

Contingency proportion systematically influences contingency learning

Noah D. Forrin¹ · Colin M. MacLeod¹

Published online: 3 October 2017
© The Psychonomic Society, Inc. 2017

Abstract In the color-word contingency learning paradigm, each word appears more often in one color (high contingency) than in the other colors (low contingency). Shortly after beginning the task, color identification responses become faster on the high-contingency trials than on the low-contingency trials—the contingency learning effect. Across five groups, we varied the high-contingency proportion in 10% steps, from 80% to 40%. The size of the contingency learning effect was positively related to high-contingency proportion, with the effect disappearing when high contingency was reduced to 40%. At the two highest contingency proportions, the magnitude of the effect increased over trials, the pattern suggesting that there was an increasing cost for the low-contingency trials rather than an increasing benefit for the high-contingency trials. Overall, the results fit a modified version of Schmidt’s (2013, *Acta Psychologica*, 142, 119–126) parallel episodic processing account in which prior trial instances are routinely retrieved from memory and influence current trial performance.

Keywords Attention in learning · Attention and memory · Perceptual implicit memory

People learn a multitude of associations that help them to navigate their daily lives. For example, we learn to associate the four seasons with different weather patterns (and learn new associations when traveling abroad). We learn that a red traffic light is ordinarily followed by a green light, although occasionally (and only at certain intersections) an advanced left-turn signal appears

before the light turns green for through traffic. Learning such associations facilitates quick and/or accurate responding. In the case of waiting at a red light, we anticipate the green light to follow and then are ready to press the accelerator when the light turns green. However, learning the more common association may impair responding when the less common event occurs. For example, despite intending to go straight, we may momentarily take our foot off the brake or even begin pressing the accelerator when the advance left-turn signal appears because we are anticipating the light turning green.

Thus, experience suggests that learned associations between stimuli not only may facilitate responding to a stimulus that has a high probability of being paired with another stimulus but may also impair responding to a stimulus that has a low probability of being paired with another stimulus. In the present research, we investigated these possibilities in the context of contingency learning (see De Houwer & Beckers, 2002), using the paradigm developed by Schmidt, Crump, Cheesman, and Besner (2007; see also Musen & Squire, 1993). In their standard contingency learning task, each trial consists of one of three words (*plate*, *month*, *under*) presented in one of three colors (red, yellow, green). Participants respond by identifying the color, typically by pressing a key that corresponds to the color (but see Forrin & MacLeod, 2017, for vocal responding). Critically, each word is presented with high probability in one color and with low probability in each of the other two colors. For example, *plate* might have an 80% chance of appearing in red and a 10% chance of appearing in each of yellow and green. In this example, *plate*_{red} is referred to as a high-contingency pairing, and *plate*_{yellow} and *plate*_{green} are referred to as low-contingency pairings.

The contingency learning effect refers to the finding that response times tend to be faster and more accurate for high-contingency pairings than for low-contingency pairings—despite the fact that the word information actually is not required

✉ Noah D. Forrin
nforrin@uwaterloo.ca

¹ Department of Psychology, University of Waterloo,
Waterloo, Ontario N2M 2R8, Canada

to make the correct response. The effect can occur when participants are unaware of the high-contingency pairs (Schmidt et al., 2007), which suggests that the learning can be implicit, although the effect is amplified when participants are made aware of the contingencies (Schmidt & De Houwer, 2012b) or are given the explicit goal of learning them (Schmidt & De Houwer, 2012a). Schmidt et al. (2007, Experiment 4) provided evidence consistent with the idea that the contingency learning effect is driven by participants learning the associations between the irrelevant word information and the response (i.e., the keypress).

Schmidt (2013; see also Schmidt, De Houwer, & Rothermund, 2016) has formulated a Parallel Episodic Processing (PEP) model to explain the contingency learning effect (inspired, in part, by Logan's, 1988, instance model of automaticity). According to the PEP model, participants encode each trial of the task into episodic memory and then retrieve previous episodes (in parallel) on subsequent trials. Critically, the activation for a given response is weighted by the "proportion of episodes pointing to that response" (Schmidt & De Houwer, 2016a, p. 87). For example, if *plate* had appeared in red on 80% of previous trials, in yellow on 10%, and in green on 10%, the response activation for a "red" response to *plate* would be higher than for either a "yellow" or "green" response, resulting in faster response times for $\text{plate}_{\text{red}}$ and slower responses for $\text{plate}_{\text{yellow}}$ or $\text{plate}_{\text{green}}$. The contingency learning effect for errors can likewise be explained by the PEP framework: In this example, participants would be more likely to accidentally respond "red" on a $\text{plate}_{\text{green}}$ (low-contingency) trial than to respond "green" on a $\text{plate}_{\text{red}}$ (high-contingency) trial because in the first case the response activation for the "red" response is relatively high, whereas in the second case the activation for the "green" response is relatively low.

The PEP model therefore yields the prediction that the higher the probability of occurrence of a particular color-word contingency, the faster and more accurate participants' responses should be to that contingency. Schmidt and De Houwer (2016a) tested this claim (which they termed the "pure proportion" account) across two experiments. In Experiment 1, participants responded to high-contingency, medium-contingency, and low-contingency trials. Consistent with their account, response times followed a high < medium < low pattern of results, and error rates mirrored this pattern (cf. Schmidt & Besner, 2008; Lin & MacLeod, 2017). In Experiment 2, participants responded to high-contingency and low-contingency trials in a learning phase, and then several novel words—presented only once each—were added to a test phase (constituting what they called no-contingency, "neutral" trials). Both response times and error rates followed a high < low < neutral pattern of results, again consistent with the pure proportion account.

The primary aim of the present experiment was to comprehensively test the pure proportion account by varying the

probability of high-contingency and low-contingency trials in the contingency learning task across a wide range of probabilities. We used what has become the standard contingency learning paradigm, which features three words and three colors (see, e.g., Schmidt & De Houwer, 2012a, 2012b, 2016a, 2016b). A robust sample participated in one of five different groups ($n = 50$ per group). In the highest high-contingency probability group, there was an 80% chance of one of the three high-contingency pairs occurring on any given trial (and a 20% chance of one of the six low-contingency pairs occurring). The other four groups were 70%, 60%, 50%, and 40% high-contingency. In the 40% high-contingency group, although low-contingency pairs occurred more commonly than high-contingency pairs overall, each word still had a color that it was paired with more often (40% of the time) than it was paired with either of the other two colors (30%, 30%). The wide range of proportions used here also presented an opportunity to explore the *threshold* for contingency learning. To our knowledge, our experiment is the first to use less than 60% high-contingency trials (50% and 40%). These lower proportions of high-contingency trials allowed us to estimate the proportion at which the contingency learning effect would no longer express itself.

