

The Acquisition of Simple Associations as Observed in Color–Word Contingency Learning

Olivia Y.-H. Lin and Colin M. MacLeod
University of Waterloo

Three experiments investigated the learning of simple associations in a color–word contingency task. Participants responded manually to the print colors of 3 words, with each word associated strongly to 1 of the 3 colors and weakly to the other 2 colors. Despite the words being irrelevant, response times to high-contingency stimuli and to low-contingency stimuli quickly diverged. This high–low difference remained quite constant over successive blocks of trials, evidence of stable contingency learning. Inclusion of a baseline condition—an item having no color–word contingency—permitted separation of the contingency learning effect into 2 components: a cost due to low contingency and a benefit due to high contingency. Both cost and benefit were quick to acquire, quick to extinguish, and quick to reacquire. The color–word contingency task provides a simple way to directly study the learning of associations.

Keywords: association, contingency learning, cost, benefit

People are remarkably adept at learning the associations that are all around them, whether they do so intentionally or unintentionally. Consequently, this ability and its intellectual underpinnings have fascinated thinkers from at least the time of Plato. The best known analysis of association is Aristotle’s four laws—contiguity, frequency, similarity, and contrast—an analysis that went on to shape the thinking of philosophers (e.g., John Stuart Mill) and indeed psychologists (e.g., B. F. Skinner). Particularly noteworthy, the basics of behaviorism have strong ties to Aristotelian ideas about association, focusing as they do on links between stimuli and stimuli and between stimuli and responses. The principles of contiguity (i.e., events occurring close to each other, whether in space or in time, tend to become linked) and frequency (i.e., the more often events are linked, the more strongly they will be connected) are deeply ingrained in psychological thinking.

Not surprisingly, then, association has played a central role in thinking about learning and memory from the earliest writings in psychology (e.g., Calkins, 1894; Ebbinghaus, 1885/1913). Yet remarkably, work on memory in the cognitive literature has largely failed to study associations directly. Although researchers talk about memory in associative terms, they do so primarily by relating stimulus events in the world to hypothetical constructs in the mind. Only occasionally, as in free recall, have direct associations

between stimuli been considered, but even then there are models that deny the need to do so (e.g., the search of associative memory model of Raaijmakers & Shiffrin, 1981). Serial learning, for much of its history thought to rely on the formation of interitem associations, has not been explained that way since the dawn of the cognitive revolution. Even paired-associate learning, which might seem the quintessential case of associative learning, has emphasized acquisition of the responses to the stimuli over the linkage between the paired stimuli and responses.

So how might one directly study associative learning? A simple situation that puts the emphasis squarely on stimulus to stimulus and stimulus to response connections is needed. One possible attack is through the learning of contingencies, learning that involves both contiguity and frequency. *Contingency learning* can be defined as the acquisition of knowledge about covariations or correlations between stimuli, between responses, or between stimuli and responses. Like other organisms, people are prepared to learn about the co-occurrences of stimuli, especially those that frequently co-occur. Such learning influences performance in many ways, including both the speed and the accuracy of responding (see, e.g., De Houwer & Beckers, 2002). There also is considerable evidence that this kind of learning occurs rapidly, sometimes even within a single trial (e.g., Lewicki, 1985), and that it can occur beneath awareness (Gehring, Gratton, Coles, & Donchin, 1992). Association is a fundamental element of cognition (see, e.g., Shanks, 2010).

In recent years, a simple procedure has begun to be used to explore the fundamental features of human contingency learning. The paradigm, called *color–word contingency learning*, was developed by Schmidt, Crump, Cheesman, and Besner (2007). Its roots are in the work of Musen and Squire (1993) exploring the familiar Stroop (1935) interference paradigm. In the color–word contingency task, participants make identification responses to the colors in which unrelated words are presented, for example, pressing a key for the color red when any word (e.g., *TABLE* or

This article was published Online First July 31, 2017.

Olivia Y.-H. Lin and Colin M. MacLeod, Department of Psychology, University of Waterloo.

This research was supported by Natural Sciences and Engineering Research Council of Canada Discovery Grant A7459. The experiments are taken from the dissertation of Olivia Y.-H. Lin under the supervision of the second author.

Correspondence concerning this article should be addressed to Colin M. MacLeod, Department of Psychology, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada. E-mail: cmacleod@uwaterloo.ca

HORSE) is presented in red. In fact, the words themselves could simply be ignored and performance could still be perfect. But this paradigm examines responding to stimulus events that have differing contingencies, determined by the proportion of trials on which a given word appears in a given color. On high-contingency trials, a word is most often presented in a specific color (e.g., *PLATE* is usually presented in green); on low-contingency trials, that word is presented in another color (e.g., *PLATE* is only relatively rarely presented in yellow or in red). The result is two sets of stimulus–stimulus (color–word) associations: strong and weak.

To illustrate more concretely, imagine a block of 30 trials in which participants press one of three keys corresponding to three colors (e.g., red, green, and yellow). Color is the response-relevant dimension, but color is conveyed in the form of one of three words (*TABLE*, *HORSE*, or *PLATE*). Each word—the irrelevant dimension—is presented in one of the colors eight times (*TABLE* usually in red, *PLATE* usually in green, and *HORSE* usually in yellow) and in the other colors one time each (e.g., *TABLE* appears once in green and once in yellow). These two stimulus–stimulus situations correspond, respectively, to the high-contingency and low-contingency conditions. Contingency learning involves developing differential responding to these two situations, whether consciously or unconsciously.

