Research Article

Contingency Learning With Evaluative Stimuli

Testing the Generality of Contingency Learning in a Performance Paradigm

James R. Schmidt and Jan De Houwer

Ghent University, Ghent, Belgium

Abstract. In two experiments, we tested the generality of the learning effects in the recently-introduced color-word contingency learning paradigm. Participants made speeded evaluative judgments to valenced target words. Each of a set of distracting nonwords was presented most often with either positive or negative target words. We observed that participants responded faster on trials that respected these contingencies than on trials that contradicted the contingencies. The contingencies also produced changes in liking: in a subsequent explicit evaluative rating task, participants rated positively-conditioned nonwords more positively than negatively-conditioned nonwords. Interestingly, contingency effects in the performance task correlated with this explicit rating effect in both experiments. In Experiment 2, all effects reported were independent of subjective and objective contingency awareness (which was completely lacking), even when awareness was measured at the item level. Our results reveal that learning in this type of performance task extends to nonword-valence contingencies and to responses different from those emitted during the performance task. We discuss the implications of these findings for theories about the processes that underlie contingency learning in performance tasks and for research on evaluative conditioning.

Keywords: contingency learning, evaluative conditioning, performance tasks, acquisition, contingency awareness, S-R learning

How we learn the relations between events in our environment is one of the key theoretical questions in the cognitive and social cognitive research domains (e.g., Allan, 2005; Fiedler, 1991; Rescorla, 1967). In particular, research on human contingency learning aims to understand the conditions under which and mechanisms by which participants learn to relate events that covary statistically (Schmidt, in press). Contingency learning is studied in a wide variety of paradigms. One particularly useful one is the color-word contingency learning paradigm introduced by Schmidt and colleagues (e.g., Schmidt & Besner, 2008; Schmidt & De Houwer, 2011). In contrast to the study-test procedures used in many learning paradigms (e.g., Musen & Squire, 1993), the color-word contingency learning paradigm is a very simple performance task that requires only one simple phase in which learning occurs and is measured simultaneously. In particular, each distracter word is presented most often in one target color (e.g., MOVE most often in blue, SENT most often in green, etc.). Learning is indicated by faster and more accurate responses to the color on high contingency trials (e.g., MOVE in blue) relative to low contingency trials (e.g., MOVE in green). Recent studies suggest that learning effects in this paradigm result from facilitation on high contingency trials rather than interference on low contingency trials (Schmidt & Besner, 2008). Furthermore, learning is dependent on having available memory resources (Schmidt & Besner, 2008), but does not seem to require contingency awareness (Schmidt, Crump, Cheesman, & Besner, 2007). Most critically, in addition to being a very simple paradigm to program and use, the resulting contingency effect is extremely reliable (i.e., very small sample sizes are required) and is acquired very quickly (e.g., significant in the first block of 18 trials in Schmidt, De Houwer, & Besner, 2010), making it a highly advantageous approach to study learning.

Although the color-word contingency learning paradigm has many practical advantages as a tool for studying learning, little is known about how specific the observed effects are and thus about the extent to which research with this paradigm informs us about learning in general. If the effects are limited to contingencies between colors and words, for instance, then research with the color-word paradigm is unlikely to tell us much about the many different instances of learning in daily life. We therefore created a variant of the color-word contingency paradigm that differs from the standard paradigm in two important ways. First, we implemented contingencies between neutral and valenced stimuli. More specifically, participants were presented with distracting nonwords and target evaluative words that they had to rate as positive or negative. Each distracting nonword was presented most often with either a positive or a negative word (e.g., NIJARON with HUG and KADIRGA with

GUNS). On high contingency trials the nonword is presented with the expected valence response (e.g., NIJARON with HUG) and on low contingency trials the nonword is presented with the opposite valence (e.g., NIJARON with GUNS). Contingency learning is evidenced by faster responses on high relative to low contingency trials. Second, we examined the impact of the contingencies on liking of the nonwords. After the reaction time task, we tested whether the nonwords that co-occurred most often with positive targets were liked more than nonwords that co-occurred most often with negative targets. This would demonstrate that the effects of the contingencies generalize well beyond the speeded responses that are performed during the reaction time task and would thus provide important information about the processes underlying these learning effects.

Experiment 1 further assessed whether participants learn specific nonword-word pairs (e.g., NIJARON → HUG) or more general nonword-valence relationships (e.g., NIJARON → positive) by including trials in which nonwords were presented with a different target word of the same valence as the high contingency response (e.g., NIJARON → FLOWER). With a similar goal in mind, nonwords in Experiment 2 were presented most often with many different target words of a particular valence, rather than with one specific target word.

In addition to testing the generality of learning in the color-word contingency task, our studies could also contribute to the literature on evaluative conditioning. Evaluative conditioning (see Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010, for a review) refers to a change in liking of a stimulus that results from the pairing of that stimulus with another stimulus (De Houwer, 2007). For instance, if one neutral picture is repeatedly presented along with a smiling face and another neutral picture is repeatedly presented with an angry face, then participants will subsequently rate the former neutral picture more positively than the latter. If we observe that the nonword-valence contingencies result in a change in liking of the nonwords, this effect would therefore qualify as evaluative conditioning and would offer interesting new ways to study evaluative conditioning.