The pure proportion account generates the straightforward prediction that the contingency learning effect ought to increase as the proportion of high-contingency trials increases (and as the proportion of low-contingency trials, correspondingly, decreases) due to progressively improving high-contingency trial performance *and* progressively worsening low-contingency trial performance. As the proportion of high-contingency trials increases, high-contingency trial performance ought to improve because the proportion of prior episodes retrieved on a given trial that point to a high-contingency response would correspondingly increase, thereby increasing the response activation for a high-contingency response. By the same logic, low-contingency trial performance ought to worsen with increasing high-contingency trial proportion because there would be a progressively smaller proportion of prior episodes that point to low-contingency responses, decreasing the response activation for low-contingency responses.

Thus, the pure proportion account suggests that there will be a performance trade-off between high-contingency and low-contingency trials simply because as the proportion of high-contingency trials increases, the proportion of low-contingency trials necessarily decreases. Increases in response activation for high-contingency responses therefore are accompanied by decreases in response activation for low-contingency responses.

The present experiment is the first to parametrically vary high-contingency proportion to map out a function for the contingency learning effect across "proportion space." As already noted, prior research that varied the proportion of the contingencies has focused on comparing performance on high-contingency trials and low-contingency trials to a

baseline condition in a within-subjects design, to assess performance costs and benefits (Lin & MacLeod, 2017; Schmidt & Besner, 2008; Schmidt & De Houwer, 2016a). In contrast to those experiments, we did not compare high-contingency and low-contingency performance to a baseline condition, and we manipulated the proportion of high-contingency/low-contingency trials between subjects, eliminating the potential intrusion of strategy changes across different high-contingency proportions. Although Schmidt et al. (2007) found a larger contingency learning effect with 80% high-contingency trials (Experiment 1a) than with 60% high-contingency trials (Experiment 1b), they did not compare the magnitude of these effects (60 ms vs. 27 ms) statistically.¹

The second aim of the present work was to investigate the relation between contingency awareness and performance. Schmidt and De Houwer (2012b) demonstrated that participants who were aware of the high-contingency pairs—indexed via a postexperiment questionnaire—showed larger contingency learning effects than those who were unaware, and argued that awareness “benefited performance by leading participants to attend more to the predictive dimension (i.e., the word)” (p. 1765). The increased attention on words may speed high-contingency response times because individuals respond to words more rapidly than to colors (see the relative speed of processing account of Morton & Chambers, 1973, as applied to the Stroop task), and the predictive word may therefore frequently give participants a “head-start” in responding to the color on the high-contingency trials. In terms of Schmidt et al.’s (2016) PEP model, contingency awareness may increase the response activation for high-contingency pairs without correspondingly decreasing the response activation for low-contingency pairs. The present experiment provided a robust data set that allowed us to test Schmidt and De Houwer’s account that the contingency awareness *benefits* performance (presumably without imposing a concurrent cost on low-contingency trials).

Our third aim was to explore the pattern of contingency learning with practice. Does the contingency learning effect change as learning progresses? The PEP model predicts that the contingency learning effect will increase as the imbalance in memory favoring high-contingency instances over low-contingency instances increases with experience. But prior research, typically using 80% high-contingency trials, has demonstrated that, after emerging early in the task (after as few as 18 trials; Schmidt, De Houwer, & Besner, 2010), the contingency learning effect is stable across the rest of the trials

(e.g., Schmidt & De Houwer, 2012a, 2012b, 2016a; Schmidt et al., 2007; Schmidt et al., 2010). Because the change with experience may, however, be relatively small—after all, the contingency learning effect itself is modest—it is possible that previous studies did not have sufficient power to observe this change. If so, our large sample size should remedy this.

Method

Participants

Two hundred and fifty University of Waterloo undergraduate students participated in exchange for course credit, with 50 assigned to each of the five high-contingency proportion groups.

Apparatus

The experiment was carried out using E-Prime software. Participants responded using a QWERTY keyboard by pressing “J” for red, “K” for yellow, and “L” for green. Circular colored tabs of the corresponding color were pasted on those three keys with the assignment of colors to keys consistent across participants.

Materials and design

On each of the 300 trials, one of three five-letter words (*month*, *under*, *plate*) was individually presented in one of the three different print colors (red, yellow, green), on a black background in bold, 18 pt. Courier New font. Each of the three words was presented most frequently in one of the three colors. For example, *month* might have been presented 80% of the time in red and 10% of the time in each of yellow and green. In this case, $month_{red}$ was a high-contingency pairing and $month_{yellow}$ and $month_{green}$ were low-contingency pairings. The three color-word high-contingency pairings were randomly determined for each participant.

The five groups were identical, except for the probability of a high-contingency pairing occurring on each trial (.80, .70, .60, .50, and .40). For example, each trial in the .80 probability condition had an 80% chance of being a high-contingency trial and a 20% chance of being a low-contingency trial (in keeping with the methodology of previous contingency learning research; see, e.g., Schmidt & De Houwer, 2012a, Schmidt & De Houwer, 2012b). Trials therefore were selected at random with replacement (i.e., a trial was not removed from the list of possible future trials after it was presented). Consequently, the overall proportions of high-contingency trials presented to each participant were normally distributed around the five expected means corresponding to each condition of the experiment: (1) $M = .802$, $SD = 0.022$; (2) $M = .698$,

¹ In a similar task that involved identifying target letters that had high-contingency or low-contingency flankers, Miller (1987) found that participants were faster (and more accurate) at identifying targets with high-contingency flankers than those with low-contingency flankers. Notably, this difference was significantly larger when there were 92% high-contingency flankers compared to 56% high-contingency flankers, a result that is also consistent with the “pure proportion account” (Schmidt & De Houwer, 2016a).

$SD = 0.024$; (3) $M = .605$, $SD = 0.024$; (4) $M = .498$, $SD = 0.026$; (5) $M = .397$, $SD = 0.026$.

Procedure

Participants were instructed that, for each trial, they would see a word presented in color in the middle of the screen. They were told to respond to the color of the word as quickly and as accurately as possible by pressing the “J” (red), “K” (yellow), or “L” (green) key. The experimenter informed participants that colored tabs were placed on these keys to assist their responding and that they should respond using their dominant hand (index, middle, and ring fingers). Participants were not informed about the high-contingency and low-contingency pairings.