Despite the words being irrelevant for responding, the color–word contingencies are in fact learned. Very quickly, response times (and accuracies) separate for the two conditions, with faster and more accurate responding for the high-contingency stimuli relative to the low-contingency stimuli. This performance difference has been labeled *the contingency effect*. Schmidt and De Houwer have documented a number of features of color–word contingency learning, demonstrating the learning to be rapid, stable, resource demanding, and augmented by conscious awareness (see Schmidt & De Houwer, 2012a, 2012b, 2012c; Schmidt, De Houwer, & Besner, 2010). Their task provides a simple situation for quite directly investigating the learning of associations.

Cost and Benefit

Intuition suggests that the learning of contingencies should influence people’s behavior in a positive way, speeding them up and decreasing their errors. In other words, strong associations should help performance. However, it is also known that contingency-based performance can become so routine that any deviation from the familiar contingencies that have been learned can disrupt performance. It seems quite possible, then, that weak associations might hurt performance. Studies in the realm of attention and vigilance have observed such a cost in responding (i.e., slower and more error-prone responses) when events occur—or co-occur—with relatively low probability (Hon & Tan, 2013; Hon, Yap, & Jabar, 2013; Laberge & Tweedy, 1964; Wolfe et al., 2007). These costs are evident even when the accuracy of target detection is high. It is noteworthy that in a related learning situation—sequence learning—there has also been evidence of both cost and benefit occurring (e.g., Nattkemper & Prinz, 1997).

A fundamental question in any paradigm involving two such conditions is whether the difference between the conditions—in this case, the contingency effect—is the result of a benefit for the better performing condition, a cost for the poorer performing

condition, or both (cf. Jonides & Mack, 1984). This is a principal focus of the work reported in this article. Intuition, coupled with the attention literature, may suggest that the answer should be “both.” It is interesting that, until very recently, the only study to tackle the question did not reach that conclusion (Schmidt & Besner, 2008). On the basis of including what Schmidt and Besner referred to as a medium-contingency condition, where one word was presented equally often in each color to serve as a baseline, the authors argued forcefully that “the contingency effect is entirely facilitative for response latencies” (p. 517; see also Schmidt et al., 2010). However, it is intriguing that in the attention literature (e.g., Hon et al., 2013), the findings actually are more consistent with the response difference between high-probability and low-probability targets’ being driven by slowed responding to the low-probability targets rather than by speeded responding to the high-probability targets. And in studies that have manipulated contingency in the Stroop task (Hazeltine & Mordkoff, 2014) and in the flanker task (Carlson & Flowers, 1996; Miller, 1987), patterns have also suggested both benefit and cost.

Schmidt (2013; Schmidt, De Houwer, & Rothermund, 2016) has proposed a model of color–word contingency learning—parallel episodic processing (PEP)—along the lines of the instance theory of automaticity (Logan, 1988). Under the PEP account, information about each stimulus event (or instance)—including in our case the color, the word, the response, and the response time—is encoded into episodic memory. As trials progress, episodic memory builds up a database of prior instances. Encountering a new trial routinely results in retrieval of prior instances, weighted preferentially for more recent instances. Consequently, instances retrieved for a high-contingency trial will ordinarily be more relevant than for a low-contingency trial. The likelihood of high-contingency instances that are fast and that match the current trial is relatively great, thereby facilitating performance on high-contingency trials. For low-contingency trials, fewer instances will be relevant and more of these will be slow. As well, there will often be partial overlap with high-contingency instances, resulting in conflict. Together, these factors will impede performance on low-contingency trials. Put simply, there should be both benefit and cost.

The Current Study

Our experimental approach was closely modeled after the typical color–word contingency paradigm that we laid out in describing the work of Schmidt et al. (2007) but with the addition of a no-contingency baseline (as in Schmidt & Besner, 2008). In Experiment 1a, three words were assigned to the usual high-contingency versus low-contingency manipulation. Critically, we added a fourth word, which had no contingency and appeared equally often in each of the colors, thereby serving as a baseline against which to assess cost and benefit. In Experiment 1b, non-words were substituted for the words, replicating and generalizing the pattern. Experiment 2 extended these findings by examining the acquisition, extinction, and reacquisition of the contingency effect.

These experiments allowed us to explore the learning of simple contingencies, which we see as a kind of model world in which to study the acquisition of associations. After all, it has to be the connections between the colors and the words than underlie the

contingency effect, and the existence of the contingency effect is itself evidence of associative learning.