Experiment 1

The primary goal of Experiment 1 was to assess whether learning effects can be observed in our evaluative version of the color-word contingency paradigm. We also assessed whether evaluative conditioning in this paradigm results from the semantic linking of the neutral nonword to the evaluative target word with which it is paired (i.e., stimulus association) or from the co-occurrence of the nonword and the valenced response (i.e., stimulus-valence learning). To do this we used three conditions (for similar manipulations, see De Houwer, 2003; Schmidt et al., 2007; Schmidt & Cheesman, 2005). On *stimulus match* trials, the distracter was presented with its most frequent target (e.g., NIJARON → HUG). On *valence match* trials, the distracter was presented with a target of the same valence as its most frequent target (e.g., NIJARON → FLOWER). On *valence mismatch* trials, the distracter was presented with a target of the *opposite* valence as its most frequent target (e.g., NIJARON → GUNS). If conditioning is driven by stimulus associations, then performance should be better for stimulus match trials relative to the other two conditions. If conditioning is driven by stimulus-valence learning, however, then performance should be poorer for valence mismatch trials relative to the other two conditions. Response time learning effects were further analyzed across blocks to study acquisition speed.

Method

Participants

Nineteen Ghent University undergraduates participated in Experiment 1 in exchange for course credit or €4.

Apparatus

Participants responded with an AZERTY keyboard. Stimulus and response timing were controlled by E-Prime (Experimental Software Tools, 2002).

Materials and Design

The experiment used four 7-letter distracting nonwords (nijaron, kadirga, fevkani, lokanta) and four 7-letter target Dutch words, two positively valenced (bloemen [flowers], knuffel [hug]) and two negatively valenced (misdaad [crime], geweren [guns]).

Procedure

Participants started with the reaction time task. They sat approximately 60 cm from the screen. Stimuli were presented in white on a black background. On each trial, participants saw a fixation cross for 250 ms, followed by blank screen for 50 ms, followed by a distracting nonword just above or just below fixation for 250 ms, followed by the addition of the target to the screen in the remaining location (just below or just above fixation) for 2,000 ms or until a response was made. Participants were instructed to respond as quickly as possible by pressing one key for positive targets and another key for negative targets. Mapping of valence to key was counterbalanced: the left "F" key for positive responses and the right "J" key for negative responses, or vice versa. After the target presentation, a blank screen for correct responses or "XXX" in red for incorrect and missed responses was presented for 500 ms. Each of the four nonwords was presented most often (8 out of 11 times) with one target, and equally often (1 of 11 times) with the other three targets. There were two distracters that were most often followed by a positive word and two most often followed by a negative word. Orthogonal to this, each trial was presented with the target above

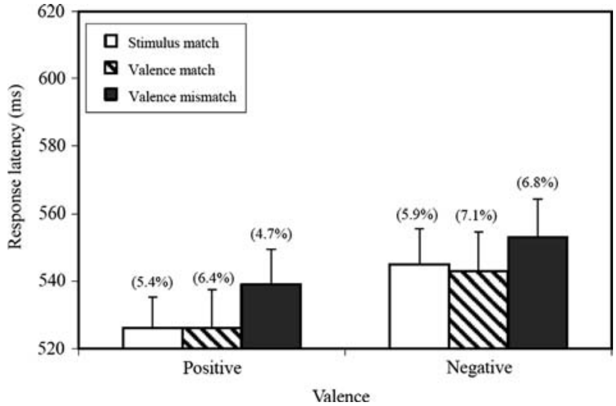

Figure 1. Experiment 1 response latencies with standard errors and percentage errors for trial type and valence.

fixation on half of the trials and below fixation on the other half of the trials. The location that the target and distracter appeared in varied randomly from trial to trial.[1] These manipulations create a block size of 88 trials. There were five blocks for a total of 440 trials.

In the final task,[2] participants were asked to explicitly rate how much they liked each word on a scale from 1 (dislike) to 8 (like). Responses were made with the number pad on the keyboard. Both the distracting nonwords and the target valenced words were rated by participants. All stimuli were presented in a random order.

Results

For the reaction time task, mean correct response times and error percentages were analyzed (see Figure 1). Missed responses were deleted from analyses (less than 1% of the data). Because complete repetitions (both the target word and distracter nonword are the same as the previous trial) greatly speed responding and are disproportionately represented in the stimulus match condition, these trials were also trimmed from analysis.[3] Because we were testing one-tailed hypotheses (e.g., that high contingency trials are responded to *faster* than low contingency trials), one-tailed *t*-tests are reported throughout the paper.