On each trial, participants first saw a white fixation “+” for 150 ms. A blank screen then appeared for 150 ms followed by a word in color (the target). Participants had 2,000 ms to respond to the target. After a correct predeadline response, the next trial started immediately. After an incorrect response or a 2,000-ms timeout, “XXX” was presented in white for 500 ms before the next trial.

Once the 300 trials were completed, participants were informed that each word had been presented most often in one color and were asked whether they had noticed these color-word relations (the wording of this question was identical to that found in Schmidt & De Houwer, 2012a, 2012b, translated from Dutch to English): “In this experiment, each word was presented most often in a certain color. Specifically, one word was presented most often in red, one word was presented most often in yellow, and one word was presented most often in green. Did you notice these relationships?” After the subjective awareness question, participants were given three forced-choice questions that assessed their objective awareness of each of the high-contingency color-word pairings. For each question, participants were asked: “In what color was [month/under/plate] usually presented?” (Only one word was presented for each question, and the order of the words was random). Participants responded using the same keys (with tabs) as they had used throughout the experiment. Participants were coded as being objectively aware of the contingencies when they correctly indicated the color in which each of the three words had appeared most frequently; if they made at least one incorrect response, they were coded as objectively unaware.

Results

Data trimming

First, we removed responses that were timeouts (0.10% of high-contingency trials; 0.13% of low-contingency trials). Anticipatory response times less than 200 ms were also removed (0.01% of high-contingency trials; 0.01% of low-

contingency trials). We had decided in advance to exclude the data of any participant who had response times or error rates that were extreme outliers (more than 3.0 SD s above or below the mean response time/error rate). There were two participants with extremely high error rates and one participant with extremely slow response times. Thus, the data of 247 participants are included in the analyses below. Including these three outlier participants did not influence the statistical significance of any of the results.

Relation between high-contingency proportion and the magnitude of the contingency learning effect

As described in the Method section, participants were assigned to each of five between-subject conditions that differed according to the probability of a high-contingency stimulus occurring on a given trial (.80, .70, .60, .50, .40). This design resulted in the participants being presented with a wide range of proportions of high-contingency trials (range: .86–.34), which was well-suited for the regression analyses reported here.

Response times The response times of only correct responses were analyzed. All trials on which errors occurred (3.31% of responses) were removed. Mean error proportions were 2.97% for high contingency and 4.17% for low contingency. Of main interest, and consistent with our hypothesis, a linear regression revealed that the size of the contingency learning effect was significantly positively related to the proportion of high-contingency trials, $R^2 = .25$, $F(1, 245) = 79.86$, $p < .001$. The linear function estimates the threshold for the contingency learning effect in response times to be approximately 41% high-contingency items (see Fig. 1). That is, below 41% high contingency, contingency learning does not occur, insofar as there no longer is a response-time performance advantage for high-contingency (vs. low-contingency) trials. It is unlikely, however, that the true relation between high-contingency proportion and the contingency learning effect is linear (otherwise there would be negative contingency learning effects below a high-contingency proportion of 41%, which seems implausible, theoretically). Indeed, a quadratic relation, $R^2 = .25$, $F(2, 244) = 41.28$, $p < .001$, fit the data as well as the linear function did, and yielded a similar estimate of the threshold of the contingency learning effect (approximately 39% high-contingency items; see Fig. 1). The significant quadratic function reflects the fact that the magnitude of the contingency learning effect “accelerated” as high-contingency proportion increased. Thus, the relation appears to be exponential in nature, likely reaching asymptote at zero.

Errors As anticipated, high-contingency proportion was also significantly positively related to the contingency learning effect for errors, $R^2 = .12$, $F(1, 245) = 34.89$, $p < .001$. The linear function estimated the threshold for the contingency learning

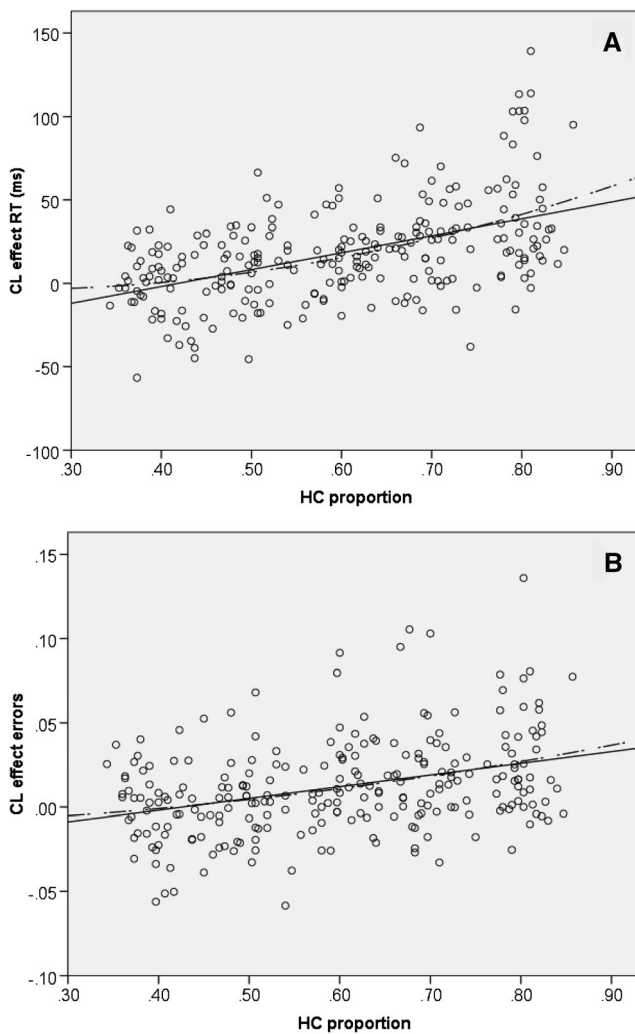


Fig. 1 Contingency learning (CL) effect as a function of the proportion of high-contingency (HC) trials (**a**) in response time (in ms), and (**b**) in error rate. Mean contingency learning (CL) effect is shown for each participant. Also shown is an estimation of an overall linear function (solid straight line) and quadratic function (dotted curving line)

effect in error rates to be approximately 42% high-contingency items (see Fig. 2), consistent with the estimate based on response times. Again, a quadratic function, $R^2 = .13$, $F(2, 244) = 17.57$, $p < .001$, fit the error data as well as the linear function did, and gave an identical estimate for the threshold of the contingency learning effect (42% high-contingency items).

