a surprise than would his behavior during cousin Doug's visit. In the R-W model, though, an outcome can only be unsurprising and therefore cause little change in the perceived cue-outcome relationship when its occurrence is already quite predictable from the presence of other cues. For the R-W model to predict the current result, then, it would be necessary to assume that, prior to the task, participants had already developed strong associations between politeness and interacting with rich, well-connected people, but also between politeness and 5 additional, unique variables that served as the compelling constraint on whether Graham was rude or polite in the remaining inconsistent items; and a further 18 strong associations to accommodate the 6 different constraints used in each of the 3 other behavioral contingency tasks. Finally, some of the compelling determinants built into the materials used here are ones that participants almost surely have come to associate with the target behaviors through experience—that one shouldn't expect politeness from a drunk, or that even the very obnoxious early bird is apt to do little when nursing a hangover the morning after a party, for instance. Others though—like whether a morning person would get involved in an all-night marathon to raise money for a friend's treatment of a rare and serious disease, or whether she would refuse her Thursday morning garbage duty in her apartment co-op because of a maggot and rat problem caused by the disregard of others—seem less likely to simply be well-established prior associates and more the product of reasoning or drawing inferences about how people would likely behave and why.

Further work needs to be done, though, to clearly establish that weighting of evidence according to its quality or value reflects an inferential reasoning process as opposed to just an opposing or competitive influence of known associates (or generalization from similar cues). While it capitalizes on a somewhat different notion of quality, one possibility for demonstrating the involvement of reasoning in weighting evidence would be to provide participants with information suggesting that some of the evidence they received (or will receive) is from a discredited source, or alternatively is outdated, and examine whether that subset of evidence exerts less influence on contingency judgments compared to participants who are told nothing additional about the quality of the evidence.

A second finding presented in the current work that is inconsistent with associative accounts of HCL is that outcome predictions were a product of not just the perceived contingency but also individual learning items that were similar to the prediction test scenarios. For example, in deciding whether Graham would let his sister have the only available seat on the bus they boarded, participants were influenced by both his tendency to be rude to his relatives and one particular learning item they recalled in which Graham, on his way to school, did get up and offer his seat to an elderly gentleman. Under associative accounts of HCL, responding should be very simply a function of the current associative strength regardless of the specific nature of the test or measurement procedure (e.g. Vadillo & Matute, 2007). It may be tempting to argue that participants' outcome predictions did not violate this principle, that they were indeed just a product of associations they had acquired during learning—specifically, one association between relatives and being rude, and a separate association between being on a crowded bus and being not rude. To be consistent with associative accounts in which cues compete for limited associative strength, though, this argument would require that the latter crowded bus-politeness association would be weaker, and therefore have less of an effect on outcome predictions, as the association between relatives and politeness increased in strength. However, there was no evidence that the prior similar instance effect was smaller when the objective relationship between the target cue and outcome was stronger. In addition, post-session interviews with some participants suggested that their predictions did not result from some automatic combination of two differently valenced associations, but rather that they were aware of both the contingency and the single instance as potential sources on which to base their prediction and that they explicitly deliberated about combining them to arrive at their judgment.

Perhaps the most problematic result for associative theories of HCL is that presented in Chapter 4. Two groups of participants were asked to learn the same contingencies (e.g. being active when the time of day is early) from not only items that had the same statistical properties, but evidence that was *exactly the same*. Furthermore, it is worth emphasizing that the contingencies presented to both groups were perfect (i.e., $\Delta P = 1.0$). The only difference between the two groups was that they were asked to learn the contingencies from vignettes that described a single person, or a random collection of different people. The results quite clearly showed that participants perceived the cue-outcome contingencies to be quite strong when they were presented through a single person, but very weak when they were presented through a random group. One possible explanation of this result is that participants' domain-specific knowledge led them to look for and find the contingencies they expected and miss the ones they didn't anticipate. In other words, the specific content of the contingencies used in the current tasks, like being rude to one's relatives or being more active early in the day, are surely patterns that participants have encountered in many individuals. While they are quite conceivably patterns that could also be true of the behavior of groups—rowing teams and medical doctors are often early birds—there is likely little prior knowledge or reason for participants to believe that these patterns would be expressed by a random collection of people. Moreover, participants reported after the task that they were engaging in less comparative processing between the learning items in the random group condition, which is consistent with reports that information about individuals is processed more integratively than information about groups (see Hamilton & Sherman, 1996).