Experiments 1a and 1b

By comparison to the no-contingency item as a baseline, deviations that represent speeding up or slowing down (or reduced or increased errors) can be differentiated, providing indices of benefit and cost. We expected to observe both a benefit for high-contingency stimuli and a cost for low-contingency stimuli relative to the baseline stimuli. Experiment 1a used words as the irrelevant dimension, in keeping with previous color-word contingency learning studies. In Experiment 1b, where the irrelevant dimension was made up of nonwords, we examined whether familiar, meaningful stimuli are necessary for the contingency effect to develop, at the same time seeking to replicate and generalize the pattern in Experiment 1a. The literature on the Stroop effect has shown that nonwords cause less interference than do words (Klein, 1964; see MacLeod, 1991, for a review). No research in the color-word contingency learning paradigm has examined the influence of nonwords (although Schmidt & De Houwer, 2012a, did use nonwords as “primes” in a sequential word-word contingency situation). If nonwords produce similar cost and benefit in color-word contingency learning, then this would provide evidence (cf. Schmidt et al., 2007) that semantic processing of the stimuli is not necessary, indicating that the contingency effect appears to rely on more primitive learning, the learning of simple associations that are strong (high contingency) or weak (low contingency).

Method

Participants. Students at the University of Waterloo completed one session in exchange for course credit. In Experiment 1a, there were 31 participants ($M_{\text{age}} = 19.43$ years; 25 female). In Experiment 1b, there were also 31 participants ($M_{\text{age}} = 20.77$ years; 25 female). All had normal or corrected-to-normal vision.

Apparatus. All experiments were programmed using E-Prime Version 2.0 (Schneider, Eschman, & Zuccolotto, 2012), with all stimuli presented and responses recorded using a PC computer with a standard keyboard and a 17-in. flat-screen color monitor.

Materials and design. The design and procedure were identical for the two experiments; all that differed was whether words or nonwords were used. There were 30 practice trials, each consisting of a row of asterisks (*****) presented with equal probability in each of the three colors (yellow, green, red). This practice phase was included to acquaint participants with the response keys and with the experimental procedure more generally.

There were eight blocks of experimental trials, with 48 trials per block. In the experimental phase of Experiment 1a, four words (*MONTH*, *UNDER*, *PLATE*, and *CLOCK*) were presented in three colors: red, yellow, and green. In Experiment 1b, four pronounceable nonwords were substituted for the words: *FLABE*, *THROG*, *DWIPS*, and *BRASK*.

The proportions of three of the words (or nonwords) were manipulated such that each appeared in one color 83.33% of the time (high contingency) and in each of the other two colors 8.33% of the time (low contingency; total = 16.67%). An additional word (or nonword) was presented in each of the colors with equal probability (33.33%). This equal-probability item acted as the

baseline (no contingency). Of the total of 384 experimental trials for each participant, there were 240 observations in the high-contingency condition, 48 observations in the low-contingency condition, and 96 observations in the no-contingency baseline condition.

Procedure. The contingencies stayed the same across all eight experimental blocks. Other than the trial stimuli, all displays—such as fixation and feedback—were presented in white on a black background in 18-point Courier New font. All stimulus words (or nonwords) were presented in lower case at the center of the screen in 16-point Courier New font on a black background. All combinations of colors and words (or nonwords) were counterbalanced across participants; which color went with which word (or nonword) was randomly determined for each participant. On each trial, participants saw a central fixation cross for 250 ms, followed by a word (or nonword) in the same location until response or until 2,000 ms had elapsed. On trials that were incorrect or too long, feedback consisting of three Xs (XXX) was shown in white at the center of the screen for 1,000 ms. Participants used three fingers on their right hand to respond to the color of the word (or nonword) by pressing key *J* (red), *K* (yellow), or *L* (green). To assist them in remembering the response assignments, we placed colored stickers on these keys. The color-to-key mappings were constant across participants and experiments.

Results

Experiment 1a (words).

Errors and outliers. Overall, 2.97% of responses were errors. Mean error proportions were 2.38% for high contingency, 4.10% for low contingency, and 3.90% for no contingency. A repeated-measures analysis of variance (ANOVA) conducted on the three types of contingency revealed a significant main effect, $F(1.55, 32.54) = 14.50$, $MSE = .00$, $p < .001$, $\eta_p^2 = .408$, with a Greenhouse-Geisser correction. Two key comparisons followed. High-contingency stimuli were significantly less error-prone than were baseline stimuli, $t(27) = 5.12$, $SE = .004$, $p < .001$; although in the expected direction, low-contingency stimuli were not significantly more error-prone than were baseline stimuli, $t(21) = 1.63$, $SE = .010$, $p = .119$.

Response times (RTs). First, incorrect responses were removed from the analysis of RTs. Then response times less than 200 ms were removed as anticipations, after which RTs were trimmed using a common 2.5-standard deviation criterion by subject and type of contingency (see Van Selst & Jolicoeur, 1994, for discussion of trimming; see, e.g., Schmidt et al., 2010, for application of the same trimming procedure). Overall, 3.00% of the correct data were trimmed (3.03% of high contingency trials, 2.94% of low contingency trials, and 2.94% of baseline trials). We note that for this experiment and for both subsequent experiments, exactly the same findings were obtained without trimming.