Relation between high-contingency proportion and performance on high-contingency and low-contingency trials

The positive relation between the proportion of high-contingency trials and the size of the contingency learning effect could reflect either (or both) of two factors: (1) Higher proportions of high-contingency trials could be associated with faster and less error-prone responding to high-contingency trials, and/or (2) higher proportions of high-

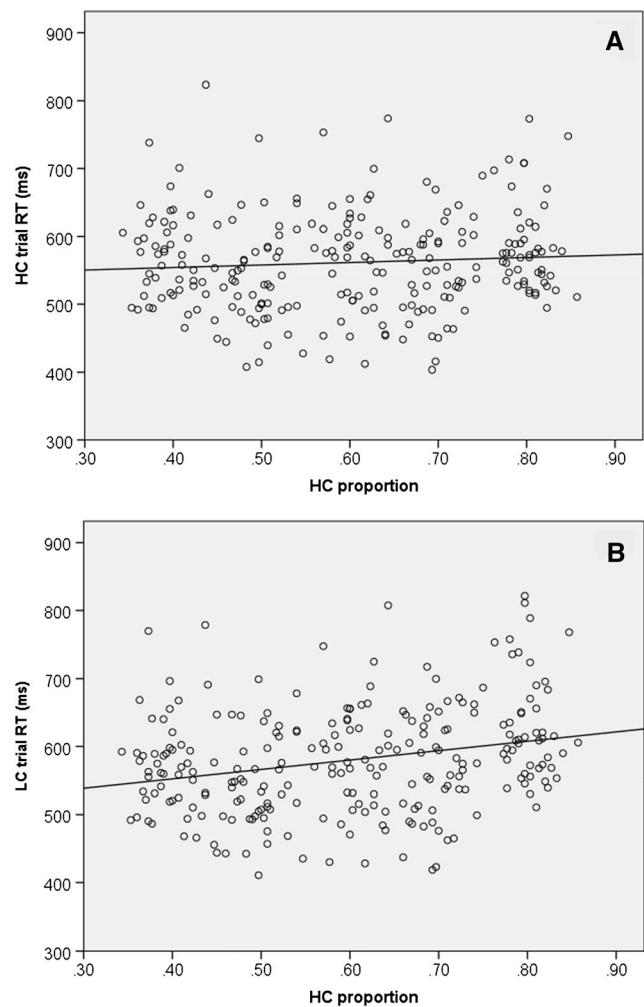


Fig. 2 **a** High-contingency response times (in ms); **b** Low-contingency response times (in ms), each as a function of the proportion of high-contingency (HC) trials. Means are shown for each participant. Also shown is an estimation of an overall linear function (solid straight line) and quadratic function (dotted curving line)

contingency trials could be associated with slower and more error-prone responding to low-contingency trials.

Response times. First, a linear regression revealed that high-contingency proportion was nonsignificantly related to response times on high-contingency trials (see Fig. 2), $R^2 = .01$, $F(1, 245) = 1.38$, $p = .24$. Thus, there was no support for the idea of an increasing benefit of contingency learning as high-contingency proportion increased; indeed, the numerical trend was actually in the opposite direction. In contrast, a second linear regression revealed that high-contingency proportion was significantly positively related to response times on low-contingency trials (see Fig. 2), $R^2 = .07$, $F(1, 245) = 17.10$, $p < .001$. This result was consistent with the possibility that participants' responses to low-contingency trials tended to slow down as low-contingency trials became increasingly rare, suggesting a growing cost with increasing high-contingency proportion.

Errors A linear regression revealed that high-contingency proportion was significantly negatively related to errors on high-contingency trials, $R^2 = .03$, $F(1, 245) = 8.38$, $p = .004$ (see Fig. 3), and was significantly positively related to errors on low-contingency trials, $R^2 = .03$, $F(1, 245) = 7.54$, $p = .006$ (see Fig. 3). These results indicate that, as high-contingency trials became more frequent (and hence low-contingency trials became increasingly rare), participants' responses to high-contingency trials became more accurate whereas their responses to low-contingency trials became less accurate.

The relation between contingency awareness and contingency learning

Table 1 shows the proportion of participants in each group (i.e., .80, .70, .60, .50, and .40) who reported being aware of the high-contingency pairs, in terms of both subjective awareness

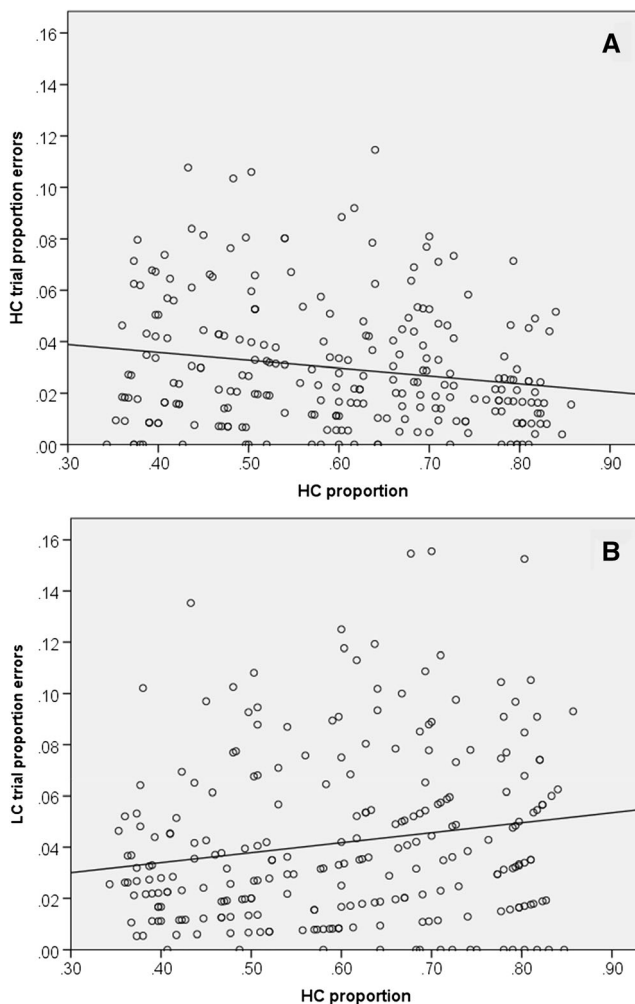


Fig. 3 a high-contingency error rates; b low-contingency error rates, both as function of the proportion of high-contingency (HC) trials. Means are shown for each participant. Also shown is an estimation of an overall linear function

and objective awareness measures.² The empirical pattern is clear—particularly for objective awareness—and not surprising: As the high-contingency proportion increased, so did contingency awareness. Subjective awareness seemed to be a rougher index, in that it only differentiated .80 and .70 from .60, .50, and .40. Objective awareness rates increased monotonically with increasing high-contingency proportion; all of these rates were significantly greater than chance (16.67%) for participants in the .80, .70, .60, and .50 high-contingency proportion groups ($p < .01$), but not for those in the .40 group ($p = .32$).