Table 1. Experiment 1 high and low contingency response times by block

	Block				
	1	2	3	4	5
High contingency (ms)	566	525	526	525	536
Low contingency (ms)	563	539	530	537	556
Effect (ms)	*−3*	*14*	*4*	*13*	*20*

Response Latencies

An ANOVA for response latencies with the factors of contingency (stimulus match vs. valence match vs. valence mismatch) and valence (positive vs. negative) revealed a significant main effect of contingency, $F(2, 36) = 3.629$, $MSE = 382$, $p = .037$, $\eta_p^2 = .17$. It also revealed a significant main effect of valence, $F(1, 18) = 12.252$, $MSE = 636$, $p = .003$, $\eta_p^2 = .41$, indicating faster overall responses for positive relative to negative targets. There was no interaction, $F(2, 36) = .177$, $MSE = 389$, $p = .839$, $\eta_p^2 < .01$. The data were then averaged across valence to assess the main effect of contingency. Planned comparisons did not reveal a significant difference between stimulus match (536 ms) and valence match trials (535 ms), $t(18) = .264$, $SE_{diff} = 5$, $p = .602$, $\eta_p^2 < .01$, suggesting the absence of stimulus-specific effects. This comparison had high power (.8) to detect an effect as small as 11 ms. Planned comparisons further revealed that valence mismatch trials (546 ms) were responded to slower than both stimulus match, $t(18) = 2.778$, $SE_{diff} = 4$, $p = .006$, $\eta_p^2 = .30$, and valence match trials, $t(18) = 2.100$, $SE_{diff} = 5$, $p = .025$, $\eta_p^2 = .20$. When stimulus and valence match trials were treated as one trial type (high contingency; 536 ms), responding was predictably faster than valence mismatch (low contingency) trials, $t(18) = 2.845$, $SE_{diff} = 3$, $p = .005$, $\eta_p^2 = .31$. The data for high and low contingency trials divided by block are presented in Table 1. As is typical of response time tasks, overall responding increased in speed across blocks, as indicated by a significant main effect of block, $F(4, 72) = 4.140$, $MSE = 2186$, $p = .004$, $\eta_p^2 = .19$. However, block did not interact with contingency, $F(4, 72) = 1.062$, $MSE = 717$, $p = .381$, $\eta_p^2 = .06$, indicating that learning occurred very early on in the task.[4]

---

1 This random variation in target location is important. Participants need to (partially) attend to *both* the target and the distracter to determine which stimulus is the target. Increasing attention to the distracter allows it to have a larger effect on target processing. Indeed, initial attempts at finding evaluative conditioning with a flanker experiment with a consistent target location produced lacklustre (though consistent) results (but see, Miller, 1987, for a contingency learning flanker paradigm that does work).

2 Before the rating task, participants also completed a variant of the affect misattribution procedure (Payne, Cheng, Govorun, & Stewart, 2005) that used the conditioned nonwords and the unconditioned valenced targets as primes. This procedure did not produce any effects, even when the prime was an unconditioned valenced word and is therefore not discussed further. Full details of the task we used can be obtained on request from the authors.

3 This trim, as in our previous report with a non-evaluative task (Schmidt et al., 2007), does affect the pattern of results in one notable way. Because a large proportion of stimulus match trials are complete repetitions, this condition will appear as being faster than the valence match condition if this sequential confound is not controlled for.

4 Inspection of the data in Table 1 seems to indicate that the learning effect was not present in Block 1. A post hoc comparison between Block 1 and the remaining four blocks was only marginal, $F(1, 18) = 3.559$, $MSE = 326$, $p < .075$, $\eta_p^2 = .17$, and nonsignificant after an (appropriate) correction for this unplanned test. The data in Experiment 1 are simply noisy due to the small sample size ($n = 19$) for the small blocks. Experiment 2 has much more power ($n = 39$) and shows no hint of a smaller effect in the first block.

Percentage Error

An ANOVA for response latencies with the factors of contingency (stimulus match vs. valence match vs. valence mismatch) and valence (positive vs. negative) did not reveal a main effect of contingency, $F(2, 36) = .922$, $MSE = 15$, $p = .407$, $\eta_p^2 = .05$, valence, $F(1, 18) = 1.535$, $MSE = 22$, $p = .231$, $\eta_p^2 = .08$, or an interaction, $F(2, 36) = .381$, $MSE = 18$, $p = .686$, $\eta_p^2 = .02$. Thus, errors were generally uninformative, but did reveal that the response latency results are unlikely to be due to a speed-accuracy tradeoff.

Explicit Rating

Valence ratings were significantly more positive for the positively-conditioned nonwords (5.55) than for the negatively-conditioned nonwords (3.68), $t(18) = 2.101$, $SE_{\text{diff}} = .41$, $p = .025$, $\eta_p^2 = .20$, indicating a learned preference for the positively-conditioned relative to negatively-conditioned nonwords. The effect was predictably much larger for the inherently affective real words (7.55 vs. 1.47, respectively), $t(18) = 24.750$, $SE_{\text{diff}} = .25$, $p < .001$, $\eta_p^2 = .97$.

Correlation

To reduce the impact of outliers, nonparametric correlation (Spearman's ρ) was used (the parametric test returned similar results) to examine the relation between learning as indexed by reaction times and by ratings. The correlation was significant, $\rho(17) = .447$, $p = .027$.

