Canadian Psychology / Psychologie canadienne

0708-5591/19/$12.00
© 2019 Canadian Psychological Association

2019, Vol. 60, No. 1, 3–13
http://dx.doi.org/10.1037/cap0000170

CPA Gold Medal Award for Distinguished Lifetime Contributions to Canadian Psychology

Learning Simple Associations

COLIN M. MACLEOD
University of Waterloo

Abstract
People are extraordinarily good at establishing connections in memory—associative learning. Indeed, this is quite possibly their most frequently deployed cognitive skill. Learning contingencies—that one event is more likely when in the presence of another—is a crucial form of associative learning that allows one to make successful predictions, increasing the speed and accuracy of responding to events in the world. This article describes a program of research investigating the fundamental building blocks of contingency learning underlying people's acquisition of correlational and causal information in the world around them.

Keywords: association, contingency, learning

The learning of associations is a fundamental—perhaps *the* fundamental—aspect of learning. Thorndike (1911) and Pavlov (1927) recognized this when they began the scientific study of associative learning over 100 years ago. To function successfully, organisms must be able to relate objects and events in the world to each other and in turn to relate those objects and events to behaviors. Whether a planarian is acquiring the connection between the presence of a light and a physical reaction such as contraction (McConnell, Jacobson, & Kimble, 1959) or a chimpanzee is acquiring a prosocial behavior resembling altruism (Horner, Carter, Suchak, & de Waal, 2011), associative learning is prevalent throughout all species with nervous systems. Even the single-celled amoeba, with no nervous system, learns to behaviorally anticipate a change in temperature (Saigusa, Tero, Nakagaki, & Kuramoto, 2008). In the last great unifying theory in psychology, Hebb (1949) built his cell assembly explanation of learning in the nervous system around the idea of association. Although distinguishing between "simple" associative learning and "higher order" cognitive learning is not straightforward (cf. Heyes, 2012), it is clear that association plays a central role in learning.

In humans, this ability to acquire and retain links may be their greatest cognitive strength. From the earliest moments of infancy, they display a remarkable ability to learn and remember relations between things. At just 1−3 days old, newborns are already learning statistical regularities between pairs of shapes, as shown in habituation paradigms (Bulf, Johnson, & Valenza, 2011). By 8 months, this statistical learning is evident in sophisticated segmentation of speech streams (Saffran, Aslin, & Newport, 1996). And the importance of associative learning only grows as infants set out on their most impressive voyage—their learning of language (e.g., Werker, Cohen, Lloyd, Casasola, & Stager, 1998). Throughout life, whether learning word definitions or face-name pairings or a myriad of other kinds of information, associative learning remains a core skill. One has only to consider the breakdown of this skill in patients with frontal lobe damage to realize just how crucial it is (e.g., Petrides, 1985).

In recent years, I have become intrigued by one form of associative learning—the learning of contingencies. How does one learn, often without intention and sometimes even without awareness, that elite golfer Tiger Woods typically wears a red shirt on the final Sunday of a tournament, or that the sudden appearance of high winds usually signals the coming of a substantial change in temperature? Surrounded by a great many correlations—and causations—people are really quite exquisitely tuned to absorbing these, as when a poker player learns—possibly without ever becoming conscious of having learned—that an opponent ordinarily makes a little less eye contact when holding a good hand.

Not surprisingly, researchers have argued that this kind of learning is special. Thus, Hasher and Zacks (1979, 1984; Zacks & Hasher, 2002) have posited that the encoding of information about event frequency occurs automatically as a consequence of attending to events. Conceiving of simple associations as primitive events, then, would suggest that learning about the frequency of those events should have a high priority. In a related vein, Kelly and Martin (1994) proposed that people's ability to be sensitive to probabilistic patterns around them— given that probabilistic patterns are everywhere—is critical in a host of problem-solving situations, from foraging to depth perception to language process-

---

This article was invited in association with the author being awarded the 2018 Gold Medal for Distinguished Lifetime Contributions to Canadian Psychology from the Canadian Psychological Association. The research reported here was supported by Natural Sciences and Engineering Research Council of Canada Discovery Grant A7459. I thank Noah Forrin, my primary collaborator on almost all phases of this work. In preparing this article, I also benefitted from comments by Derek Besner, Ramona Bobocel, Bill Hockley, and James Schmidt.

Correspondence concerning this article should be addressed to Colin M. MacLeod, Department of Psychology, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada. E-mail: cmacleod@uwaterloo.ca

ing. Like frequency, then, probability has a high priority for successful learning. Contingencies, in turn, can be thought of as probabilistic frequencies.

Contingency learning is, therefore, an essential skill, so it has been studied extensively in the past, primarily in animal-conditioning studies (see Hall, 2002; R. R. Miller & Escobar, 2002) but increasingly in studies of human learning (see De Houwer & Beckers, 2002; Shanks, 2007). Ordinarily, the contingencies to be learned have taken the form of sequences—for example, Event B often follows Event A such that A becomes a signal predictive of B. As a way to measure the learning of the contingencies, experiment participants are often asked to anticipate—to predict—what will come next, although they are sometimes simply timed on their response to a second event as a function of the first event (see Sternberg & McClelland, 2012). As contingency learning grows, predictions should be more accurate and response times should be faster when contingencies are present than when they are absent. And the stronger the contingency, the more robust this pattern should be.

Contingencies do not, however, require a sequential paradigm, as is certainly clear from the animal literature. In 2007, Schmidt, Crump, Cheesman, and Besner introduced a simple simultaneous task (for related paradigms, see Carlson & Flowers, 1996; Levin & Tzelgov, 2016; J. Miller, 1987; Mordkoff & Halterman, 2008). It was this that caught my attention, perhaps primed by my past research exploring the Stroop task (see MacLeod, 1991). In their color-word contingency paradigm, participants respond on each trial to the color of a word by pressing a corresponding button. In the prototypical experiment, there are three colors (e.g., red, yellow, and green) and three words (e.g., *month, under,* and *plate*). Each trial consists of a word in color, with the color being response-relevant and the word being response-irrelevant. The contingency is between the words and the colors: Again prototypically, each word is presented 80% of the time in one of the colors and only 10% of the time in each of the other two colors. Table 1 illustrates how this would look for a typical block of 30 trials. The *contingency learning effect* is calculated as the difference in response time (or accuracy) between low-contingency (LO) and high-contingency (HI) trials.

Table 1
Illustration of a Block of 30 Trials in the Color−Word Contingency Learning Paradigm

Word and color	No. of trials	Contingency
month		
RED	8	HI
YELLOW	1	LO
GREEN	1	LO
under		
RED	1	LO
YELLOW	8	HI
GREEN	1	LO
plate		
RED	1	LO
YELLOW	1	LO
GREEN	8	HI

Note. HI = high contingency; LO = low contingency.

The Schmidt and De Houwer Studies

In a series of articles, Schmidt, De Houwer, and their colleagues have investigated a considerable number of interesting aspects of this learning situation. Initially, Schmidt et al. (2007) saw the learning in this situation as evidence of learning without awareness, given that a substantial subset of their participants could not report the contingencies after the experiment yet these participants did not behave differently from those who could report the contingencies, indicative of a kind of implicit control. Schmidt, De Houwer, and Besner (2010) went on to show that the onset of the learning is extremely rapid (they suggested that this occurs in fewer than 18 trials) and is quite stable thereafter. Moreover, when the contingencies were "turned off" (i.e., when every word appeared equally often in every color), unlearning was also very rapid as well. Schmidt et al. (2010, Experiment 2) further demonstrated that this type of contingency learning is resource-dependent in that it was sharply attenuated when participants had to carry a memory load while doing the task: The load presumably commandeered a considerable proportion of the resources normally used for the learning.