Response times Both subjective awareness, $r(246) = .20$, $p = .002$, and objective awareness, $r(246) = .36$, $p < .001$, were significantly correlated with the size of the contingency learning effect; the two awareness measures also were significantly intercorrelated, $r(246) = .24$, $p < .001$. To further examine whether contingency awareness was related to the size of the contingency learning effect, independent of high-contingency proportion (a factor with which they were both correlated), both awareness measures were added as predictors to the regression analyses reported in the previous section, for which the proportion of high-contingency trials was the only predictor of the contingency learning effect. With all three predictors included, the regression model was significant, $R^2 = .27$, $F(3, 243) = 29.45$, $p < .001$, and accounted for a significantly higher proportion of the variance in the size of the contingency learning effect relative to the regression model in which high-contingency proportion was the only predictor, $R^2_{\text{change}} = 0.02$, $F(2, 243) = 3.44$, $p = .03$. Both high-contingency proportion and objective awareness were significant predictors of the contingency learning effect, $\beta = .41$, $t(246) = 6.42$, $p < .001$, and $\beta = .15$, $t(246) = 2.37$, $p = .02$, respectively, but subjective awareness was not a significant predictor, $\beta = .05$, $t(246) = 0.81$, $p = .42$.³

Next, we conducted the same regression with high-contingency response times as the dependent measure. The regression model was significant, $R^2 = .03$, $F(3, 243) = 2.69$, $p = .05$, and accounted for a significantly greater proportion of variance in high-contingency response times relative to the regression model in which high-contingency proportion was the only predictor, $R^2_{\text{change}} = 0.03$, $F(2, 243) = 3.33$, $p = .04$. Objective awareness was a significant predictor of the high-contingency response times, $\beta = -0.15$, $t(246) = 2.10$, $p = .04$, whereas neither high-contingency proportion nor subjective awareness was a significant predictor, $\beta = 0.11$, $t(246) = 1.57$, $p = .12$, and $\beta = 0.12$, $t(246) = 1.76$, $p = .08$,

² Overall, 27.9% of participants were both subjectively and objectively aware, 18.6% were subjectively aware but objectively unaware, 19.0% were subjectively unaware but objectively aware, and 34.4% were both subjectively and objectively unaware.

³ The wording of the subjective awareness question, taken from prior contingency learning research (e.g., Schmidt et al., 2007), may have biased participants toward responding “yes,” perhaps attenuating the relation between subjective awareness and the contingency learning effect. A more open-ended subjective awareness measure may be preferable for future research.

respectively. Thus, objectively aware participants tended to respond more quickly to high-contingency trials than did objectively unaware participants.

Last, we conducted the regression with low-contingency response times as the dependent measure. The regression model was significant, $R^2 = .08$, $F(3, 243) = 7.30$, $p < .001$, but did not account for a greater proportion of variance in low-contingency response times relative to the regression in which high-contingency proportion was the only predictor, $R^2_{\text{change}} = 0.02$, $F(2, 243) = 2.30$, $p = .10$. Here, high-contingency proportion was the only significant predictor of response times, $\beta = .26$, $t(246) = 3.64$, $p < .001$. Subjective awareness was a marginally significant predictor, $\beta = 0.13$, $t(246) = 1.93$, $p = .06$ (in the opposite direction than would be expected), and objective awareness was not a significant predictor, $\beta = -0.08$, $t(246) = 1.18$, $p = .24$.

In sum, these regressions demonstrate that the magnitude of the contingency learning effect in response times was significantly related to two predictors: the proportion of high-contingency trials and the participant's objective awareness. In terms of the relation between the proportion of high-contingency trials and the size of the contingency learning effect, the "action" was in participants' responses to low-contingency trials: As the proportion of high-contingency trials increased—and low-contingency trials became increasingly rare—participants responded to low-contingency trials increasing slowly (a robust effect). In terms of the relation between objective awareness and the size of the contingency learning effect, the "action" was in participants' responses to high-contingency trials: Participants who were objectively aware tended to respond more quickly to high-contingency trials compared with those who were objectively unaware (a modest effect).

Errors Both subjective awareness, $r(246) = .13$, $p = .04$, and objective awareness, $r(246) = .22$, $p = .001$, also were significantly correlated with the size of the contingency learning effect in errors. A regression model that included high-contingency proportion, objective awareness, and subjective awareness as predictors of the contingency learning effect was significant, $R^2 = .12$, $F(3, 243) = 11.93$, $p < 0.001$, but accounted for a nonsignificantly higher proportion of the variance, $R^2_{\text{change}} = 0.004$, $F(2, 243) = 0.52$, $p = .60$, relative to the model that only included the high-contingency proportion predictor. High-contingency proportion was a significant predictor of the contingency learning effect, $\beta = .32$, $t(246) = 4.55$, $p < 0.001$, but objective awareness and subjective awareness were not significant predictors, $\beta = .06$, $t(246) = 0.87$, $p = .39$, and $\beta = .03$, $t(246) = 0.42$, $p = .68$, respectively.

When the regression was conducted with high-contingency error rates as the dependent measure, R^2_{change} was nonsignificant ($p = .18$), as was the case with low-contingency error rates as the dependent measure ($p = .80$). Thus, increasing the proportion of high-contingency trials tended to increase the size of the contingency learning effect—both by decreasing

error rates for high-contingency trials and by increasing error rates for low-contingency trials (see Fig. 3)—whereas contingency awareness was unrelated to the contingency learning effect in errors.

Block analyses

Here we tested the PEP model's prediction that the contingency learning effect would increase over blocks of the experiment. Participants' response-time data were divided into six blocks of 50 trials each—the smallest block size for which there were no missing block data for any of the participants. The contingency learning effect was significant in each of the six blocks, and the magnitude of the effect did indeed tend to increase over blocks (see Table 2).