Figure 1A presents the RT data for the three conditions as a function of block. We carried out a 3 (contingency: high, low, baseline) \times 8 (blocks) repeated-measures ANOVA. There was a significant main effect of contingency, $F(1.47, 44.16) = 17.32$, $MSE = 7,778.59$, $p < .001$, $\eta_p^2 = .366$. The main effect of block also was significant, $F(4.41, 132.30) = 2.86$, $MSE = 7,102.03$, $p = .022$, $\eta_p^2 = .087$, but there was no Contingency \times Block interaction, $F(7.13, 213.77) = 1.31$, $MSE = 4,729.54$, $p = .247$, $\eta_p^2 = .042$. Put simply, the pattern across blocks was very consis-

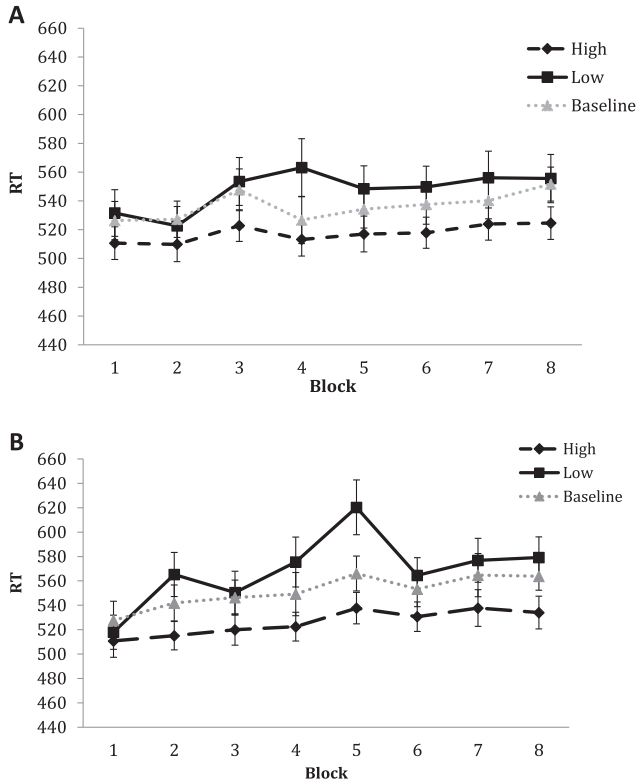


Figure 1. Experiment 1: Mean response times (RTs) for the high-contingency, low-contingency, and no-contingency (baseline) stimuli as a function of block. Panel A shows the data for Experiment 1a using words; Panel B shows the data for Experiment 1b using nonwords. Error bars depict one standard error of the corresponding mean.

ment, as Figure 1A shows. This also replicated the Schmidt et al. (2007, 2010) finding that the contingency effect emerges very quickly in the first block and remains quite stable thereafter.

The two comparisons of primary interest were then carried out. There was a reliable benefit, such that participants were 22 ms faster responding to high-contingency stimuli ($M = 517$ ms) than to no-contingency baseline stimuli ($M = 539$ ms), $t(30) = 5.28$, $SE = 4.31$, $p < .001$. There also was a reliable cost, with responding to low-contingency stimuli ($M = 555$ ms) 16 ms slower than to no-contingency stimuli, $t(30) = 2.11$, $SE = 7.60$, $p = .043$.

Experiment 1b (nonwords).

Errors and outliers. Overall, 3.17% of responses were errors. Mean error proportions were 2.31% for high contingency, 6.25% for low contingency, and 3.76% for no contingency (baseline). A repeated-measures ANOVA conducted on the three types of contingency revealed a significant main effect, $F(1.27, 38.07) = 14.96$, $MSE = .001$, $p < .001$, $\eta_p^2 = .333$. There was a significant benefit, with high-contingency stimuli less error-prone than baseline stimuli, $t(30) = 4.05$, $SE = .004$, $p < .001$. This time there also was a significant cost, with more errors for low-contingency stimuli than for baseline stimuli, $t(30) = 2.95$, $SE = .008$, $p = .006$.

Response times. First, incorrect responses were removed from the analysis of RTs. Next, all response times less than 200 ms were removed as anticipations. Finally, RTs were trimmed using the

2.5-SD criterion by subject and type of contingency. Overall, 2.81% of the correct data were trimmed (2.70% of high-contingency trials, 2.87% of low-contingency trials, and 3.07% of baseline trials).

Figure 1B presents the RT data as a function of contingency condition and block. We again carried out a 3 (contingencies) \times 8 (blocks) repeated-measures ANOVA. There was a main effect of contingency, $F(1.66, 49.93) = 32.49$, $MSE = 4,233.97$, $p < .001$, $\eta_p^2 = .520$. There also was a main effect of block, $F(7, 210) = 6.95$, $MSE = 3,744.64$, $p < .001$, $\eta_p^2 = .188$. This time, there was a Contingency \times Block interaction, $F(6.74, 202.15) = 2.14$, $MSE = 4,180.41$, $p = .043$, $\eta_p^2 = .067$. The source of the interaction likely was the “bumpiness” over blocks in the low-contingency condition, which is not readily interpretable. It is worth noting that, in every condition, there was a main effect of block accompanied by a linear trend across blocks; we report just the linear trend analyses here: for high-contingency stimuli, $F(1, 30) = 14.05$, $MSE = 1,373.84$, $p = .001$, $\eta_p^2 = .319$; for low-contingency stimuli, $F(1, 30) = 17.53$, $MSE = 3,461.09$, $p < .001$, $\eta_p^2 = .369$; and for baseline stimuli, $F(1, 30) = 12.41$, $MSE = 2,447.78$, $p = .001$, $\eta_p^2 = .293$. As we have seen in most of the contingency learning experiments carried out in our laboratory, participants characteristically slow down across blocks, perhaps tiring of the task.