While there is also much future work needed here to investigate the mechanism(s) responsible for this effect, even without a clear answer to this question it is difficult to see how an associative account would predict that learning the same perfect contingency from a set of evidence that provided identical information about the cue, the outcome and their co-occurrence could be so dramatically attenuated. The associative learning phenomena that perhaps come readily to mind as most similar to that documented here are those involving selectivity in the formation of associations, such as Garcia and Keolling's (1966) demonstration that taste cues are much more readily associated with illness than shock, and the reverse is true for audiovisual cues. However, such associative learning phenomena are reducing the associability of a cue by pairing it with a different or inappropriate outcome. In the current work, participants in the single person and random group were learning about the same cue paired with the same outcome. Perhaps the only way that an associative account could predict this result would be to argue that in the random group condition, the salience of the cues was very low, which would attenuate learning. It should be fairly straightforward to evaluate this argument by simply instructing participants in both conditions to evaluate the information for the particular cue-outcome contingency contained in each set of materials, in order to make the salience of the cue more equivalent between the two conditions. If the perceived contingency in the group condition is weaker despite this instruction, this would suggest the involvement of some non-associative process, such as being swayed against believing strongly in a contingency for which one has a difficult time generating a good causal story. Explanations based on cue salience also do not deal with participants' self-reported differences in how they were processing the learning items and the consequences of comparative versus individualistic processing for contingency learning.

The results of the studies described here appear to lend themselves best to some sort of hybrid account of HCL that involves both associative and higher order cognitive processes, which is quite consistent with the current direction of the field (e.g., Vadillo & Matute, 2007; Tangen & Allan, 2004; Cheng, 1997; Price & Yates, 1995). Under many conditions, participants did appear to extract the statistical properties of their experience and acquire knowledge of contingent relationships in a manner consistent with the operation of an associative learning mechanism. However, the patterns of performance observed here suggest that such mechanisms do not operate outside the influence of participants' beliefs and models about the world and the inferences they draw from them, and that responding is not an automatic or inevitable product of such mechanisms.

Serial position or order effects in contingency learning

Many researchers have been interested in investigating how the order of evidence presentation affects contingency judgments, and the consistency between such serial position effects and the various theoretical accounts of HCL. Order effects are typically studied by blocking the presentation of learning items such that the contingency across early items is stronger, weaker or opposite to that across later items and evaluating how contingency or causal judgments differ between these conditions or change across the session (e.g., Vadillo & Matute, 2007; Dennis & Ahn, 2001; López, Shanks, Almaraz, & Fernández, 1998; Wasserman, Kao, Van Hamme, Katagiri & Young, 1996; Yates & Curley, 1986). If contingency learning proceeds by extracting information from experiences according to a statistical rule such as ΔP (e.g., Cheng & Holyoak, 1995; Cheng & Novick, 1992), the result should be unaffected by any blocking or ordering of evidence. In contrast, as Vadillo & Matute describe,

> Associative learning algorithms, on the contrary, are supposed to be highly sensitive to the precise order in which information was provided. Specifically, associative models generally assume that cue-outcome associations are constantly being updated as more information is provided. This means that the associative strength is strongly determined by the most recent contingencies. These models predict that, when contradictory information is received in different phases, what is learned in the last phase will overwrite what was learned previously…

In other words, given the typical study procedure, associative accounts predict a recency effect because the recent information is "surprising" against what was learned earlier, and it is unanticipated outcomes that really drive learning or alter perceived contingencies.