Response times A Contingency (high vs. low) \times Block (one to six) ANOVA revealed the expected main effect of contingency, $F(1, 246) = 95.61$, $MSE = 2805.52$, $p < .001$, $\eta^2 = .28$, and a significant linear contrast of block, $F(1, 246) = 9.74$, $MSE = 6818.64$, $p = .002$, $\eta^2 = .04$, indicating that participants' response times became slower over the experiment (a result that we have quite consistently seen in our other contingency learning studies, presumably due to fatigue). Of central interest, the linear contrast of the Contingency \times Block interaction was also significant, $F(1, 246) = 7.31$, $p = .007$, $\eta^2 = 0.03$, indicating that the size of the contingency learning effect increased linearly across the six blocks of the experiment.⁴ This trend is displayed in Fig. 4.⁵

⁴ Note that assessment of whether this linear increase reflected an increasing benefit on high-contingency trials or an increasing cost on low-contingency trials (or both) was not possible because this experiment did not incorporate a "no contingency" condition as a baseline (for experiments including such a baseline, see Lin & MacLeod, 2017).

⁵ In a supplemental analysis, we added Group (.80, .70, .60, .50, .40 high contingency) as a between-subjects factor to the ANOVA to examine whether this increase in the magnitude of the contingency learning effect over blocks differed depending on the probability with which high-contingency trials occurred. Consistent with this possibility, the ANOVA revealed a marginally significant linear contrast of the Contingency \times Block \times Group three-way interaction, $F(4, 242) = 2.32$, $MSE = 1830.45$, $p = .06$, $\eta^2 = .04$. We therefore conducted ANOVAs for each group separately to investigate which group(s) showed an increasing contingency learning effect over blocks. The linear contrast of the Contingency \times Block interaction was significant (and robust) for the .70 high-contingency group, $F(1, 48) = 12.15$, $MSE = 1940.41$, $p = .001$, $\eta^2 = .20$, but was nonsignificant for the .80 high-contingency group, $F(1, 49) = 2.29$, $MSE = 2775.83$, $p = .14$, $\eta^2 = .04$, and was absent for the .60, .50, and .40 groups ($F_s < 1$). Thus, participants in the .70 high-contingency group drove the overall effect of the contingency effect increasing significantly in magnitude over blocks. For the .80 high-contingency group, contingency learning may be too rapid to observe a practice effect (see also Schmidt et al., 2010, for evidence of a rapid onset). In contrast, for the .60 high-contingency group (and below), practice effects may take extremely high statistical power to detect because the overall contingency learning effect is quite weak and less consistent across participants. The .70 high-contingency group may therefore represent a "Goldilocks zone" in which the contingency learning effect is robust enough to be fairly consistent across participants but not so robust as to have an immediate and sustained impact on performance. Further research is needed to test this possibility.

Table 1 The proportion of participants in each high-contingency (HC) group who were subjectively aware and the proportion who were objectively aware of the color-word contingencies

Awareness	.40 HC	.50 HC	.60 HC	.70 HC	.80 HC
Subjectively aware	.34	.32	.33	.63	.70
Objectively aware	.12	.36	.42	.65	.80

Errors Table 3 shows participants' mean error rates for high-contingency and low-contingency trials over the six blocks (see also Fig. 4). An ANOVA revealed the expected significant main effect of contingency, $F(1, 246) = 36.36$, $MSE = 0.003$, $p < .001$, $\eta^2 = .13$, and a significant linear contrast of block, $F(1, 246) = 34.70$, $MSE = 0.002$, $p < .001$, $\eta^2 = .12$, suggesting that error rates increased over blocks, likely due to fatigue. Of main interest, the linear contrast of the Contingency \times Block interaction was nonsignificant ($F < 1$). Thus, in contrast to the response-time data, there was no evidence in the error data that the contingency learning effect increased linearly over blocks.

Discussion

In this study, we have mapped out the relation between the proportion of high-contingency trials and the magnitude of the contingency learning effect (the difference in responding to the high-contingency and low-contingency stimuli) for the standard color-word contingency learning paradigm. As the proportion of high-contingency trials declines steadily from 80% to 40%, so too does the size of the contingency learning effect, disappearing at around 40% high contingency. The function appears to be exponential (see Fig. 1).

These findings can be accommodated by the Parallel Episodic Processing (PEP) model of Schmidt (2013; Schmidt et al., 2016). As Schmidt has repeatedly found (e.g., Schmidt et al., 2010) and as we have also observed

Table 2 Response times in milliseconds (with *SEs*) for high-contingency (HC) trials and low-contingency (LC) trials across the six blocks (of 50 trials each)

Block	HC	LC	CL effect	$t(246)$	p	d
1	573 (5.32)	588 (6.49)	15	3.64	<.001	0.16
2	550 (5.07)	562 (5.89)	12	3.70	<.001	0.14
3	549 (4.75)	567 (5.40)	18	5.21	<.001	0.23
4	559 (5.20)	577 (6.19)	18	4.18	<.001	0.20
5	565 (5.53)	593 (6.29)	28	6.50	<.001	0.31
6	573 (5.88)	595 (6.26)	22	5.43	<.001	0.23

Note. Also shown for each block are the size of the contingency learning (CL) effect, the paired-samples t statistic, the p value, and the effect size (Cohen's d)

Table 3 Errors as proportions (with *SEs*) for high-contingency (HC) trials and low-contingency (LC) trials across the six blocks (of 50 trials each)

Block	HC	LC	CL effect	$t(246)$	p	d
1	.022 (.002)	.036 (.003)	.014	3.66	<.001	.32
2	.027 (.002)	.037 (.003)	.010	2.94	.004	.24
3	.030 (.002)	.036 (.003)	.006	1.38	.17	.12
4	.033 (.003)	.044 (.004)	.011	2.59	.01	.22
5	.035 (.003)	.042 (.004)	.007	1.66	.10	.14
6	.032 (.003)	.053 (.004)	.021	4.89	<.001	.39

Note. Also shown for each block are the size of the contingency learning (CL) effect, the paired-samples t statistic, the p value, and the effect size (Cohen's d)

elsewhere (Forrin & MacLeod, 2017; Lin & MacLeod, 2017), contingency learning onsets very quickly. The routine retrieval of prior instances recovers increasingly many high-contingency instances and relatively fewer low-contingency instances as trials progress, resulting in diverging response functions for high-contingency and low-contingency stimuli.