The issue of awareness came back under investigation by Schmidt and De Houwer (2012d). They showed that instructing one group of participants regarding the contingencies—by telling them in advance which words appeared most often in which colors—enhanced their learning relative to another group of uninformed participants, as indexed by a larger difference between HIs and LOs. Schmidt and De Houwer (2012a) drove this point home when they explicitly told participants that there were contingencies and instructed them to intentionally try to learn the contingencies. Again, the contingency learning effect was larger for the instructed, intentional learners. Although Schmidt and De Houwer (2012a, 2012d) concluded that conscious knowledge could be moderating implicit learning, another possibility is that participants can engage in both implicit and explicit learning, which would also fit with the finding in Schmidt et al. (2007) that reducing the strength of the contingency (from 80% to 50% HI) reduced participant awareness but also reduced the magnitude of the contingency learning effect.

Further studies from their laboratory extended the phenomenon to the case of sequential presentations (Schmidt & De Houwer, 2012c) and to different classes of stimuli (e.g., evaluative stimuli; Schmidt & De Houwer, 2012b). In 2013, Schmidt introduced a computational model of contingency learning that he called the parallel episodic processing (PEP) model (Schmidt, 2013). This model (see also Schmidt, De Houwer, & Rothermund, 2016, and Schmidt, 2018, for further development of the model, and Schmidt, 2016, for another application of the model) has its roots in Logan's (1988, 2002) instance theory of automaticity, Hintzman's (1984) MINERVA 2 theory of recognition memory, and Medin and Schaffer's (1978) context theory of classification. Very simply, as shown in Figure 1, each trial lays down an instance of itself in memory. Each trial also routinely summons recent instances from memory to assist current processing, likely by biasing response selection. Because HI trials have many more instances—and hence many more recent instances—in memory, they benefit more from this routine instance retrieval, resulting in faster responding for HI trials than for LO trials. As is so often the case in

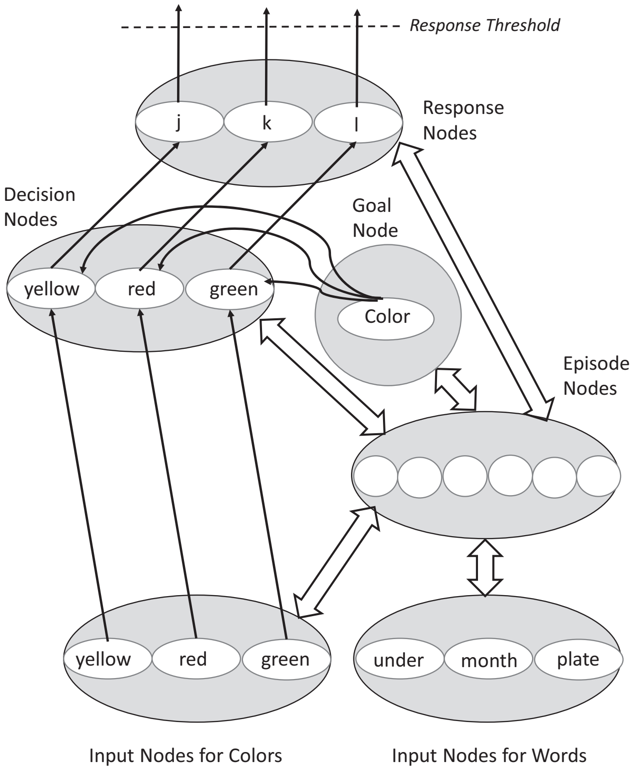

Figure 1. The parallel episodic processing (PEP) model (Schmidt, 2013, 2018; Schmidt et al., 2016). The goal node keeps the task—color naming—"in mind." Input nodes for color feed into the decision nodes and then into the response nodes. Input nodes for both words and colors feed into the episode nodes, creating instances of each experienced trial. These episode nodes then feed forward to the decision and response nodes (note that response nodes *j, k,* and *l* correspond to the color response keys for yellow, red, and green). From 'Best not to bet on the horserace: A comment on Forrin and MacLeod (2017) and a relevant stimulus-response compatibility view of colour-word contingency learning asymmetries' by Schmidt, J. R., 2018, *Memory & Cognition, 46*, p. 333. Copyright [2018] by Springer Nature. Adapted with permission.

cognition, memory is used to speed processing by circumventing having to think about what to do.

Most recently, as we were beginning our work on contingency learning, Schmidt and De Houwer homed in on the issue of the rapid, stable learning of contingencies (Schmidt & De Houwer, 2016b) and suggested that the apparent stability of the effect might be due to "learning to learn" (see Harlow, 1949, p. 51) the task of responding to the colors—that initial learning was actually learning to map the colors to the key responses and only after that was there true contingency learning involving word-color connections. Accordingly, when they first had participants practice responding to the colors, then introduced the words and contingencies, they now observed an increase in the size of the contingency learning effect across blocks, although it was still robust even during the first block of trials. This increase was particularly evident when the onset of the word preceded (by 150 ms) rather than co-occurred with, the color, a finding in keeping with the conditioning literature: A conditioned stimulus is most effective in delayed (forward) conditioning, where it precedes and co-occurs with the unconditioned stimulus, serving optimally as a signal (see, e.g., Chance, 2008).

The final experiment that I consider from the Schmidt and De Houwer laboratory at Ghent University leads directly into the first project in our laboratory at the University of Waterloo. They (Schmidt & De Houwer, 2016a) varied the contingency of individual words to colors and included as baselines words that they described as "medium contingency" (p. 79; i.e., words that appeared equally often in every color; Experiment 1) or "neutral" (p. 84; novel words that appeared in color only during a test phase; Experiment 2). They consistently showed that contingency proportion was the key predictor of performance. Our initial work, done in parallel and independently of their work, also began by addressing the baseline issue, our goal being to dissect the source of the contingency learning effect.

Cost, Benefit, and Baseline

Olivia Lin and I were first interested in whether the contingency learning effect—the HI–LO difference—resulted from speeding of the HI trials (a benefit), slowing of the LO trials (a cost), or possibly both. In Lin and MacLeod (2018), based on the first chapter of her dissertation (Lin, 2015), we adopted an approach similar to that of Schmidt and De Houwer (2016a, Experiment 1). We reasoned that a suitable baseline for indexing cost–benefit would be the case where a word appears equally often in each color, such that word and color are not contingent. Just imagine adding the word *clock,* presented four times in each color, to a version of Table 1 in which each of the other three words was now presented 10 times in one color and one time in each of the other two colors. We did this experiment twice, once using words and once using nonwords, and obtained identical patterns; consequently, I have combined these in Figure 2, with a total of 62 participants' having taken part.

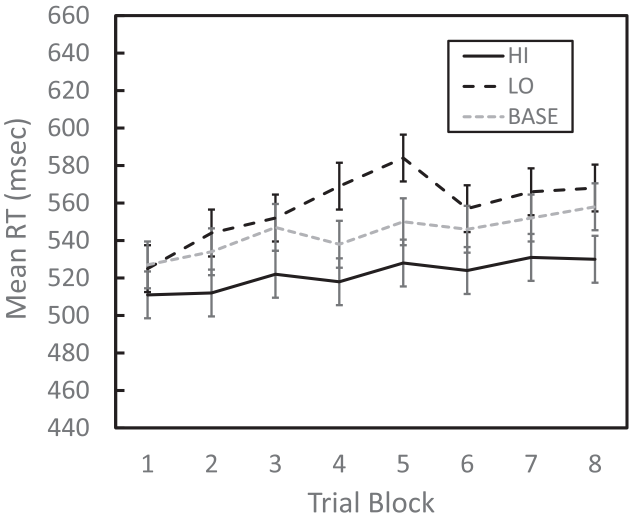

Figure 2. Mean response time for the high-contingency (HI), low-contingency (LO), and no-contingency baseline (BASE) stimuli as a function of trial block. Error bars are standard errors of the corresponding means. RT = response time. The data, combining the data for words (Experiment 1a) and nonwords (Experiment 1b), are from Lin and MacLeod (2018).