As in Experiment 1a, the two key comparisons were carried out. There was a benefit, such that participants were 26 ms faster responding to high-contingency stimuli ($M = 526$) than to no-contingency baseline stimuli ($M = 552$), $t(30) = 6.57$, $SE = 3.93$, $p < .001$. And again low-contingency stimuli ($M = 568$) were slower—by 16 ms—than were no-contingency baseline stimuli, $t(30) = 2.86$, $SE = 5.58$, $p = .008$, entirely in keeping with the cost observed in Experiment 1a.

Discussion

Participants responded faster to high-contingency stimuli relative to baseline no-contingency stimuli, reflecting a benefit of high contingency. This is consistent with the findings of Schmidt and Besner (2008). But, in contrast to their findings—and consistent with Schmidt’s (2013; Schmidt et al., 2016) PEP model—the low-contingency stimuli also diverged from the baseline no-contingency stimuli, indicative of a cost. Moreover, the use of nonwords in Experiment 1b indicates that familiar stimuli such as words are not necessary for this type of contingency learning to occur. People can learn contingencies with nonmeaningful units, and this learning is as quick and stable as it is with meaningful units, displaying all of the same pattern elements. What is being learned about the irrelevant dimension, in this case the words or nonwords, is more primitive, beneath the level of meaning. This is a key reason why we see color–word contingency learning as tapping directly into the learning of simple associations.

Experiment 2

In Experiment 2, we set out to determine what happens to the cost versus the benefit when contingencies are turned on and off, mindful of the extinction and (re)acquisition functions so familiar in conditioning studies. To do so, we examined how the learning in an initial contingency phase affected performance when there

was a switch to a no-contingency phase, in which all stimuli appeared with equal probability, and then back to the same contingency phase. Unlike in prior studies in the literature, the presence of a no-contingency baseline item through all phases provided a bridge between the phases—and indeed between the experiments: We expected the two contingency phases here to replicate the patterns observed in Experiment 1.

Based on the research reported in Schmidt et al. (2010, Experiment 1), we also expected that, when the contingencies were canceled, any brief initial carryover of the contingency effect would quickly disappear. In their experiment, the carryover was gone by the end of the first no-contingency block. Of course, because their experiment did not include a no-contingency baseline, they could not differentiate benefit and cost, nor could they observe how these behaved as a function of transfer from contingency to no-contingency and back again. The PEP model does not predict a differential pattern for the benefit versus the cost.

Method

Participants. Thirty participants ($M_{age} = 20.03$ years, excluding one participant, who did not report their age; 26 female) from the University of Waterloo completed one session in exchange for course credit. All had normal or corrected-to-normal vision.

Design and procedure. First, there were 48 practice trials, during which participants were shown a row of asterisks (*****) in red, yellow, or green, with equal probability. Following practice, there were three main phases: first contingency phase, no-contingency phase, and second contingency phase. There were six blocks of 48 trials in each of these phases. All stimuli were presented in 18-point Courier New font on a black background.

Contingency phases. These were carried out exactly as in Experiment 1a except that there were only six blocks, not eight.

No contingency phase. All four words appeared in each of the three colors equally often (33.33%). All combinations of colors and words were counterbalanced.

Results

Errors and outliers. Overall, there were 4.76% errors. For the first contingency phase, the difference in errors between high-contingency stimuli ($M = 3.50\%$) and baseline stimuli ($M = 4.40\%$) was not significant, $t(29) = 1.49$, $SE = .01$, $p = .147$, nor was the difference between low-contingency stimuli ($M = 5.65\%$) and baseline stimuli, $t(29) = 1.72$, $SE = .01$, $p = .10$, although each effect was in the expected direction. Therefore neither benefit nor cost reached significance in the error data. For the no-contingency phase, as expected, there were no reliable differences in errors over the three conditions (high contingency: 4.63%; low contingency: 5.70%; baseline: 5.74%): high contingency versus baseline, $t(29) = 1.86$, $SE = .004$, $p = .073$, and low contingency versus baseline, $t(29) = .08$, $SE = .01$, $p = .940$. For the second contingency phase, high-contingency trials ($M = 4.22\%$) had more errors than did baseline trials ($M = 1.40\%$), $t(29) = 7.60$, $SE = .004$, $p < .001$, an unexpected cost for the high-contingency stimuli. Low-contingency trials ($M = 6.48\%$) also had more errors than did baseline trials, $t(29) = 6.31$, $SE = .01$, $p < .001$, representing the anticipated cost in errors.