As the PEP model would anticipate, this divergence should be increasingly apparent (in terms of both response time and accuracy) as high-contingency proportion increases, exaggerating the difference in activation between high contingency and low contingency. That is precisely what we observed. Moreover, the threshold for obtaining a contingency learning effect was approximately 40-30-30 (40% high-contingency pair, 30% for each of the two low-contingency pairs; see Fig. 1), which suggests that a contingency learning effect will emerge when the difference in the proportions of high-contingency and low-contingency trials exceeds 10%. (Of course, the smaller this difference in proportions the more participants would be needed to achieve adequate statistical power to reliably obtain a contingency learning effect.)

Although we did not examine this case here—the two low-contingency pairs need not have equal proportions (see Schmidt & De Houwer, 2016a). It seems likely that a different threshold for obtaining the contingency learning effect would emerge if the proportions of the low-contingency pairs differed (e.g., 40-35-25) because this would result in a larger difference in response activation between the highest (40%) and the lowest (25%) contingency pairs relative to when the low-contingency proportions are identical (i.e., 40-30-30).

There is, however, one feature of our data that conflicts with the PEP model as currently framed. The model predicts that, as high-contingency proportion increases, the growth in the contingency learning effect should result from increasingly better performance on the increasingly frequent high-contingency trials. Although this empirical pattern was observed for error rates, in terms of response times we found that the growth in the contingency learning effect appeared

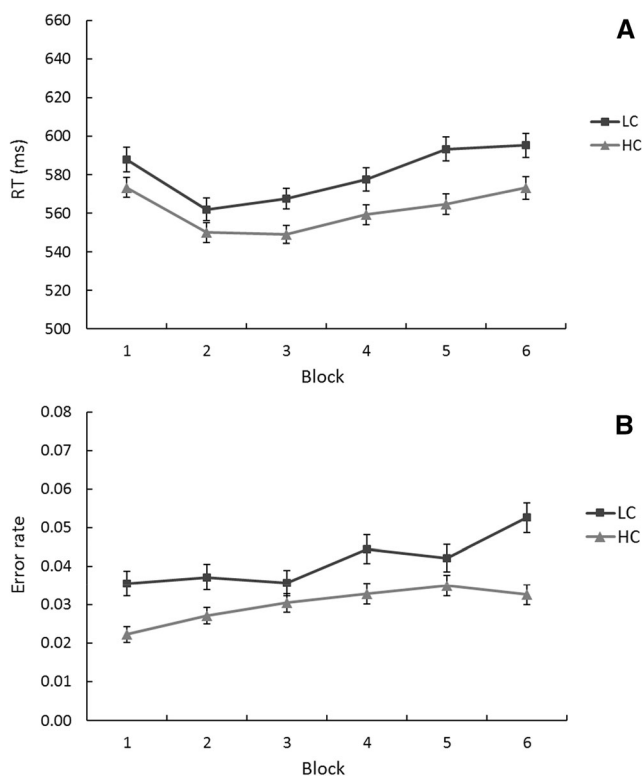


Fig. 4 **a** Mean response times for the high-contingency (HC) and low-contingency (LC) trials as a function of block. **b** Mean error rates for the high-contingency and low-contingency trials as a function of block. Error bars are standard errors

to derive instead from increasingly poor performance on the increasingly rare low-contingency trials. We do not see this as a fatal flaw in the model, however. Instead of seeing performance as the outgrowth of the retrieval of increasing numbers of high-contingency instances, the model simply has to recognize a larger role for the increasingly rare low-contingency instances in determining performance. This substantial influence of rare events is certainly not unprecedented: It is, for example, well established in the attention literature (e.g., Hon & Tan, 2013; Wolfe, Horowitz, & Kenner, 2005).

Decreasing the proportion of low-contingency trials may have slowed response times for those trials not only because of their decreasing response activation (in line with the PEP model) but also because of increasing response competition (i.e., interference) from the high-contingency responses. Interference has long been known to slow responses to incongruent trials in the Stroop task (Stroop, 1935; for a review see MacLeod, 1991), the finding that color naming is slower for incongruent stimuli (e.g., GREEN_{blue}) than for congruent (e.g., BLUE_{blue}) or control (e.g., XXXX_{blue}) stimuli. Moreover, increased facilitation of congruent stimuli in the Stroop task is related to increased interference of incongruent stimuli, which can result in a positively skewed, ex-Gaussian distribution of incongruent stimuli (see Spieler, Balota, & Faust, 2000). Similarly, then, response times for improbable

low-contingency pairs may be impaired by response competition from probable high-contingency pairs, stretching the tail of the low-contingency distribution.

In sum, there are two factors that could account for poor performance on low-contingency trials: First, they have few instances in memory (and correspondingly low response activation), and second, they have to overcome response competition (interference) from the high-contingency response. Although these mechanisms are not separable in the standard color-word contingency paradigm as used in the present experiment, Schmidt and De Houwer's (2016a) finding that there is worse performance for low-contingency trials than for medium contingency trials suggests that increasingly improbable trials are responded to increasingly slowly when response competition is held constant (interference from high-contingency trials ought to have been equivalent in the medium contingency and low-contingency conditions). And in a second experiment, they did not find evidence of response interference for low-contingency items compared to a novel-word (i.e., no-contingency) baseline. Likewise, the inverse relation observed here between the proportion of low-contingency trials and their response times may have had more to do with decreasing response activation than with increasing response interference. Further research certainly is needed to explore this activation versus interference issue. In particular, manipulations of contingency that result in non-equivalent low-contingency proportions should be undertaken to examine the influence of low-contingency proportion.

The nonsignificant relation between high-contingency proportion and high-contingency response times is difficult for the PEP model to account for. One possibility is that the experimental instructions led to a floor effect for high-contingency trial response times. Recall that the instructions emphasized not only speed but also accuracy (as is standard for contingency learning—and indeed most response-time-based—experiments). This may have encouraged some participants—including those who received a high proportion of high-contingency trials—to constrain their response speed to minimize their error rate, which would have weakened the relation between the manipulated proportion of high-contingency trials and the response times for those trials.

The second aim of this study was to examine the relation between contingency awareness and contingency learning. We found that both subjective and objective awareness were significantly correlated with the magnitude of the contingency learning effect. Although Schmidt and De Houwer (2012b) found a nonsignificant correlation between objective awareness and the size of the contingency learning effect, they noted that this could have been attributable to a restriction of range (i.e., most of their participants were objectively aware). Consistent with their speculation, the wide range of proportions in the present experiment remedied the restricted range issue and a significant correlation was in fact

observed. Indeed, objective awareness remained a significant predictor when included in a regression with both high-contingency proportion and subjective awareness. That subjective awareness became a nonsignificant predictor may be because it is a less precise measure of awareness: Of participants who claimed to be subjectively aware, 40% were objectively unaware (i.e., they did not correctly identify the high-contingency pairs).