The answer to the cost–benefit question was clear: There was a benefit for HIs and a cost for LOs, based on deviations from the no-contingency baseline. Moreover, as in the Schmidt and De Houwer (eg., 2016a) work, the effects were evident almost from the outset and were quite stable over succeeding blocks of trials. In a second experiment, we observed this same pattern over six blocks, but when we turned off the contingencies for six blocks (all words appeared equally often in every color), contingency learning was immediately extinguished, with no evidence of it, even in the first no-contingency block. When in a final six blocks we restored the contingencies, the contingency learning pattern reappeared right away in the first block. It is difficult to determine whether there was any "savings" for the original learning in the first six blocks, given how rapid the learning always is from the onset of contingencies. In that regard, one of our continuing goals is to find ways to slow down the learning—possibly by increasing the set size of words and colors, by decreasing the proportion contingency, and so forth—so that we can track the early stages of learning and also examine the longevity of the learning using relearning–savings measures.

Consider these results in terms of the PEP model (Schmidt, 2013; Schmidt et al., 2016). The focus of the model is on the speeding of responses to HI trials due to the routine retrieval from memory of matching instances. We certainly reproduced this result relative to the no-contingency baseline. But we also observed apparent slowing of the LO trials relative to the same baseline, something the model did not emphasize. LO trials may actually incur a cost because few of the retrieved instances will match them; indeed, most of the retrieved instances will be HIs and hence will conflict with the LO trials, causing the LO trials to be surprising.

Is There a Contribution of Repetition?

One of our first thoughts concerning contingency learning related to the considerable overlap between contingency on the one hand and frequency and repetition on the other hand. Put simply, and as illustrated in Table 1, the individual high-contingency stimuli are each much more likely to occur than are any of the low-contingency stimuli. And this greater frequency for HIs also means that immediate repetitions of HIs are considerably more likely than are immediate repetitions of LOs. Could the contingency learning effect actually be an effect of differential immediate repetition favoring the HIs? Schmidt et al. (2010) reanalyzed the data of Schmidt et al. (2007) to address this question post hoc and determined that the contingency learning effect was still robustly present even when these repetitions were excluded from the data set.[1]

The first contingency learning study that Noah Forrin and I undertook addressed the repetition question experimentally with a direct manipulation (Forrin & MacLeod, 2015). Two groups of 40 participants each were compared, one group where immediate repetitions of individual stimuli were allowed (the standard procedure) and one where they were disallowed by the controlling program. We sought to determine how critical repetition was for the advantage of HIs over LOs. What we found was comforting with respect to contingency learning and in line with the Schmidt et al. (2010) post hoc repetition analyses: Although repetitions apparently had a small influence on the learning, that influence was not statistically significant. As Figure 3 portrays, response time was, not surprisingly, a little slower overall when repetitions were prevented (obviously, repetitions tend to be responded to more quickly, thereby lowering mean response time when they are present); however, the contingency learning effect was similar whether repetitions were allowed (33 ms) versus disallowed (27 ms). The contingency learning effect is therefore not a proxy for the influence of stimulus repetition being more likely for HIs.

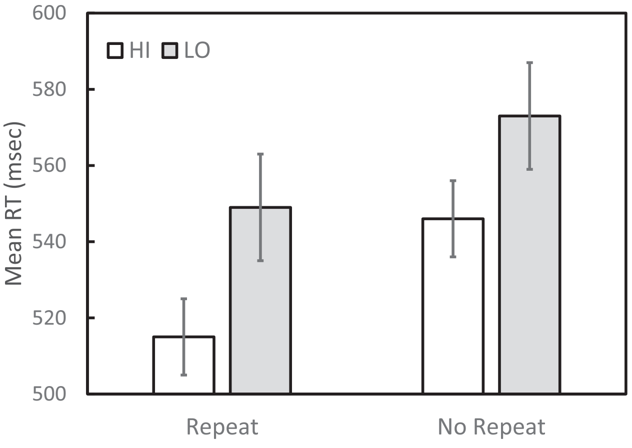

Figure 3. Mean response time for the high-contingency (HI) and low-contingency (LO) conditions as a function of whether individual trial repetitions were allowed or disallowed. Error bars are standard errors of the means. The data are from Forrin and MacLeod (2015, unpublished).

Single Versus Multiple Associations

In a way, our newest line of research in this domain is the flip side of the repetition question. In virtually all of the studies to be described in this article, both our work and that from the Schmidt and De Houwer laboratory, one word is connected to each color to create the high-contingency condition.[2] In an ongoing series of studies, Brady Roberts, Noah Forrin, and I are investigating what happens when more than one word has a high-contingency connection to the same color (Roberts, Forrin, & MacLeod, 2018). Imagine Table 1 doubled so that a second set of three words also connects to the same three colors (e.g., both the words *month* and *table* appear 80% of the time in red, both the words *under* and *horse* appear 80% of the time in yellow, and both the words *plate* and *phone* appear 80% of the time in green). At issue is whether the size of the contingency learning effect will be unaffected or whether this "fan effect" (cf. J. R. Anderson, 1974) will diminish contingency learning. It is certainly possible that the one-to-one mapping of high contingency that we have been using is a special

---

[1] James Schmidt informed me that, based on a recent more sophisticated analysis examining repetitions as a function of lag (e.g., 1 back, 2 back, etc.), it is conceivable that the contingency effect is effectively a collection of individual-trial binding effects (J. R. Schmidt, personal communication, August 6, 2018; i.e., that "binding effects" and "contingency effects" are equivalent but studied on different time scales, as suggested by Schmidt, De Houwer, & Rothermund, 2016, pp. 84–85).

[2] See Schmidt et al. (2007) for a variation involving two color responses connected to the same response key. More relevant, Schmidt and De Houwer (2012c) reported a variation involving three words per color, but they did not examine the influence that more words per color might have had on the magnitude of the contingency learning effect.

case. More theoretically, under the PEP model, there would be fewer instances of any one HI stimulus when two or more words (rather than only one word) are mapped to the same color.

This line of work is just under way, and the results of our first experiment are tantalizing. Figure 4 shows the data for 50 participants in each of two groups. There was strong contingency learning in both groups, whether one word or two words were connected with high contingency to each color. It was also the case, however, that the extent of that learning differed between the two arrangements: The magnitude of the learning was significantly greater when just one word was connected as a HI to a given color (43 ms) than when two words shared the same HI connection (27 ms). Of course, it may be that this difference is (at least partly) a consequence of reduced repetitions in the two-word case, which would be expected to influence the HI trials more than the LO trials. But given the findings concerning repetition in the preceding section, and in line with this being a true effect of contingency per se, it is interesting that the difference between the two conditions appears to be in the LO trials, not the HI trials. Perhaps the increased prevalence of LO trials in the six-word situation makes those trials less surprising than the more unique LO trials in the three-word condition.