Discussion

Experiment 1 was successful in producing a reaction time learning effect using our new reaction time paradigm. Specifically, participants identified the valence of a target word faster if the distracting nonword was presented most often with the same relative to different valence as the target. It was further found that participants did not learn the specific target associated with the distracting nonword, as indicated by the lack of a difference between stimulus match and valence match trials. Instead, participants learned the valence associated with the distracting nonword, as indicated by the difference between valence match and valence mismatch trials.

We also observed that explicit ratings of the positively-conditioned nonwords were more positive than the explicit ratings of the negatively-conditioned nonwords, indicating that participants did learn the conditioned valence of the nonwords. Further, the response time and explicit rating effects were positively correlated. Finally, learning did not seem to be modulated by block, indicating that it emerged very quickly. However, given the small sample size, this should be interpreted cautiously (see Footnote 4).

Experiment 2

In Experiment 2, several changes to the paradigm were made to test the generalizability of our reaction time task to differing task parameters. Instead of having only four target words (two positive and two negative), Experiment 2 had 24 target words (12 positive and 12 negative). Additionally, instead of pairing each nonword most often with a particular target word (as in Experiment 1), each of four nonwords was presented most often with *all* words of a particular valence. Specifically, two nonwords were presented four times more frequently with each of the 12 positive target words and two other nonwords were presented four times more frequently with the negative target words. This was the first time we have studied learning in this way. The explicit rating task was again included to examine changes in liking. At the end of the experiment, participants were tested for contingency awareness.

Method

Participants

Thirty-nine Ghent University undergraduates participated in Experiment 2 in exchange for €4.

Apparatus

The apparatus of Experiment 2 was identical to that of Experiment 1.

Materials and Procedure

Experiment 2 was identical to Experiment 1 with a few exceptions. Instead of four target words, there were 24, 12 positively valenced and 12 negatively valenced (see Appendix). The experiment also included subjective and objective awareness measures as well as confidence ratings in guesses on the objective awareness test. The response time task consisted of one block of 480 trials (although this was later divided into six blocks of 80 trials for a block analysis). Two of the nonwords were presented eight times with each positive target and twice with each negative target. The two other nonwords were presented eight times with each negative target and twice with each positive target. The target and distracter were presented equally often in the top versus bottom position for each distracter-target pair. The same four target words from Experiment 1 were used in the explicit rating task along with the four nonword distracters.

Following the explicit rating task, contingency awareness was assessed. First, participants were presented with a screen telling them that each nonword was presented most often with either positive or negative target words and they were asked if they noticed these relations (subjective awareness) by pressing the "j" key for "ja" ["yes"] or the "n" key for "nee" ["no"]. Next, participants were presented

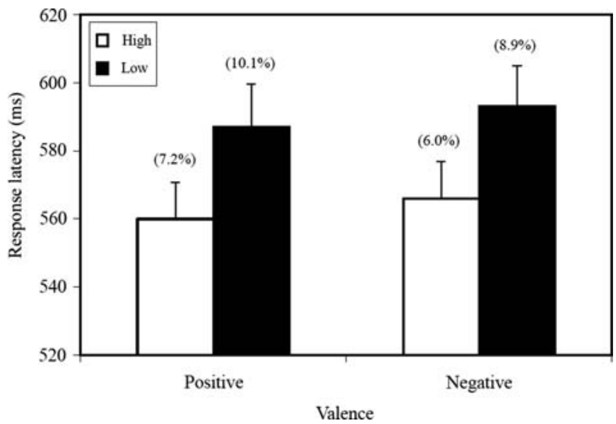

Figure 2. Experiment 2 response latencies with standard errors and percentage errors for contingency and valence.

with a trial for each nonword and were asked to guess whether the nonword was presented most often with positive or negative targets (objective awareness) by pressing the "p" or "n" key, respectively. Following each guess, participants were then asked to rate how confident they were with this guess on a scale from 1 ("niet heel zeker" ["not very sure"]) to 5 ("heel zeker" ["very sure"]).

Results

For the response time task, mean correct response times and error percentages were analyzed (see Figure 2). Missed responses were deleted from analyses (less than 1% of the data).

Response Latencies

Response latencies were analyzed using an ANOVA with the factors contingency (high vs. low) and valence (positive vs. negative) as within variables. It revealed a significant main effect of contingency, $F(1, 38) = 31.303$, $MSE = 922$, $p < .001$, $\eta_p^2 = .45$. The main effect of valence was not significant, $F(1, 38) = 1.644$, $MSE = 819$, $p = .207$, $\eta_p^2 = .04$, nor was the interaction, $F(1, 38) = .003$, $MSE = 481$, $p = .959$, $\eta_p^2 < .0001$. The data were again analyzed by block (arbitrarily set to six blocks of 80, see Table 2). Overall responding was faster across blocks, as indicated by a significant main effect of block, $F(5, 190) = 20.286$, $MSE = 2{,}856$, $p < .001$, $\eta_p^2 = .35$. Again, block did not interact with contingency, $F(5, 190) = .331$, $MSE = 1{,}308$, $p = .894$, $\eta_p^2 < .01$. Hence, there was no evidence that learning varied over blocks.