The picture that emerges from studies regarding how information order affects contingency judgments, though, is somewhat murky. Some evidence suggests that recent information has a stronger effect on judgments (Vadillo & Matute, 2007; Collins & Shanks, 2002; López et al., 1998), which would be consistent with associative accounts. However, others have demonstrated that early information is more influential (Dennis & Ahn, 2001; Yates & Curley, 1986), and even no evidence of a presentation order effect (Wasserman et al., 1996). While none of the data from the current studies addresses order effects in HCL, or why they might occur, the task introduced here may be of some value in investigating why early or late information seems to be more influential.

Associative accounts predict recency effects in contingency judgments, *in the manner in which they are typically studied.* In other words, when the recent information conveys a different contingency than the early information, which makes the outcome surprising given the cues presented, this later information causes important shifts in perceptions. However, according to the R-W model, if later evidence was informationally equivalent to that experienced early, the outcome would be anticipatable and perceptions would change little in the face of this later evidence. The studies described here introduce a different way of manipulating the surprisingness of evidence, and manipulating when that information is presented could be useful in establishing whether recent information is only more influential when it is surprising. Specifically, in the second manuscript, performance was compared for two groups of participants who received 6 items that provided evidence against a contingency, but these items were only surprising for one of the two groups. If recency effects are dependent on the surprisingness of the evidence, then presenting these 6 exceptions early or late should only affect contingency judgments in the group who receives the surprising exceptions.

Dennis and Ahn (2001) argued that the strong influence of early information, or primacy effect, they observed occurred because it provided the learner with a hypothesis that they could subsequently evaluate as they encountered new evidence. Their discussion is somewhat unclear, though, as to whether they are arguing that early information provides a working hypothesis, or whether it simply exerts an anchoring effect on contingency judgments (although they seem more inclined toward the former). Again, the current task may prove useful in distinguishing between these two possibilities. If primacy effects occur because early information provides a working hypothesis, it is quite reasonable to expect that this effect would be modulated by the quality of the early information. In other words, people ought to be less willing to form early hypotheses based on highly suspect as opposed to good quality information. In contrast, anchoring effects in judgment need not be rational or influenced by the quality of the anchor. For instance, Tversky and Kahneman (1974) asked participants to estimate the percent of African nations belonging to the UN. Just prior to providing their estimates, participants were shown the results of a random number generator. Although the number generated was random and unrelated to the topic of judgment, it nonetheless influenced participants' estimates of the percentage of UN members. The current work has discussed various ways in which evidential quality in the current materials might be readily manipulated, and it would be interesting to investigate whether primacy effects in contingency judgments are influenced by the quality of that early evidence.

Future directions

The stimuli in the current task differed from those typically used in two different ways. First, expression of the cue and outcome was variable and unique from trial to trial, and second, learning items contained a good deal of additional information that might have been interpreted as additional cues, a context or setting for the event, or simply irrelevant to the outcome. It would be interesting to separately evaluate how variability in cue or outcome manifestation and the presence of extraneous detail affect the accuracy and rate of learning. Further, although somewhat difficult to envision at the moment, it would be interesting to consider how associative models might accommodate qualitative variability in cues or outcomes, and the effect of transient cues or contextual information that changes from trial to trial.

I am also interested in using the current task, or modifications thereof, to better understand what is involved in what I will for now refer to as cue amplification. Although not the happiest of examples, some of the experiences we collect through our day-to-day activities are those in which we learn about some rather egregious or horrific behaviors, like sexual abuse or incest, pedophilia, or school shootings. Often, these experiences are also potentially revealing about causes of the behavior, or at very least, some predictors of it. Indeed, concern is quite high about how to anticipate any one of these negative outcomes. As with other examples of human behavior, it is almost surely the case that there are multiple, and interactive causes or predictors at work. Yet, anecdotal experience suggests that people not only drastically overestimate the value of one or a couple of predictors, but that they are often rather confident in those beliefs even in the face of little evidence. To provide one example, consider the argument that seems to resurface with each school shooting—that these terrible acts would not occur were there no bullying or violent video games to play. To inflate the predictive value of a cue, and furthermore to be overconfident in its value based on only a small sample of evidence would be quite inconsistent with associative accounts of HCL, but may well occur with outcomes that have a very strong emotional valence and when the conditions of learning support people's ability to generate causal stories about the relationship between the predictor(s) and the outcome. Relatedly, this sort of cue amplification may also be most likely to occur under conditions of conscious thought or deliberation, as opposed to incidental or unconscious processing (e.g., Dijksterhuis & Nordgren, 2006).