Frequency and Contingency

The multiple-contingency study just described also points to the closely related issue of frequency versus contingency, a topic that deserves a concerted empirical attack. Frequency of individual stimuli is typically highly correlated with contingency—individual HIs are considerably more frequent than are individual LOs—so one must understand how frequency contributes to the observed contingency learning effect, just as we have done for repetition. Olivia Lin (2015) began to investigate frequency in her dissertation, pointing to the perception and attention literature as inspiration. As one illustration, Hon et al. (2013) constructed a task where accuracy was kept very high: Response times to high-probability targets became invariant across training, whereas response times to low-probability targets got slower. Hon et al. theorized that this could be due to the rarity of the low-probability targets, causing them to be surprising when they did occur.

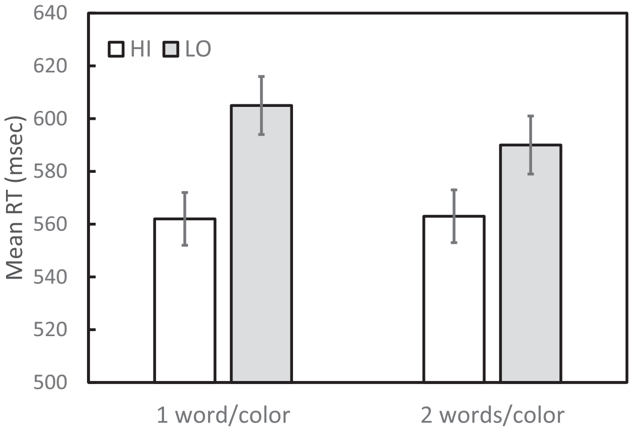

Figure 4. Mean response times when trials involve three words, with one word highly contingent on each of the three colors, and when there are six words, with two words highly contingent on each of the three colors. Error bars are standard errors of the means. HI = high contingency; LO = low contingency. The data are from Roberts, Forrin, and MacLeod (2018, unpublished).

I very briefly describe just two of Lin's (2015) experiments to give the "flavor" of how one might go about teasing apart frequency and contingency, despite their intricate connection. These two experiments offer very different perspectives—and findings—demonstrating that solving this particular problem will not be straightforward. In Experiment 7 of the dissertation, she constructed blocks of trials that maintained the usual 80%/20% balance of HI and LO trials—but with a twist. In these blocks, one set of items had 8–1–1 frequencies, following the usual procedure, but another set had 16–2–2 frequencies. Would performance differ for these items, which were presented twice as often in each block? The answer is that the contingency effect was unaltered by the frequency manipulation. This may have been because frequency and contingency did not compete with each other, in that both item sets had the same relative frequencies. This may also indicate that one can learn some event frequencies—here, the HIs and LOs—preferentially without necessarily learning other event frequencies—here, the blockwide frequencies (for an animal learning analog, see Rescorla, 1968).

Now consider another of Lin's (2015) experiments, Experiment 6 of the dissertation. Here, frequency and contingency do, in a sense, compete, and the outcome is very different. In a HI-contingency−LO-frequency condition, a given word was associated with only one color (high contingency), but that word itself occurred with relatively low frequency (e.g., *MONTH* appeared six times in a block always in red and never in yellow or green). In a LO-contingency−HI-frequency condition, each word appeared equally often in two colors (low contingency), but now the word occurred with relatively high frequency (e.g., *UNDER* and *PLATE* both appeared 12 times in a block, six times in yellow and six times in green, but never in red).

In this study, it is intriguing that there was no reliable effect of contingency; that is, participants responded to low-contingency items (M = 518 ms) about as quickly as they did to high-contingency items (M = 525 ms). Apparently, when a particular high-contingency color−word stimulus pairing occurs as often as does a particular low-contingency color−word stimulus pairing, there may be no contingency effect. This fits with instance theory (Logan, 1988, 2002) and Schmidt's (2013, 2018; Schmidt et al., 2016) PEP model because the number of instances (word plus color) of each of these stimuli—whether LO contingency−HI frequency or HI contingency−LO frequency—is identical in memory. This implicates frequency as playing a key role in the contingency effect: Part of the reason for high-contingency items being responded to more quickly is that they are simply experienced more often across trials. Contrasting Lin's (2015) Experiment 6 with her Experiment 7 makes it clear that the frequency−contingency relation is a complex one, worthy of further study.

Varying Contingency Proportion

Perhaps the most fundamental question to ask about contingency learning concerns the influence that the degree of contingency has on learning. Noah Forrin and I (Forrin & MacLeod, 2018a) decided to take on this question using a parametric variation in the proportion of HI trials. With the notable exception of Schmidt et al. (2007), almost all of the research thus far had used the 80% HI, 20% LO (at the item level, 80-10-10) procedure depicted in Table 1. We included this condition as a benchmark but expanded the range to encompass 70-15-15, 60-20-20, 50-25-25, and 40-30-30. As usual, we tested large samples (50 participants for each of the five groups) to obtain a clear data pattern. Figure 5 displays a simplified version of the data from that study, showing the difference in response time between HI and LO trials as a function of the congruency proportion.

It is apparent from these data that the extent of contingency learning is a very direct outgrowth of the contingency proportion. People's responses distinguish HI from LO until the difference between HI and LO becomes close to numerically indistinguishable (i.e., 40-30-30). Particularly noteworthy is the observation that the decline in the congruency learning effect with decreasing contingency proportion was not significantly related to the time to respond to HI trials; instead, it was significantly related to the time to respond to LO trials. It appeared that as LO trials moved from relatively common (40-30-30) to relatively uncommon (80-10-10), they incurred an increasing cost. Consistent with the view that participants are building expectations about the likelihood of particular color–word pairings, which fits with the PEP model, as LO trials become rarer, they also become considerably less expected. The result is that when these LO trials do appear, they are quite surprising and hence slow to respond to.

In fact, over our experiments thus far, both published and unpublished, we have often seen the pattern that changes in the contingency learning effect (measured by LO–HI) depend more on changes in responding to LO trials than on responding to HI trials. This is certainly not unprecedented: In the attention literature, rare events have often been shown to powerfully affect responding (e.g., Hon & Tan, 2013; Wolfe, Horowitz, & Kenner, 2005). It is noteworthy that in Forrin and MacLeod (2018a) we observed that awareness of the contingencies, as measured by a postexperiment questionnaire, declined regularly as the proportion of HI trials declined. This was true both for subjective awareness, where in a postexperiment questionnaire participants were told that each word appeared most often in one color and were asked whether they had noticed this, and for objective awareness, where subsequently in the questionnaire participants were asked to identify the color in which each word had most often appeared. Objective awareness in particular was positively correlated with the size of the contingency learning effect, which is consistent with the earlier findings of Schmidt and De Houwer (2012a,2012b; Schmidt et al., 2007), showing a larger effect for subjectively aware and for informed participants. And when LO and HI trials were examined separately, this positive correlation stemmed from the HI trials, consistent with the prediction of the PEP model, where speeding of HI trials, not slowing of LO trials, should be what is observed.[3]

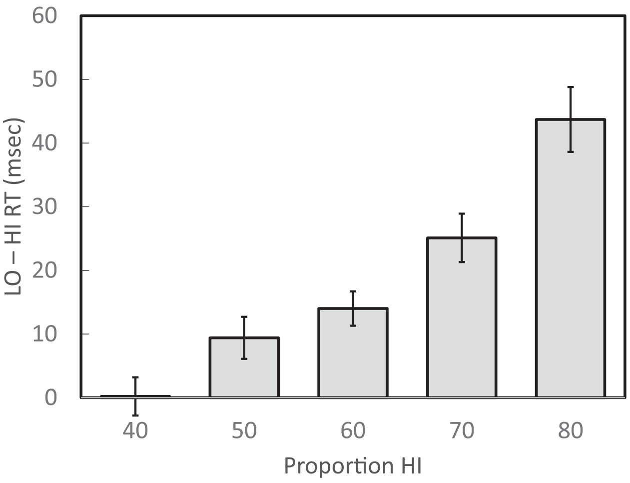

Figure 5. Differences in mean response time between high-contingency (HI) and low-contingency (LO) items—the contingency learning effect—as a function of congruency proportion. Error bars are standard errors of the means. RT = response time. The data are from Forrin and MacLeod (2018a).