Table 2. Experiment 2 high and low contingency response times by block

	Block					
	1	2	3	4	5	6
High contingency (ms)	617	560	556	548	546	553
Low contingency (ms)	644	580	588	578	569	574
Effect (ms)	*28*	*20*	*31*	*31*	*24*	*22*

Percentage Error

An ANOVA for errors with the factors of contingency (high vs. low) and valence (positive vs. negative) revealed a significant main effect of contingency, $F(2, 38) = 8.147$, $MSE = 43$, $p = .007$, $\eta_p^2 = .18$. The main effect of valence was marginal, $F(1, 38) = 2.946$, $MSE = 21$, $p = .094$, $\eta_p^2 = .07$. There was no interaction, $F(1, 38) = .056$, $MSE = 14$, $p = .814$, $\eta_p^2 < .01$.

Explicit Rating

Valence ratings were again significantly more positive for the positively-conditioned nonwords (4.95) than for the negatively-conditioned nonwords (3.55), $t(38) = 3.729$, $SE_{diff} = .37$, $p < .001$, $\eta_p^2 = .27$, indicating a learned preference for the positive relative to negative contingency nonwords. The effect was predictably much larger for the inherently affective real words (7.37 vs. 1.56, respectively), $t(38) = 21.336$, $SE_{diff} = .27$, $p < .001$, $\eta_p^2 = .92$.

Correlation

Contingency learning was observed in the response latency, error, and explicit rating data. The correlation between the reaction time and explicit rating effects was again significant, $\rho(37) = .362$, $p = .012$. The correlation between the percentage error and explicit rating effects was also significant, $\rho(37) = .427$, $p = .003$.

Awareness

Of the 39 participants, 14 (36%) reported noticing contingencies (subjective awareness). However, the sample as a whole showed no sensitivity to contingencies on the objective awareness measure, with 51.9% accuracy on the two-choice task not being significantly different from chance (i.e., 50%), $t(38) = .650$, $SE = 3.0$, $p = .255$, $\eta_p^2 = .01$. This was also true when looking at the data of only subjectively-aware participants (55.4%), $t(13) = 1.147$, $SE = 4.6$, $p = .136$, $\eta_p^2 = .09$, and subjectively-unaware participants (50.0%), $t(24) = .000$, $SE = 3.8$, $p = .500$, $\eta_p^2 < .0001$. However, it is worth pointing out that the subjectively-aware participants were at least trending numerically in the correct direction.

More critically, the size of the response time contingency effect was not correlated with subjective awareness, $\rho(37) = .019$, $p = .454$, nor with objective awareness, $\rho(37) = -.182$, $p = .866$. Similarly, the error contingency effect was not correlated with subjective awareness, $\rho(37) = .209$, $p = .100$, nor with objective awareness,

$\rho(37) = -.251$, $p = .938$. Finally, the size of the explicit rating effect was not correlated with subjective awareness, $\rho(37) = .191$, $p = .122$, nor with objective awareness, $\rho(37) = -.187$, $p = .873$. Thus, contingency awareness seems unrelated to contingency learning in this evaluative task.

Item-Level Awareness

Some authors (e.g., Baeyens, Eelen, & Van den Bergh, 1990; Pleyers, Corneille, Luminet, & Yzerbyt, 2007; Stahl & Unkelbach, 2009) suggest an item-level awareness measure is better than a participant-level awareness measure. In our experiment, an item-level measure entails comparing the evaluative effect for nonwords that participants indicated the correct valence for in the objective awareness test (aware) with the evaluative effect for nonwords that participants indicated the incorrect valence for (unaware). For this analysis, two participants had to be deleted from the unaware condition as they made no incorrect valence guesses. The critical interaction between contingency (high vs. low) and item awareness (aware vs. unaware) was not significant for response times, $F(1, 36) = .430$, $MSE = 864$, $p = .516$, $\eta_p^2 = .01$, nor for errors, $F(1, 36) = 1.113$, $MSE = 21$, $p = .298$, $\eta_p^2 = .03$. The critical interaction between conditioned valence (positive vs. negative) and item awareness (aware vs. unaware) was not significant for explicit ratings, $F(1, 5) = 1.000$, $MSE = 3$, $p = .363$, $\eta_p^2 = .17$, though very few participants had data for all cells for this final comparison.

Confidence

As a final measure, confidence in guesses on the objective awareness task was measured. Confidence in correct guesses (3.1) was no better than confidence in incorrect guesses (3.1), $t(36) = .231$, $SE_{diff} = .18$, $p = .410$, $\eta_p^2 < .01$. Thus, even when participants are more confident they are still completely unaware of the correct contingency relationships and are guessing at random.