Response times. First, incorrect responses were removed from the analysis of RTs. Next, all response times less than 200 ms were removed as anticipations. Finally, RTs were trimmed using the 2.5-SD criterion by subject and type of contingency. For the first contingency phase, trimming resulted in removing 3.07% of high-contingency trials, 2.65% of low-contingency trials, and 2.90% of baseline trials; for the no-contingency phase, trimming resulted in removal of 3.01% of high-contingency trials, 3.10% of low-contingency trials, and 3.54% of baseline trials; and for the second contingency phase, trimming resulted in removal of 3.17% of high-contingency trials, 2.68% of low-contingency trials, and 2.80% of baseline trials.

Figure 2 presents the RT data as a function of condition, block, and phase. To examine the contingency effect, we conducted

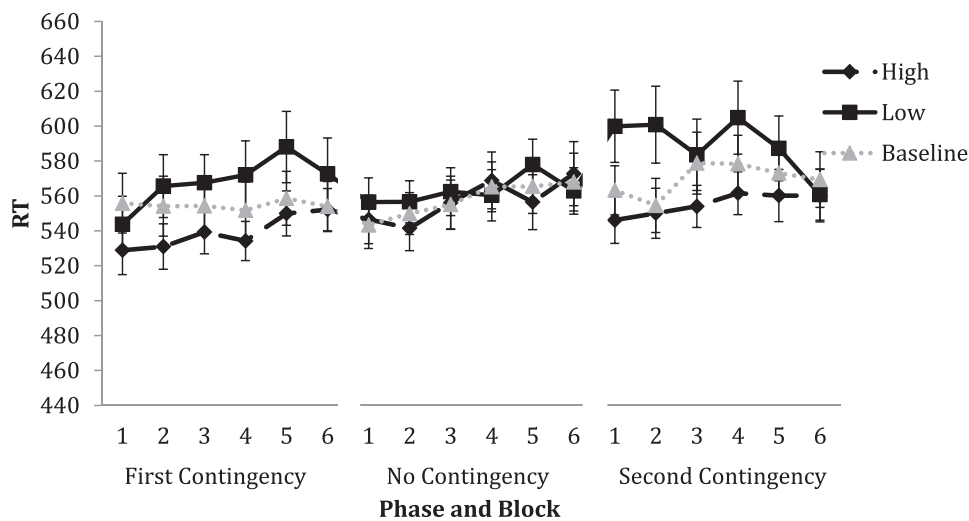


Figure 2. Experiment 2: Mean response times (RTs) for the high-contingency, low-contingency, and no-contingency (baseline) stimuli as a function of block. Data are shown separately for three phases: first contingency phase, no-contingency phase, and second contingency phase. Error bars depict one standard error of the corresponding mean.

separate 3 (contingencies) \times 6 (blocks) repeated-measures ANOVAs for each of the three phases.

First contingency phase. There was a significant main effect of contingency, $F(1.66, 48.09) = 10.44$, $MSE = 4,387.50$, $p < .001$, $\eta_p^2 = .265$. There was no effect of block, $F(5, 145) = 1.41$, $MSE = 3,904.51$, $p = .224$, $\eta_p^2 = .046$, and there was no interaction between contingency and block, $F(5.42, 157.19) = .93$, $MSE = 3,746.44$, $p = .470$, $\eta_p^2 = .031$. Basically, the contingency effect was consistent across blocks. High-contingency stimuli ($M = 539$ ms) were responded to 15 ms faster than were baseline stimuli ($M = 554$ ms), $t(29) = 3.18$, $SE = 4.76$, $p = .003$, evidence of a significant benefit. Low-contingency stimuli ($M = 568$ ms) were responded to 14 ms slower than were baseline stimuli, $t(29) = 2.16$, $SE = 6.20$, $p = .039$, evidence of a significant cost.

No-contingency phase. There was no main effect of contingency ($F < 1$) and no interaction between contingency and block, $F(10, 290) = 1.06$, $MSE = 1,253.32$, $p = .397$, $\eta_p^2 = .035$. Essentially, the contingency effect disappeared right away once the contingencies were canceled. There was a main effect of block, $F(5, 145) = 2.29$, $MSE = 2,908.60$, $p = .05$, $\eta_p^2 = .073$, and a linear trend over blocks, $F(1, 29) = 8.11$, $MSE = 3,772.73$, $p = .008$, $\eta_p^2 = .219$, evidence of a reliable slowing over blocks. Unlike for Schmidt et al. (2010), there was no evidence of any transfer of the contingencies learned in the first contingency phase to the no-contingency phase, even in the first block after the contingencies were canceled.

Second contingency phase. The contingencies learned in the first contingency phase did reappear immediately, though, when they were reinstated in the second contingency phase. The main effect of contingency was significant, $F(1.48, 42.88) = 16.49$, $MSE = 4,348.40$, $p < .001$, $\eta_p^2 = .363$. There was no main effect of block ($F < 1$), nor an interaction of contingency with block, $F(5.74, 166.47) = 1.77$, $MSE = 4,220.48$, $p = .112$, $\eta_p^2 = .057$. Again, the contingency effect was consistent over blocks. High-contingency stimuli ($M = 555$ ms) were 14 ms faster than were baseline stimuli ($M = 569$ ms), $t(29) = 3.01$, $SE = 4.57$, $p = .005$, a reliable benefit. Low-contingency stimuli ($M = 590$ ms) were 21 ms slower than were baseline stimuli, $t(29) = 4.04$, $SE = 5.11$, $p < .001$, a reliable cost.