Relative Speed of Processing and Stimulus–Response Compatibility

In the voluminous literature on the Stroop (1935) color–word interference phenomenon (see MacLeod, 1991, for a review), until the early 1970s the longstanding explanation of interference was the relative speed of processing account (see Dyer, 1973; Morton & Chambers, 1973). Why is it so slow to name the color of an incompatible word, such as saying "red" to the word *BLUE* printed in red? The reason, according to the relative speed account, is that processing of the word is faster than is processing of the color, so the word "gets in the way" (entering the response buffer first). That is also why interference does not occur in the other direction, when the task is to read the word: Processing of the word occurs prior to processing of the color. Indeed, Cattell (1886) had long ago demonstrated this pattern in his dissertation. But research in the 1980s (e.g., Dunbar & MacLeod, 1984; Glaser & Glaser, 1982) seriously undermined this simple explanation, and the relative speed of processing account of Stroop interference rightly fell out of favor.

In the Stroop case, the responses on the two dimensions—word and color—are semantically related, a feature that complicates the explanation. But in the color–word contingency learning paradigm, the responses on the word and color dimensions are unrelated. Perhaps the simple relative speed account could capture this less complicated situation? Maybe contingency learning occurs in this task because the processing of the words is faster than is the processing of the colors, so the words can influence color response time. That is what Noah Forrin and I set out to determine (Forrin & MacLeod, 2017). In our first experiment ($N = 32$), we had participants respond to the colors as usual in one block of trials but to the words in the other block of trials. We were initially surprised when we observed equivalent contingency learning in the two

---

[3] In our data, there was a strong negative correlation between the proportion of LO trials and response times for LO trials (i.e., fewer LO trials, slower response times to LO trials). Consequently, as the proportion of LO trials decreased, the magnitude of the contingency learning effect increased. This is not predicted by the PEP model, but we do not see this as a problem: The model, in its retrieval of instances, is quite capable of handling slowing by rare events, given retrieval of few LO instances coupled with competition from retrieval of multiple HI instances.

blocks, having expected reduced learning when responding to words. But we quickly realized that responding by key-pressing (as in previous studies) required a novel mapping of the responses to the keys, likely overwhelming any difference in processing speed between the two response modes.

We tested this in the next two experiments (*N* = 60 in each). In Experiment 2, we switched to vocal responding, a much more "natural" response mode for words. The data pattern changed dramatically: Now there was contingency learning when saying the color names aloud but no evidence of contingency learning when saying the words aloud. This result is entirely in line with the relative speed of processing prediction—that the fast words should influence the slow colors but not vice versa. And in the third experiment, when we gave the color information a head start by presenting it separately first, the contingency learning effect was reduced for color naming and reappeared for word reading, also coinciding with the prediction of the relative speed account.

Since the Forrin and MacLeod (2017) article, Schmidt (2018) has presented a modified explanation based on his PEP model. In this response, he argued in favor of a key role for relevant stimulus–response compatibility. Although accepting that words are dealt with faster than are colors when responding is vocal, he maintained (p. 328) that "the advantage that words have over colours with a vocal response (reading/naming) is not a benefit in stimulus-processing speed but a benefit in the compatibility between targets and responses (i.e., response-selection speed)." Essentially, one must take into consideration the stimulus–response mapping. Put more simply, words are responded to faster in only certain contexts, such as reading aloud. I agree that this more nuanced perspective on the results of Forrin and MacLeod (2017) is to be preferred.

Generalizing Across Dimensions

From this point forward, I consider ongoing, as-yet unpublished research from our laboratory, the aim being to further characterize basic processes in contingency learning. Consider first a simple question: Can one learn contingencies that occur across different dimensions? A straightforward way to begin addressing this question would be to explore learning cross-modal contingencies. Sébastien Lauzon, Noah Forrin, and I addressed this issue using auditory–visual connections (Lauzon, Forrin, & MacLeod, 2017). On the visual dimension, we switched from words to shapes (substituting for the usual three words the three shapes square, circle, and triangle). On the auditory dimension, we used three distinct tones (high, medium, and low). The layout of a block of trials corresponded closely to that shown in Table 1 for words and colors, and responding was again done using three keys on the keyboard. Participants completed two sets (several blocks) of trials, one set responding to the shape and one set responding to the tone of each stimulus. As Figure 6 shows (*N* = 40), contingency learning was readily established in both directions.[4]

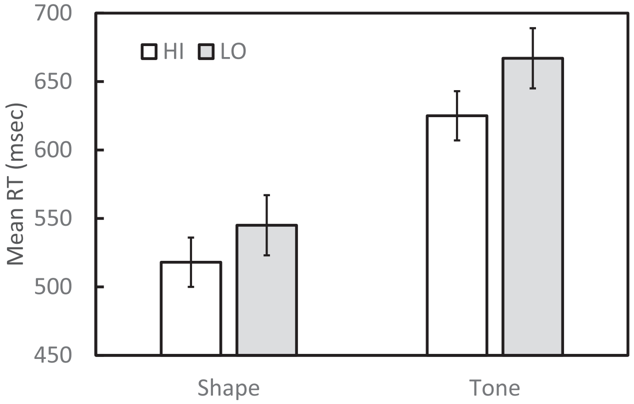

Figure 6. Mean response time (RT) in an auditory−visual contingency learning task. Participants identified shapes during one block of trials and tones during the other block, with both showing robust and equivalent learning. Error bars are standard errors of the means. HI = high contingency; LO = low contingency; RT = response time. The data are from Lauzon, Forrin, and MacLeod (2017, unpublished).

Olivia Lin, Noah Forrin, and I approached this dimensional issue in a different way (Lin, Forrin, & MacLeod, 2016). Would contingency learning be affected by whether the two dimensions of stimuli are integrated (the word is in color) versus separated (the color information is presented adjacent to the word)? Certainly, in the Stroop situation, this makes a substantial difference: There is considerably less interference for separated than for integrated stimuli (MacLeod, 1998). In the integrated condition, our stimuli were words in colors, as in Table 1. There were, in fact, two conditions using separated stimuli: (1) a color bar was centered on the screen and the word appeared simultaneously either above or below the bar and (2) the word was centered and the color bar appeared simultaneously above or below the word. As Figure 7 shows, there certainly was contingency learning regardless of stimulus format but, like the Stroop findings, the effect was much smaller for separated dimensions (11 ms) than for integrated dimensions (41 ms).[5] It is noteworthy that, once again, the difference was in the LO trials, which were considerably slower in the integrated case. This would seem to fit well with the relative speed of processing account too: Separating the word information by displaying it above or below the focal color should slow down processing of the word, giving it less of an opportunity to influence responding to the color.

These two studies provide evidence of the robustness of contingency learning across different stimulus types and modalities. A circumstance under which a contingency (greater than 40-30-30) cannot be learned, and learned readily, has yet to be seen. But sometimes the learning can be harder than others, as the most recent line of research demonstrates in the next section.