Discussion

The results of Experiment 2 again show a learning effect in response times with our paradigm, this time with a much larger set of targets and where distracters were contingent with all words of the conditioned valence and not just one word in particular. The response time effect again was not modulated by block, suggesting that learning was quite fast. As in Experiment 1, the learning effect again transferred to the explicit rating task. Furthermore, the response time effect and explicit rating effect were again correlated, suggesting that the two are likely driven (at least in part) by the same learning mechanism. Analyses of the awareness data suggested that learning occurred independently of awareness (though see the General Discussion for some potential caveats).

General Discussion

In two experiments, nonword-valence contingencies that were present during a response time task influenced both performance during that task and subsequent explicit ratings of stimulus liking. We can thus conclude that effects in the color-word contingency learning paradigm are not restricted to color-response contingencies or speeded identification responses. Our results also suggest that similar processes underlie learning in both types of paradigms. In line with this idea, learning in the nonword-valence paradigm has many features in common with learning in the color-word paradigm: Acquisition is (a) rapid (Experiments 1 and 2), (b) response based rather than stimulus based (Experiment 1), and (c) largely independent of contingency awareness (Experiment 2; see Schmidt et al., 2007, for similar findings in the color-word paradigm). As such, our findings support the idea that the processes that produce learning in the color-word contingency learning paradigm are general and could underlie a variety of learning phenomena in daily life. Finally, Experiment 2 further demonstrated that nonword-valence learning can even occur with a large set of stimuli rather than a small set of distracter-target pairs. In the remainder of this section, we discuss the implications of our results for theories about the processes underlying contingency learning and for the literature on evaluative conditioning.

Processes Underlying Contingency Learning

Schmidt and colleagues (2010) argued that the contingency learning effects observed in paradigms such as the color-word contingency learning task may be driven by simple memory storage and retrieval processes. According to this view, on each trial information about the stimuli that were presented and the response that was made are stored in an episodic memory trace. Processing of a distracter (e.g., MOVE) leads to the retrieval of several of these episodes (e.g., episodes where MOVE was presented), from which the likely response can be determined. The present results suggest that these episodes contain information not only about the identity of stimuli and responses but also about their valence. If valence information is stored in episodic traces, then retrieval of these episodes should lead to the retrieval of valence, thus making nonwords that were paired with positive words more likeable than nonwords that were paired with negative words. The finding of an effect on liking further demonstrates that episodic information stored during the response time learning task extends to other tasks outside the context in which it was learned. It is of course possible that processes other than episodic retrieval underlie learning in (variants of) color-word contingency learning. Whatever these processes might be, the present studies tell us that these processes should be able to influence not only identity responses in speeded identification tasks, but also explicit judgments that occur after the speeded identification task and that are correlated with the learning effect in that task. As such, our results impose important constraints on any current or future theory of learning in this type of paradigm.

Experiment 2 further demonstrated that learning in this type of paradigm can even occur when nonwords are correlated with a category of stimuli (e.g., positive words) rather than with a specific stimulus (i.e., as in Experiment 1). A similar finding has been reported in the sequence learning literature by Goschke and Bolte (2007). In their naming experiments, there was no predictable series of stimuli, but there was a predictable series of *categories* of stimuli (body parts, animals, clothing, and furniture). Participants were able to learn these category-based sequential contingencies as indicated by a decrement in performance when the series was switched to random. Of course, learning the *temporal sequence* of things is not necessarily the same as learning the *co-occurrence* of two things. Still, findings such as these suggest that episodes may contain more than simple representations of the stimuli and responses, but also more abstract categorical and valence information.

Our observation of contingency effects independent of awareness is also noteworthy. The role of contingency awareness in contingency learning is an important area of debate in the literature (e.g., Lovibond & Shanks, 2002). In research with non-evaluative paradigms, evidence for learning without awareness has been observed before (e.g., McKelvie, 1987; Nissen & Bullemer, 1987; Schmidt et al., 2007). However, there is also much debate as to what constitutes adequate evidence of learning without awareness. Given this debate, we anticipate that some readers will be satisfied with our operationalization(s) of contingency awareness, whereas others will not (e.g., because we assessed contingency awareness after learning). Hence, we refrain from making strong claims about this issue and realize that, although indicative, our data are unlikely to close the debate about whether learning without contingency awareness is possible.

Evaluative Conditioning

The current response time task may prove advantageous in future evaluative conditioning work. One advantage to our paradigm is that the contingency effect is very reliable, even with a small sample (e.g., $n = 19$ in Experiment 1) and short experiment duration. The use of a performance measure such as this has other inherent advantages, for instance, for measuring evaluative conditioning as it occurs (see also, Kerkhof et al., 2009). This could prove particularly useful in the study of the time course of learning, extinction, and other processes that may take time to evolve. Interestingly, the response latency results of the current two experiments already provide evidence for very rapid learning, as indicated by the lack of block effects. The typical paradigms used for studying evaluative conditioning are non-ideal for studying temporal questions such as these as they often require a study-test design (i.e., a design with a separate test phase after the study phase), a practice that is also typical in non-evaluative learning research (e.g., Musen & Squire, 1993).