Discussion

As in Experiment 1, the contingency effect was found to include both a benefit and a cost. Both effects were apparent in both contingency phases. The moment the contingencies were canceled, the contingency effect disappeared; the moment the contingencies were restored, the contingency effect reappeared. It is tempting to view the reemergence of the effect as evidence of the restoration of the original effect—a kind of savings—but given the rapid learning of these contingencies, it could also simply be new learning. In any event, the contingency effect is very responsive to change.

It is interesting that there was no evidence of any carryover effect in the no-contingency phase: The contingency effect simply disappeared immediately in the first block of transfer. This certainly highlights the idea that one can unlearn contingencies very quickly (Schmidt et al., 2010), with the added finding that this is true in terms of both the cost and the benefit. The PEP model does not predict differential impacts of canceling or of restoring con-

tingencies on benefit versus cost, so our results are consistent with the model.

General Discussion

In three experiments in which a baseline no-contingency condition was inserted among high-contingency and low-contingency conditions, in addition to the benefit for high-contingency stimuli, there also consistently was a reliable cost for low-contingency stimuli. This cost was evident both in response time and in error rate and was true both for words and for nonwords. Our experiments, closely modeled after the most prevalent procedure used to study contingency learning (see Schmidt et al., 2007), also reproduced several of the other patterns evident in that prior work, including rapid learning and unlearning (Schmidt et al., 2010).¹ We conclude that the typical contingency effect, the difference between high-contingency and low-contingency stimuli, derives from the sum of two components—faster responding to high-contingency stimuli coupled with slower responding to low-contingency stimuli.

While we were carrying out these studies, a related study was being conducted by Schmidt and De Houwer (2016, Experiment 1). They, too, observed both benefit and cost when they compared high- and low-contingency color–word pairings to one word (of three) that was equally likely to appear in each color, the condition that we have called “no contingency.” Their emphasis, however, was on comparing high and low contingency to entirely novel color–word pairings: They showed that both high- and low-contingency pairings benefited relative to that alternative baseline. Although that is a useful perspective, we see comparison to the novel baseline more as providing an index of familiarity, whereas comparison to the no-contingency baseline (which they called “medium contingency”) more directly addresses contingency. In

¹ A potentially interesting observation in the present experiments is worthy of subsequent pursuit. The low-contingency stimuli and the baseline stimuli seemed to diverge somewhat slowly over the first couple of blocks, unlike the virtually immediate separation of high-contingency trials from both low-contingency and no-contingency trials. If real, this possibly slow onset of the cost in performance on low-contingency stimuli could be due to the small number of presentations of these stimuli in each block (only 12.5% of all trials). This would have made it hard to learn the low-contingency stimuli—and to differentiate them from the baseline trials, which were also relatively infrequent (only 25% of all trials) compared to high-contingency stimuli (62.5% of all trials). Moreover, the low-contingency stimuli may have been increasingly surprising as trials progressed (in line with Hon & Tan, 2013). Consistent with this speculation, unlike in the first contingency phase of Experiment 2, in the second contingency phase this diverging pattern was not apparent, presumably because the first contingency phase provided sufficient experience.

There is, of course, one issue to be cautious about in any situation involving learning—the learning–performance distinction (Tolman & Honzik, 1930; see Bjork, 1999, for discussion). What we saw in the data was performance differences between high-, low-, and no-contingency trials. From these differences, we inferred that learning was occurring, but participants may not in fact have been learning high-contingency items as early as the data suggest. High-contingency trials may have reaped an immediate benefit due to the confounded feature of having more repeat trials, not due to their high contingency. It might have taken participants a few blocks to actually learn the high-contingency items, which could explain why the low-contingency items did not appear to register a cost at first. After a few blocks, when most participants had learned the high-contingency items, they were then surprised by the now unexpected low-contingency items.

our experiments, we have shown additionally that both the cost and the benefit are quickly established, that both are consistent across extended performance, and that both are rapidly extinguished and restored.

The cost-benefit pattern is entirely consistent with the prediction of the PEP model (Schmidt, 2013; Schmidt et al., 2016). Essentially, each new trial routinely brought about retrieval of prior relevant trial episodes stored in memory, where relevance was determined by the response-relevant dimension (color), the response-irrelevant dimension (word), and the response itself (key press). High-contingency trials retrieved prior episodes that were more likely to match, including the recent ones, and many of these were responded to rapidly, so high-contingency trials benefited. Low-contingency trials matched very few of the retrieved prior episodes, and these were typically responded to slowly, plus they were victims of partial overlap with high-contingency episodes; together these resulted in a cost. As a result, previous episodes were each instances of associations, providing what amounted to an associative history of performance in the task, a history which in turn affected performance of the task.

There are many directions that research on the acquisition of associations could take in the context of the contingency learning task. This paradigm readily allows for the exploration of associative strength by manipulating proportion contingency. It also could be used to assess the role of attention in associative learning by following an acquisition phase with an attentional measure such as a cueing task: Would high-contingency stimuli command more attention than would low-contingency stimuli? And it is well suited to the examination of memory for associations, for example by repeatedly changing the irrelevant dimension to create a set of well-learned versus poorly learned associations that could then be tested via explicit or implicit memory tests. Overall, we see the color-word contingency learning paradigm as a rich testing ground for the investigation of how one forms and uses associations.