Prior Experience and Impeded Learning

Of late, I have come to see contingency learning as a type of human conditioning and to realize that many fundamental conditioning principles likely apply in learning simple contingencies. Indeed, Schmidt and De Houwer (2018) evidently are thinking

---

[4] Although not significantly different, the size of the HI–LO difference was nevertheless smaller when identifying shapes (27 ms), the faster modality to respond to, than when identifying tones (42 ms), the slower modality to respond to. This fits with the relative speed of processing account but could also simply reflect the fact that slower overall response times often result in larger condition difference scores.

[5] The data in Figure 7 are for the bar-centered separate condition, but the data for the word-centered separate condition were virtually identical.

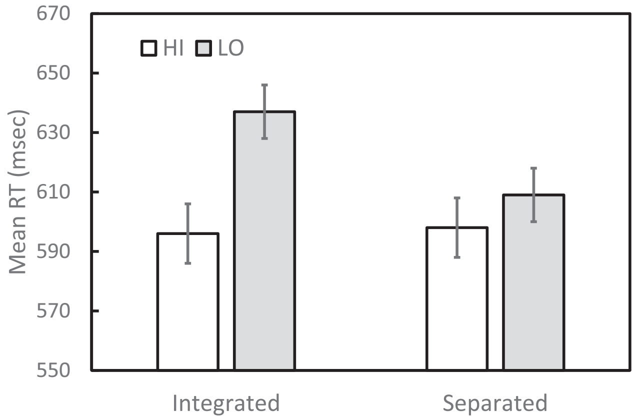

Figure 7. Mean response time in an integrated versus separated visual contingency learning task. For integrated trials, the word was printed in the color; for separated trials, a color bar was centered and the word appeared in black either above or below the color bar. Error bars are standard errors of the means. HI = high contingency; LO = low contingency; RT = response time. The data are from Lin, Forrin, & MacLeod (2016, unpublished).

along similar lines: Their most recent contingency learning work investigating blocking and overshadowing, two basic conditioning situations, clearly illustrates this. I briefly describe here an ongoing project that Noah Forrin and I are carrying out in which we ask the very Skinnerian question "Now that we can readily produce contingency learning, are there circumstances under which that learning could be made more difficult?" The first situation that occurred to us was the one called latent inhibition in the conditioning literature. Originating with Lubow and Moore (1959), this phenomenon refers to the rather counterintuitive finding that nonhuman animals have a more difficult time conditioning to a stimulus (as the conditioned stimulus) when that stimulus has previously occurred without any contingency, despite its heightened familiarity. So, if a rat has experienced a buzzer repeatedly in the recent past, and if one later tries to teach the rat that the buzzer is a conditioned stimulus for food, the rat has more trouble learning this than if the rat had no prior exposure to the buzzer. Essentially, the rat has initially learned that the buzzer is not meaningful and subsequently must overcome that to learn that the buzzer has become meaningful.

Might there be an analogous situation in human contingency learning? We decided to compare contingency learning in the typical paradigm displayed in Table 1 under two different conditions of prior experience (Forrin & MacLeod, 2018b). So the typical contingencies-on paradigm represents the test (second) phase here. The manipulation took place in the experience (first) phase, where there were two conditions. In both, the initial experience involved the same three colors but did not involve any contingency: Every word appeared equally often in every color across trials. In the different experience condition, the three words were different from those in the critical test phase; this is our baseline, where no influence of the first phase was expected on the test phase. The second experience condition was intended to capture the latent inhibition idea: Here, the words in the first phase were the same as those in the test phase, but again there was no contingency, with each word presented equally often in each color. We anticipated that this would cause latent inhibition, such that learning in the critical contingencies-on phase should be harder in this second condition.

The data shown in Figure 8, with 100 participants per group, bear out this prediction. Consider first the baseline condition, shown in the top panel. When different words are present in the experience and test phases, the experience blocks do not influence contingency learning in the test blocks: Learning is evident right away in the first block, and the advantage of HI over LO is consistent across the four test blocks. Now consider the latent inhibition condition, shown in the bottom panel. When the same words are present in the two phases, learning is not evident until the second test block. The prior noncontingent experience has disrupted the subsequent learning of a contingency, just as in the latent inhibition situation. It is interesting to contrast this to the Lin and MacLeod (2018) study described earlier, where presumably the no-contingency series of blocks in the second phase did not cause latent inhibition in the third phase because the same contingencies had been present in the first and third phases—the conditioning had been previously established. Noah Forrin and I intend to continue to explore interfering factors in contingency learning, factors such as proactive interference.

Interim Progress Report

In a short period of time, a considerable amount has been learned about human contingency learning, and hence about the

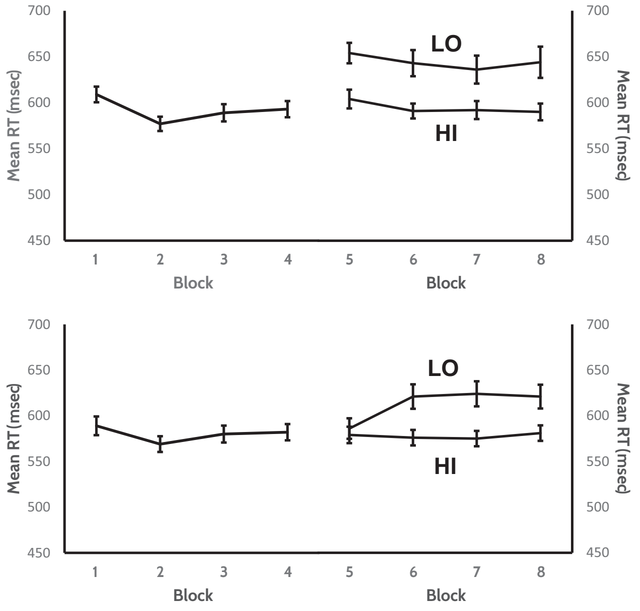

Figure 8. Mean response times for the two experience conditions (Blocks 1–4), where no contingency was present, and for the test condition (Blocks 5–8), where both high-contingency (HI) and low-contingency (LO) trials were present. The manipulation was in the experience conditions. The top panel shows the experience condition where the words in Blocks 1–4 were different from the words in Blocks 5–8, the baseline condition. The bottom panel shows the experience condition where the words in Blocks 1–4 were the same as the words in Blocks 5–8, the latent inhibition condition. Error bars are standard errors of the means. HI = high contingency; LO = low contingency. The data are from Forrin and MacLeod (2018b, unpublished).

underpinnings of human association learning. It now does appear that many of the “laws” of associative learning observed in conditioning research with nonhuman animals hold true as well for this type of human learning. Pursuing this perspective will likely be a fruitful direction for future human contingency learning research. For the present, however, this simple contingency learning situation demonstrates the remarkable skill that humans have in learning associations. Evidently, this learning is ordinarily at least partially implicit but it can also have an explicit component. And as the proportion contingency decreases, so too does the magnitude of the contingency learning effect (and participant awareness of the contingencies decreases as well). It is also clear that these associations are rapidly acquired and rapidly extinguished, and that contingency learning involves both a benefit for the high-contingency items and a cost for the low-contingency items. Some alternative explanations for the contingency learning effect have been ruled out, such as that high-contingency trials simply benefit from their greater repetition: Repetition matters, but it is not the whole story. The learning is very flexible in terms of the particular associations being learned, as illustrated when the relevant and irrelevant features are separated, either in visual space or across modalities. And this type of learning is very sensitive to context, as when prior experience can contraprepare a person for new contingency learning.