References (for Introduction and Concluding Discussion)

Allan, L. G., Siegel, S., & Tangen, J. M. (2005). A signal detection analysis of contingency data. *Learning & Behavior : A Psychonomic Society Publication, 33*(2), 250-263.

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

Allan, L. G. (1993). Human contingency judgments: Rule based or associative? *Psychological Bulletin, 114*(3), 435-448.

Allan, L. G. (2003). Assessing power PC. *Learning & Behavior, 31*(2), 192-204.

Allan, L. G., & Jenkins, H. M. (1983). The effect of representations of binary variables on judgment of influence. *Learning and Motivation, 14*(4), 381-405.

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General, 108*(4), 441-485.

Baker, A. G., Mercier, P., Vallée-Tourangeau, F., Frank, R., & Pan, M. (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*(2), 414-432.

Beckers, T., De Houwer, J., Pineño and, O., & Miller, R. R. (2005). Outcome additivity and outcome maximality influence cue competition in human causal learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(2), 238-249.

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

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*(2), 367-405.

Cheng, P. W. (., & Holyoak, K. J. (1995). *Complex adaptive systems as intuitive statisticians: Causality, contingency, and prediction.* Cambridge, MA, US: The MIT Press.

Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology, 58*(4), 545-567.

Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review, 99*(2), 365-382.

Collins, D. J., & Shanks, D. R. (2002). Momentary and integrative response strategies in causal judgment. *Memory & Cognition, 30*(7), 1138-1147.

De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and theories on human contingency learning. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 55B*(4), 289-310.

Dennis, M. J., & Ahn, W. (2001). Primacy in causal strength judgments: The effect of initial evidence for generative versus inhibitory relationships. *Memory & Cognition, 29*(1), 152-164.

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

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science, 1*(2), 95-109.

Durlach, P. J. (1983). Effect of signaling intertrial unconditioned stimuli in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes, 9*(4), 374-389.

Durlach, P. J. (1989). Role of signals for unconditioned stimulus absence in the sensitivity of autoshaping to contingency. *Journal of Experimental Psychology: Animal Behavior Processes, 15*(3), 202-211.

Feingold, A. (1988). Matching for attractiveness in romantic partners and same-sex friends: A meta-analysis and theoretical critique. *Psychological Bulletin, 104*(2), 226-235.

Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science, 4*(3), 123-124.

Gredebäck, G., Winman, A., & Juslin, P. (2000). *Rational assessments of covariation and causality.* Proceedings of the 22nd Annual Conference of the Cognitive Science Society (pp. 190–195). Hillsdale, NJ: Lawrence Erlbaum

Associates, Inc.

Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological Review, 103*(2), 336-355.

Houwer, J. D., Beckers, T., & Glautier, S. (2002). Outcome and cue properties modulate blocking. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 55*(3), 965-985.

Houwer, J. D., Beckers, T., & Glautier, S. (2002). Outcome and cue properties modulate blocking. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 55*(3), 965-985.

Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9-31). Miami, FL: University of Miami Press.

López, F. J., Almaraz, J., Fernández, P., & Shanks, D. (1999). Adquisición progresiva del conocimiento sobre relaciones predictivas: Curvas de aprendizaje en juicios de contingencia. / gradual acquisition of knowledge of predictive relationships: Learning curves in contingency judgments. *Psicothema, 11*(2), 337-349.

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

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

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*(1), 182-196.

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*(1), 182-196.