It could be argued that the change in response times and errors observed in our contingency learning paradigm reflects stimulus-response learning rather than an actual change in liking. If the change in performance is merely the result of stimuli becoming linked with physical responses (i.e., pressing one of two keys), it would not qualify as an instance of evaluative conditioning (see De Houwer, 2007). However, participants did learn the valence associated with the distracters, as indicated by the explicit rating effect. The explicit rating task used a completely different set of responses and thus learning could not have been the result of links between stimuli and physical responses. The fact that in both experiments the change in ratings correlated quite strongly with the changes in reaction time performance is in line with the idea that the latter changes also reflect changes in liking.

Still, whether the changes in response times and errors provide viable measures of changes in liking may be contestable and further research investigating these issues is certainly welcome. However, our paradigm *did* prove effective for studying explicit rating effects (which are certainly a measure of liking). Part of the effectiveness of our paradigm might be due to the use of a response task during learning, as recently suggested by Gast and Rothermund (2011). We suggest that producing an evaluative response may serve to more strongly bind stimuli and valence information into episodic memory stores. For this reason alone, our paradigm could prove to be a potent alternative to current paradigms for studying evaluative conditioning.

Acknowledgments

This research was supported by a Visiting Postdoctoral Fellowship from the Research Foundation – Flanders (FWO – Vlaanderen) to James R. Schmidt and Jan De Houwer and Methusalem Grant BOF09/01M00209 of Ghent University to Jan De Houwer. James R. Schmidt is now a postdoctoral fellow of the Research Foundation – Flanders (FWO – Vlaanderen).

References

Allan, L. G. (2005). Learning of contingent relationships. *Learning & Behavior, 33*, 127–129.

Baeyens, F., Eelen, P., & Van den Bergh, O. (1990). Contingency awareness in evaluative conditioning: A case for unaware affective-evaluative learning. *Cognition & Emotion, 4*, 3–18.

De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory & Cognition, 31*, 353–359.

De Houwer, J. (2007). A conceptual and theoretical analysis of evaluative conditioning. *The Spanish Journal of Psychology, 10*, 230–241.

Experimental Software Tools. (2002). E-Prime. Retrieved from http://www.pst-net.com

Fiedler, K. (1991). The tricky nature of skewed frequency tables: An information loss account of distinctiveness-based illusory correlations. *Journal of Personality and Social Psychology, 60*, 24–36.

Gast, A., & Rothermund, K. (2011). I like it because I said that I like it. Evaluative conditioning effects can be based on stimulus-response learning. *Journal of Experimental Psychology: Animal Behavior Processes, 37*, 466–476.

Goschke, T., & Bolte, A. (2007). Implicit learning of semantic category sequences: Response-independent acquisition of abstract sequential regularities. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 33*, 394–406.

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin, 136*, 390–421.

Kerkhof, I., Goesaert, E., Dirikx, T., Vansteenwegen, D., Baeyens, F., D'Hooge, R., & Hermans, D. (2009). Assessing valence indirectly and online. *Cognition and Emotion, 23*, 1615–1629.

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes, 28*, 3–26.

McKelvie, S. J. (1987). Learning and awareness in the Hebb digits task. *Journal of General Psychology, 114*, 75–88.

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

Musen, G., & Squire, L. R. (1993). Implicit learning of colour-word associations using a Stroop paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 789–798.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*, 1–32.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An Inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*, 277–293.

Pleyers, G., Corneille, O., Luminet, O., & Yzerbyt, V. (2007). Aware and (dis)liking: Item-based analyses reveal that valence acquisition via evaluative conditioning emerges only when there is contingency awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 130–144.

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review, 74*, 71–80.

Schmidt, J. R. (in press). Human contingency learning. In N. M. Seal (Ed.), *Encyclopedia of the Sciences of Learning*. Springer.

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., & Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the Stroop task. *Canadian Journal of Experimental Psychology, 59*, 132–138.

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. (2011). Does temporal contiguity moderate contingency learning in a speeded performance task? *Quarterly Journal of Experimental Psychology*. Advance online publication.

Schmidt, J. R., De Houwer, J., & Besner, D. (2010). Contingency learning and unlearning in the blink of an eye: A resource dependent process. *Consciousness and Cognition, 19*, 235–250.

Stahl, C., & Unkelbach, C. (2009). Evaluative learning with single versus multiple USs: The role of contingency awareness. *Journal of Experimental Psychology: Animal Behavior Processes, 35*, 286–291.