At the theoretical level, the PEP model (Schmidt, 2013, 2018; Schmidt et al., 2016) does a good job of capturing many of the findings thus far and may well be modified to incorporate some apparently discrepant findings (e.g., our repeated observation of the prevalent influence of low-contingency items). And again, one can look to the animal-conditioning literature not just for intriguing parallels at the empirical level but also for inspiration in developing a comprehensive theory of human contingency learning. Theories such as the Rescorla and Wagner (1972) conditioning account (see Van Hamme & Wasserman, 1994, for an updated, extended version) may well be adjusted to explain human contingency learning, at the same time offering intriguing new predictions to be tested. For example, could one conceptualize low-contingency items as conditioned inhibitors, and how might one test this idea?

My view is that this type of simple contingency learning is informative about one of people’s most fundamental and most frequently deployed skills—learning to associate stimuli or dimensions. Having long been interested in attention and memory (see, e.g., MacLeod, 2010, 2013), I realize that learning is the foundation on which both attention and memory rest. People tune attention based on learning (see, e.g., B. A. Anderson, 2016), and they use memory to capture learning so that it can be used at a later time, obviating the need for more costly (in terms of time and effort) online cognitive analysis. Moreover, the bulk of people’s learning certainly is associative in nature: They learn the relations between elements in their world so that they may better function in that world, whether taking account of correlations or linking causes to effects.

Résumé

Les gens sont exceptionnellement habiles pour établir des connexions dans des contextes d’apprentissage par mémoire associative. En effet, il s’agit fort probablement de la compétence cognitive la plus fréquemment déployée. L’apprentissage de contingences – qu’un événement est plus probable lorsqu’en présence d’un autre – est une forme cruciale d’apprentissage associatif qui permet de faire des prédictions avec succès, augmentant ainsi la vitesse et la précision de la réponse aux événements dans le monde. Le présent article décrit un programme de recherche sur les éléments fondamentaux de l’apprentissage de contingences lesquels sont sous-jacents à l’acquisition d’information causale et corrélationnelle dans le monde qui les entoure.

Mots-clés : association, contingence, apprentissage.

References

Anderson, B. A. (2016). The attention habit: How reward learning shapes attentional selection. *Annals of the New York Academy of Sciences, 1369,* 24–39. http://dx.doi.org/10.1111/nyas.12957

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 6,* 451–474. http://dx.doi.org/10.1016/0010-0285(74)90021-8

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition, 121,* 127–132. http://dx.doi.org/10.1016/j.cognition.2011.06.010

Carlson, K. A., & Flowers, J. H. (1996). Intentional versus unintentional use of contingencies between perceptual events. *Perception & Psychophysics, 58,* 460–470. http://dx.doi.org/10.3758/BF03206821

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind, os-XI,* 63–65. http://dx.doi.org/10.1093/mind/os-XI.41.63

Chance, P. (2008). *Learning and behavior*. Belmont, CA: Wadsworth.

De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and theories on human contingency learning. *Quarterly Journal of Experimental Psychology, 55,* 289–310. http://dx.doi.org/10.1080/02724990244000034

Dunbar, K., & MacLeod, C. M. (1984). A horse race of a different color: Stroop interference patterns with transformed words. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 622–639. http://dx.doi.org/10.1037/0096-1523.10.5.622

Dyer, F. N. (1973). The Stroop phenomenon and its use in the stlldy of perceptual, cognitive, and response processes. *Memory & Cognition, 1,* 106–120. http://dx.doi.org/10.3758/BF03198078

Forrin, N. D., & MacLeod, C. M. (2015). *Does repetition contribute to contingency learning?* Unpublished data, University of Waterloo, Waterloo, Ontario, Canada.

Forrin, N. D., & MacLeod, C. M. (2017). Relative speed of processing determines color-word contingency learning. *Memory & Cognition, 45,* 1206–1222. http://dx.doi.org/10.3758/s13421-017-0721-4

Forrin, N. D., & MacLeod, C. M. (2018a). Contingency proportion systematically influences contingency learning. *Attention, Perception, & Psychophysics, 80,* 155–165. http://dx.doi.org/10.3758/s13414-017-1424-4

Forrin, N. D., & MacLeod, C. M. (2018b). *When prior experience is detrimental to contingency learning*. Unpublished data, University of Waterloo, Waterloo, Ontario, Canada.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 875–894. http://dx.doi.org/10.1037/0096-1523.8.6.875

Hall, G. (2002). Associative structures in Pavlovian and instrumental conditioning. In R. Gallistel (Ed.), *Stevens’ handbook of experimental psychology: Vol. 3: Learning, motivation, and emotion* (3rd ed., pp. 1–45). New York, NY: Wiley.

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review, 56,* 51–65. http://dx.doi.org/10.1037/h0062474

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388. http://dx.doi.org/10.1037/0096-3445.108.3.356

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist, 39,* 1372–1388. http://dx.doi.org/10.1037/0003-066X.39.12.1372

Hebb, D. O. (1949). *The organization of behavior*. New York, NY: Wiley.

Heyes, C. (2012). Simple minds: A qualified defence of associative learning. *Philosophical Transactions of the Royal Society of London: Series B. Biological Sciences, 367,* 2695–2703. http://dx.doi.org/10.1098/rstb.2012.0217

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers, 16,* 96–101. http://dx.doi.org/10.3758/BF03202365

Hon, N., & Tan, C.-H. (2013). Why rare targets are slow: Evidence that the target probability effect has an attentional locus. *Attention, Perception, & Psychophysics, 75,* 388–393. http://dx.doi.org/10.3758/s13414-013-0434-0

Hon, N., Yap, M. J., & Jabar, S. B. (2013). The trajectory of the target probability effect. *Attention, Perception, & Psychophysics, 75,* 661–666. http://dx.doi.org/10.3758/s13414-013-0429-x

Horner, V., Carter, J. D., Suchak, M., & de Waal, F. B. M. (2011). Spontaneous prosocial choice by chimpanzees. *Proceedings of the National Academy of Sciences of the United States of America, 108,* 13847–13851. http://dx.doi.org/10.1073/pnas.1111088108

Kelly, M. H., & Martin, S. (1994). Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition, and language. *Lingua, 92,* 105–140. http://dx.doi.org/10.1016/0024-3841(94)90339-5

Lauzon, S., Forrin, N. D., & MacLeod, C. M. (2017). *Cross-modal contingency learning*. Unpublished data, University of Waterloo, Waterloo, Ontario, Canada.

Levin, Y., & Tzelgov, J. (2016). Contingency learning is not affected by conflict experience: Evidence from a task conflict-free, item-specific Stroop paradigm. *Acta Psychologica, 164,* 39–45. http://dx.doi.org/10.1016/j.actpsy.2015.12.009

Lin, O. Y.-H. (2015). *Characterizing colour-word contingency learning (Doctoral dissertation)*. Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada.

Lin, O. Y.-H., Forrin, N. D., & MacLeod, C. M. (2016). *The influence of separation versus integration on contingency learning*. Unpublished data, University of Waterloo, Waterloo, Ontario, Canada.