References

- Bjork, R. A. (1999). Assessing your own competence: Heuristics and illusions. In D. Gopher & A. Koriati (Eds.), *Attention and Performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 435–459). Cambridge, MA: MIT Press.
- Calkins, M. W. (1894). Association [Part D of Studies from the Harvard Psychology Laboratory]. *Psychology Review*, 1, 476–483.
- 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>
- De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and theories on human contingency learning. *Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology*, 55, 289–310. <http://dx.doi.org/10.1080/02724990244000034>
- Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology*. New York, NY: Teacher's College, Columbia University. (Original work published 1885)
- Gehring, W. J., Gratton, G., Coles, M. G., & Donchin, E. (1992). Probability effects on stimulus evaluation and response processes. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 198–216. <http://dx.doi.org/10.1037/0096-1523.18.1.198>
- Hazeltine, E., & Mordkoff, J. T. (2014). Resolved but not forgotten: Stroop conflict dredges up the past. *Frontiers in Psychology*, 5, 1327. <http://dx.doi.org/10.3389/fpsyg.2014.01327>
- 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>
- Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin*, 96, 29–44. <http://dx.doi.org/10.1037/0033-2909.96.1.29>
- Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *American Journal of Psychology*, 77, 576–588. <http://dx.doi.org/10.2307/1420768>
- Laberge, D., & Tweedy, J. R. (1964). Presentation probability and choice time. *Journal of Experimental Psychology*, 68, 477–481. <http://dx.doi.org/10.1037/h0048691>
- Lewicki, P. (1985). Nonconscious biasing effects of single instances on subsequent judgments. *Journal of Personality and Social Psychology*, 48, 563–574. <http://dx.doi.org/10.1037/0022-3514.48.3.563>
- 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>
- 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>
- 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>
- Musen, G., & Squire, L. R. (1993). Implicit learning of color-word associations using a Stroop paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 789–798. <http://dx.doi.org/10.1037/0278-7393.19.4.789>
- Nattkemper, D., & Prinz, W. (1997). Stimulus and response anticipation in a serial reaction task. *Psychological Research*, 60, 98–112. <http://dx.doi.org/10.1007/BF00419683>
- Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, 88, 93–134. <http://dx.doi.org/10.1037/0033-295X.88.2.93>
- 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., & Besner, D. (2008). The Stroop effect: Why proportion congruent has nothing to do with congruency and everything to do with contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 514–523. <http://dx.doi.org/10.1037/0278-7393.34.3.514>
- 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). 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. (2012b). 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. (2012c). 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. (2016). Contingency learning tracks with stimulus-response proportion. *Experimental Psychology*, *63*, 79–88. <http://dx.doi.org/10.1027/1618-3169/a000313>
- 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>
- Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools, Inc.
- Shanks, D. R. (2010). Learning: From association to cognition. *Annual Review of Psychology*, *61*, 273–301. <http://dx.doi.org/10.1146/annurev.psych.093008.100519>
- 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>
- Tolman, E. C., & Honzik, C. H. (1930). Introduction and removal of reward, and maze performance in rats. *University of California Publications in Psychology*, *4*, 257–275.
- Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *47*, 631–650. <http://dx.doi.org/10.1080/14640749408401131>
- Wolfe, J. M., Horowitz, T. S., Van Wert, M. J., Kenner, N. M., Place, S. S., & Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental Psychology: General*, *136*, 623–638. <http://dx.doi.org/10.1037/0096-3445.136.4.623>

Received July 6, 2016

Revision received April 15, 2017

Accepted April 26, 2017 ■

Call for Nominations

The Publications and Communications (P&C) Board of the American Psychological Association has opened nominations for the editorships of the *Journal of Experimental Psychology: Animal Learning and Cognition*, *Neuropsychology*, and *Psychological Methods* for the years 2020 to 2025. Ralph R. Miller, PhD, Gregory G. Brown, PhD, and Lisa L. Harlow, PhD, respectively, are the incumbent editors.

Candidates should be members of APA and should be available to start receiving manuscripts in early 2019 to prepare for issues published in 2020. Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. Self-nominations are also encouraged.

Search chairs have been appointed as follows:

- *Journal of Experimental Psychology: Animal Learning and Cognition*, Chair: Stevan E. Hobfoll, PhD
- *Neuropsychology*, Chair: Stephen M. Rao, PhD
- *Psychological Methods*, Chair: Mark B. Sobell, PhD

Candidates should be nominated by accessing APA's EditorQuest site on the Web. Using your browser, go to <https://editorquest.apa.org>. On the Home menu on the left, find "Guests/Supporters." Next, click on the link "Submit a Nomination," enter your nominee's information, and click "Submit."

Prepared statements of one page or less in support of a nominee can also be submitted by e-mail to Sarah Wiederkehr, P&C Board Editor Search Liaison, at swiederkehr@apa.org.

Deadline for accepting nominations is Monday, January 8, 2018, after which phase one vetting will begin.