Objective awareness was significantly related to faster response times for high-contingency trials, but not to slower response time for low-contingency trials. This result replicates Schmidt and De Houwer's (2012b) finding of a larger contingency learning effect in aware participants, and is consistent with their claim that this effect is driven by faster high-contingency response times. Objective awareness may lead participants to direct more attention to the nominally "irrelevant" word information (Schmidt & De Houwer 2012b) to help speed their responses, and may strengthen the association between each word and its respective high-contingency pair. Thus, objective awareness may increase response activation for high-contingency pairs (in line with the PEP model) without influencing activation for low-contingency pairs. As we noted above, Schmidt and De Houwer's (2016a) prior research suggests that increased response activation for high-contingency responses does not adversely affect low-contingency responses (via response interference).

Last, in a more fine-grained analysis, we observed the contingency learning effect to slowly increase—approximately linearly—across successive blocks of trials (i.e., a practice effect). Recently, after we had conducted this experiment, Schmidt and De Houwer (2016b) also found practice effects and hypothesized that two criteria had to be met (both of which were intended to increase the slope of the contingency learning effect)—that there is a practice phase, and that the word starts out in a neutral color. The present experiment incorporated neither of those two criteria, yet we still found a significant increase in the magnitude of the contingency learning effect over trials. Consequently, our study constitutes novel evidence of a practice effect using the "standard" contingency learning procedure. We suspect that the high statistical power conferred by our large samples helped to reveal this effect.

Although this experiment is the first to chart the size of the contingency learning effect over a wide range of high-contingency proportions, numerous experiments have tracked the Stroop effect over different proportions of congruent trials (e.g., Cheesman & Merikle, 1986; Lindsay & Jacoby, 1994; Jacoby, Lindsay, & Hessels, 2003; for a wide range of proportions, see Blais, Harris, Guerrero, & Bunge, 2012; for a review, see MacLeod, 1991). Those experiments have revealed a proportion-congruent effect (Lowe & Mitterer, 1982): Increasing the proportion of congruent trials increases the magnitude of the Stroop effect—the same empirical

pattern as we observed for the contingency learning effect in the present research. Although the resemblance between the proportion-congruent effect and the "proportion-contingent" effect observed here aligns with Schmidt and Besner's (2008) claim that the proportion-congruent effect is a contingency effect, a growing body of research has challenged this claim (e.g., Bugg, 2014; Bugg, Jacoby, & Chanani, 2011; Gonthier, Braver, & Bugg, 2016). Indeed, in accord with the possibility that different factors drive these two effects, our present finding that objective contingency awareness modulates the size of the contingency learning effect can be contrasted with the prior result that awareness does not affect the magnitude of the proportion congruent effect (Blais et al., 2012).

To conclude, the present experiment revealed two factors that were related to the magnitude of the contingency learning effect in response times. The first (and more robust) was proportion: In particular, increasing the proportion of high-contingency trials led to increasingly slow responding on low-contingency trials, which increased the size of the contingency learning effect. The second relevant factor was objective awareness: Overall, objectively aware participants tended to have faster response times to high-contingency trials than did objectively unaware participants. In exploring a broad range of contingencies in a study with considerable power, we have expanded what is known about this kind of fundamental learning and have provided a further test of Schmidt's (2013, Schmidt et al., 2016) Parallel Episodic Processing model. With the modification of increased emphasis on the role played by the relatively rare, low-contingency events, the model provides a successful account of the findings thus far. As Lin and MacLeod (2017) have argued, the color-word contingency learning paradigm provides a direct way to examine the learning of simple associations, and as such is a valuable addition to our set of cognitive tools.

Author note This research was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant A7459. We thank Mackenzie Mullen and Deanna Priori for their assistance in collecting the data. Correspondence may be addressed to nforin@uwaterloo.ca or cmacleod@uwaterloo.ca

References

- Blais, C., Harris, M. B., Guerrero, J. V., & Bunge, S. A. (2012). Rethinking the role of automaticity in cognitive control. *The Quarterly Journal of Experimental Psychology*, *65*, 268–276.
- Bugg, J. M. (2014). Evidence for the sparing of reactive cognitive control with age. *Psychology and Aging*, *29*, 115–127.
- Bugg, J. M., Jacoby, L. L., & Chanani, S. (2011). Why it is too early to lose control in accounts of item-specific proportion congruency

- effects. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 844–859.
- Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology*, 40, 343–367.
- De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and theories on human contingency learning. *The Quarterly Journal of Experimental Psychology*, 55B, 289–310.
- Forrin, N. D., & MacLeod, C. M. (2017). Relative speed of processing determines color-word contingency learning. *Memory & Cognition*. doi:<https://doi.org/10.3758/s1342>
- Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive control in the Stroop task. *Memory & Cognition*, 44, 778–788.
- 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.
- Jacoby, L. L., Lindsay, D. S., & Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin & Review*, 10, 638–644.
- Lin, O. Y.-H., & MacLeod, C. M. (2017). The acquisition of simple association as observed in color-word contingency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi:<https://doi.org/10.1037/xlm000043>
- Lindsay, D. S., & Jacoby, L. L. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 219–234.
- Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 492–527.
- Lowe, D. G., & Mitterer, J. O. (1982). Selective and divided attention in a Stroop task. *Canadian Journal of Psychology*, 36, 684–700.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109(2), 163–203.
- Miller, J. O. (1987). Priming is not necessary for selective attention failures: Semantic effects of unattended, unprimed letters. *Perception & Psychophysics*, 41, 419–434.
- Morton, J., & Chambers, S. M. (1973). Selective attention to words and colours. *The Quarterly Journal of Experimental Psychology*, 25, 387–397.
- 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.
- 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.
- 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.
- 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.
- 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.
- Schmidt, J. R., & De Houwer, J. (2012b). Learning, awareness, and instruction: Subjective contingency awareness does matter in the colour-word contingency learning paradigm. *Consciousness and Cognition*, 21, 1754–1768.
- Schmidt, J. R., & De Houwer, J. (2016a). Contingency learning tracks with stimulus-response proportion: No evidence of misprediction costs. *Experimental Psychology*, 63, 79–88.
- 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.
- 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.
- 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.
- Spieler, D. H., Balota, D. A., & Faust, M. E. (2000). Levels of selective attention revealed through analyses of response time distributions. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 506–526.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662.
- Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual searches. *Nature*, 435, 439–440.