Lin, O. Y.-H., & MacLeod, C. M. (2018). The acquisition of simple associations as observed in color-word contingency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44,* 99–106. http://dx.doi.org/10.1037/xlm0000436

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527. http://dx.doi.org/10.1037/0033-295X.95.4.492

Logan, G. D. (2002). An instance theory of attention and memory. *Psychological Review, 109,* 376–400. http://dx.doi.org/10.1037/0033-295X.109.2.376

Lubow, R. E., & Moore, A. U. (1959). Latent inhibition: The effect of nonreinforced pre-exposure to the conditional stimulus. *Journal of Comparative and Physiological Psychology, 52,* 415–419. http://dx.doi.org/10.1037/h0046700

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203. http://dx.doi.org/10.1037/0033-2909.109.2.163

MacLeod, C. M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. *Memory & Cognition, 26,* 201–211. http://dx.doi.org/10.3758/BF03201133

MacLeod, C. M. (2010). When learning met memory. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 64,* 227–240. http://dx.doi.org/10.1037/a0021699

MacLeod, C. M. (2013). The six R's of remembering. *Canadian Psychology/Psychologie canadienne, 54,* 38–49. http://dx.doi.org/10.1037/a0030955

McConnell, J. V., Jacobson, A. L., & Kimble, D. P. (1959). The effects of regeneration upon retention of a conditioned response in the planarian. *Journal of Comparative and Physiological Psychology, 52,* 1–5. http://dx.doi.org/10.1037/h0048028

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85,* 207–238. http://dx.doi.org/10.1037/0033-295X.85.3.207

Miller, J. (1987). Priming is not necessary for selective-attention failures: Semantic effects of unattended, unprimed letters. *Perception & Psychophysics, 41,* 419–434. http://dx.doi.org/10.3758/BF03203035

Miller, R. R., & Escobar, M. (2002). Learning: Laws and models of basic conditioning. In R. Gallistel (Ed.), *Stevens' handbook of experimental psychology: Vol. 3: Learning, motivation, and emotion* (3rd ed., pp. 47–102). New York, NY: Wiley.

Mordkoff, J. T., & Halterman, R. (2008). Feature integration without visual attention: Evidence from the correlated flankers task. *Psychonomic Bulletin & Review, 15,* 385–389. http://dx.doi.org/10.3758/PBR.15.2.385

Morton, J., & Chambers, S. M. (1973). Selective attention to words and colours. *Quarterly Journal of Experimental Psychology, 25,* 387–397. http://dx.doi.org/10.1080/14640747308400360

Pavlov, I. P. (1927). *Conditioned reflexes*. New York, NY: Oxford University Press.

Petrides, M. (1985). Deficits on conditional associative-learning tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia, 23,* 601–614. http://dx.doi.org/10.1016/0028-3932(85)90062-4

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology, 66,* 1–5. http://dx.doi.org/10.1037/h0025984

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

Roberts, B. R. T., Forrin, N. D., & MacLeod, C. M. (2018). *Single versus multiple contingencies in color-word contingency learning*. Unpublished data, University of Waterloo, Waterloo, Ontario, Canada.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996, December 13). Statistical learning by 8-month-old infants. *Science, 274,* 1926–1928. http://dx.doi.org/10.1126/science.274.5294.1926

Saigusa, T., Tero, A., Nakagaki, T., & Kuramoto, Y. (2008). Amoebae anticipate periodic events. *Physical Review Letters, 100,* 018101. http://dx.doi.org/10.1103/PhysRevLett.100.018101

Schmidt, J. R. (2013). The parallel episodic processing (PEP) model: Dissociating contingency and conflict adaptation in the item-specific proportion congruent paradigm. *Acta Psychologica, 142,* 119–126. http://dx.doi.org/10.1016/j.actpsy.2012.11.004

Schmidt, J. R. (2018). Best not to bet on the horserace: A comment on Forrin and MacLeod (2017) and a relevant stimulus-response compatibility view of colour-word contingency learning asymmetries. *Memory & Cognition, 46,* 326–335. http://dx.doi.org/10.3758/s13421-017-0755-7

Schmidt, J. R., Crump, M. J. C., Cheesman, J., & Besner, D. (2007). Contingency learning without awareness: Evidence for implicit control. *Consciousness and Cognition, 16,* 421–435. http://dx.doi.org/10.1016/j.concog.2006.06.010

Schmidt, J. R., & De Houwer, J. (2012a). Adding the goal to learn strengthens learning in an unintentional learning task. *Psychonomic*

Bulletin & Review, 19, 723–728. http://dx.doi.org/10.3758/s13423-012-0255-5

Schmidt, J. R., & De Houwer, J. (2012b). Contingency learning with evaluative stimuli: Testing the generality of contingency learning in a performance paradigm. *Experimental Psychology, 59,* 175–182. http://dx.doi.org/10.1027/1618-3169/a000141

Schmidt, J. R., & De Houwer, J. (2012c). Does temporal contiguity moderate contingency learning in a speeded performance task? *Quarterly Journal of Experimental Psychology, 65,* 408–425. http://dx.doi.org/10.1080/17470218.2011.632486

Schmidt, J. R., & De Houwer, J. (2012d). Learning, awareness, and instruction: Subjective contingency awareness does matter in the colour-word contingency learning paradigm. *Consciousness and Cognition, 21,* 1754–1768. http://dx.doi.org/10.1016/j.concog.2012.10.006

Schmidt, J. R., & De Houwer, J. (2016a). Contingency learning tracks with stimulus-response proportion: No evidence of misprediction costs. *Experimental Psychology, 63,* 79–88. http://dx.doi.org/10.1027/1618-3169/a000313

Schmidt, J. R., & De Houwer, J. (2016b). Time course of colour-word contingency learning: Practice curves, pre-exposure benefits, unlearning, and relearning. *Learning and Motivation, 56,* 15–30. http://dx.doi.org/10.1016/j.lmot.2016.09.002

Schmidt, J. R., & De Houwer, J. (2018). Cue competition and incidental learning: No blocking or overshadowing in the colour-word contingency learning procedure without instructions to learn. Manuscript submitted for publication.

Schmidt, J. R., De Houwer, J., & Besner, D. (2010). Contingency learning and unlearning in the blink of an eye: A resource dependent process. *Consciousness and Cognition, 19,* 235–250. http://dx.doi.org/10.1016/j.concog.2009.12.016

Schmidt, J. R., De Houwer, J., & Rothermund, K. (2016). The parallel episodic processing (PEP) model 2.0: A single computational model of stimulus-response binding, contingency learning, power curves, and mixing costs. *Cognitive Psychology, 91,* 82–108. http://dx.doi.org/10.1016/j.cogpsych.2016.10.004

Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology, 60,* 291–309. http://dx.doi.org/10.1080/17470210601000581

Sternberg, D. A., & McClelland, J. L. (2012). Two mechanisms of human contingency learning. *Psychological Science, 23,* 59–68. http://dx.doi.org/10.1177/0956797611429577

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662. http://dx.doi.org/10.1037/h0054651

Thorndike, E. L. (1911). *Animal intelligence: Experimental studies*. Advance online publication. http://dx.doi.org/10.5962/bhl.title.55072

Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgements: The role of nonpresentation of compound stimulus elements. *Learning and Motivation, 25,* 127–151. http://dx.doi.org/10.1006/lmot.1994.1008

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology, 34,* 1289–1309. http://dx.doi.org/10.1037/0012-1649.34.6.1289

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005, May 25). Rare items often missed in visual searches. *Nature, 435,* 439–440. http://dx.doi.org/10.1038/435439a

Zacks, R. T., & Hasher, L. (2002). Frequency processing: A twenty-five year perspective. In P. Sedlmeier & T. Betsch (Eds.), *Frequency processing and cognition* (pp. 21–36). http://dx.doi.org/10.1093/acprof:oso/9780198508632.003.0002

Received August 7, 2018
Revision received October 25, 2018
Accepted October 29, 2018 ■