McClelland, J. L. (1987). *Parallel distributed processing. Vol. 2: Psychological and Biological Models.* Cambridge, Mass.: MIT Press.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114*(2), 159-188.

Meiser, T. (2003). Effects of processing strategy on episodic memory and contingency learning in group stereotype formation. *Social Cognition, 21*(2), 121-156.

Meiser, T., & Hewstone, M. (2004). Cognitive processes in stereotype formation: The role of correct contingency learning for biased group judgments. *Journal of Personality and Social Psychology, 87*(5), 599-614.

Meiser, T., & Hewstone, M. (2006). Illusory and spurious correlations: Distinct phenomena or joint outcomes of exemplar-based category learning? *European Journal of Social Psychology, 36*(3), 315-336.

Mitchell, C. J., & Lovibond, P. F. (2002). Backward and forward blocking in human electrodermal conditioning: Blocking requires an assumption of outcome additivity. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 55*(4), 311-329.

Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin, 68*(1), 29-46.

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

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*(6), 1639-1655.

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

Rumelhart, D. E. (1987). *Parallel distributed processing. Vol. 1: Foundations.* Cambridge, Mass.: MIT Press.

Shanks, D. R. (1985). Forward and backward blocking in human contingency judgement. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 37*(1), 1-21.

Shanks, D. R. (1986). Selective attribution and the judgment of causality. *Learning and Motivation, 17*(4), 311-334.

Shanks, D. R. (1987). Acquisition functions in contingency judgment. *Learning and Motivation, 18*(2), 147-166.

Shanks, D. R. (1989). Selectional processes in causality judgment. *Memory & Cognition, 17*(1), 27-34.

Shanks, D. R. (1991). Categorization by a connectionist network. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(3), 433-443.

Shanks, D. R. (1995). *The psychology of associative learning*. New York, NY, US: Cambridge University Press.

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

Spellman, B. A. (1996a). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science, 7*(6), 337-342.

Spellman, B. A. (1996b). *Conditionalizing causality*. San Diego, CA, US: Academic Press.

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

Tangen, J. M., & Allan, L. G. (2004). Cue interaction and judgments of causality: Contributions of causal and associative processes. *Memory & Cognition, 32*(1), 107-124.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Vadillo, M. A., & Matute, H. (2007). Predictions and causal estimations are not supported by the same associative structure. *The Quarterly Journal of Experimental Psychology, 60*(3), 433-447.

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

Van Overwalle, F., & Van Rooy, D. (2001). How one cause discounts or augments another: A connectionist account of causal competition. *Personality and Social Psychology Bulletin, 27*(12), 1613-1626.

Wagner, A. R., Logan, F. A., & Haberlandt, K. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology, 76*(2, Pt.1), 171-180.

Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(1), 53-76.

Waldmann, M. R. (2001). Predictive versus diagnostic causal learning: Evidence from an overshadowing paradigm. *Psychonomic Bulletin & Review, 8*(3), 600-608.

Waldmann, M. R. (. (1996). *Knowledge-based causal induction*. San Diego, CA, US: Academic Press.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121*(2), 222-236.

Ward, W. C., & Jenkins, H. M. (1965). The display of information and the judgment of contingency. *Canadian Journal of Psychology/Revue Canadienne De Psychologie, 19*(3), 231-241.

Wasserman, E. A. (1990). Attribution of causality to common and distinctive elements of compound stimuli. *Psychological Science, 1*(5), 298-302.

Wasserman, E. A. (., Kao, S. (., Van Hamme, L. J. (., Katagiri, M. (., & Young, M. E. (. (1996). *Causation and association*. San Diego, CA, US: 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*(4), 406-432.

Wasserman, E. A., & Shaklee, H. (1984). Judging response-outcome relations: The role of response-outcome contingency, outcome probability, and method of information presentation. *Memory & Cognition, 12*(3), 270-286.

Yates, J. F., & Curley, S. P. (1986). Contingency judgment: Primacy effects and attention decrement. *Acta Psychologica, 62*(3), 293-302.