Received September 12, 2011
Revision received October 24, 2011
Accepted October 27, 2011
Published online March 13, 2012

James R. Schmidt

Ghent University
Henri Dunantlaan 2
B-9000 Ghent
Belgium
E-mail james.schmidt@ugent.be

Appendix

Experiment 2 Stimuli

Positive words	Negative words
cadeau (present)	executie (execution)
droom (dream)	ongeval (accident)
vrede (peace)	ziekte (disease)
geschenk (gift)	oorlog (war)
warmte (warmth)	kanker (cancer)
muziek (music)	tumor (tumor)
geboorte (birth)	bommen (bombs)
leven (life)	drugs (drugs)
liefde (love)	moord (murder)
vriend (friend)	virus (virus)
bloemen (flowers)	misdaad (crime)
knuffel (hug)	geweren (guns)

Erratum

Correction to Schmidt and De Houwer (2012)

https://doi.org/10.1027/1618-3169/a000141

In Schmidt and De Houwer (2012b), we reported a series of experiments investigating evaluative conditioning within a variant of the color–word contingency learning paradigm. In our Experiment 2 (pp. 178–180), in addition to our main analyses on response time, error rate, and explicit rating data, we also included analyses with measures of subjective awareness, objective awareness, and confidence in objective awareness guesses. Recently, however, collaborators of ours discovered a coding error for the objective awareness measure while preparing follow-up work (Gast, Richter, & Ruszpel, 2018). In particular, responses should have been coded as correct if the participant indicated the valence that the nonword prime was initially trained with (i.e., positive for nonwords that were paired most often with positive targets and negative for nonwords that were paired most often with negative targets). Instead, all positive responses were coded as correct (and negative responses as incorrect). Here, we report the corrected tests relating to objective awareness. Note that all other tests (e.g., related to subjective awareness) still hold. We also report two-tailed tests, rather than one-tailed tests (unlike the original report), given that one-tailed tests are generally regarded as inappropriate in any research in which an effect in the unexpected direction could be informative, which is generally always the case in cognition research (Lombardi & Hurlbert, 2009; Ruxton & Neuhauser, 2010).

Objective awareness was not significantly above chance (i.e., 50%), $t(38) = 1.707$, $SE = 5.6$, $p = .096$, $\eta^2 = .07$, though was higher (59.6%) than initially reported (51.9%) and trending (p. 179). Objective awareness was not significantly greater than chance for the subjectively aware (66.1%), $t(13) = 1.505$, $SE = 10.7$, $p = .156$, $\eta^2 = .15$, or unaware participants (56.0%), $t(24) = 0.923$, $SE = 6.5$, $p = .365$, $\eta^2 = .03$. It is noteworthy that there was a hint of an effect for subjectively aware participants, though they were few in number and subjective and objective awareness did not correlate significantly, $\rho(37) = .176$, $p = .285$.

Objective awareness was not correlated with the response time contingency effect, $\rho(37) = .105$, $p = .523$, but was correlated with the error rate effect, $\rho(37) = .460$, $p = .003$, and explicit rating effect, $\rho(37) = .560$, $p < .001$. Thus, there was some evidence that awareness moderated the magnitude of the contingency effect (consistent with results from a non-evaluative version of the paradigm; Schmidt & De Houwer, 2012a, 2012c). However, the contingency effect regression intercept (see Greenwald, Klinger, & Schuh, 1995) at chance guessing (.5) was robustly above zero for response times (26 ms), $t(37) = 5.172$, $SE = 5$, $p < .001$, errors (2.3%), $t(37) = 2.259$, $SE = 1.013$, $p = .030$, and explicit ratings (1.06), $t(37) = 3.167$, $SE = 0.33$, $p = .003$, consistent with implicit learning.

For item-level awareness (p. 180), the critical interaction between contingency (high vs. low) and objective awareness (correct vs. incorrect guess) was not significant for response times, $F(1, 21) = 1.992$, $MSE = 339$, $p = .173$, $\eta^2_p = .09$, or errors, $F(1, 21) = 3.055$, $MSE = 16.0$, $p = .095$, $\eta^2_p = .13$. The interaction between conditioned valence (positive vs. negative) and awareness (correct vs. incorrect guess) was also not significant, $F(1, 6) = 5.629$, $MSE = 3.65$, $p = .055$, $\eta^2_p = .48$, though note that very few participants had observations in all cells and the interaction did trend in the correct direction (3.43). Finally, participants were no more confident in correct (2.71) than in incorrect objective ratings (2.56), $t(21) = 0.745$, $SE = .20$, $p = 0.465$, $\eta^2 = .03$ (p. 180). Globally, some influence of objective awareness was observed, with some tests significant and others suggestive. However, the results for the corrected objective awareness measure (along with the originally-reported findings for subjective awareness) still suggest that the learning effects emerge even in the absence of contingency awareness.

References

Gast, A., Richter, J., & Ruszpel, B. (2018). *Is there evidence for unaware evaluative conditioning in a valence contingency learning task?* Manuscript submitted for publication.

Greenwald, A. G., Klinger, M. R., & Schuh, E. S. (1995). Activation by marginally perceptible ("subliminal") stimuli: Dissociation of unconscious from conscious cognition. *Journal of Experimental Psychology: General, 124*, 22–24. https://doi.org/10.1037/0096-3445.124.1.22

Lombardi, C. M., & Hurlbert, S. H. (2009). Misprescription and misuse of one-tailed tests. *Austral Ecology, 34*, 447–468. https://doi.org/10.1111/j.1442-9993.2009.01946.x

Ruxton, G. D., & Neuhauser, M. (2010). When should we use one-tailed hypothesis testing? *Methods in Ecology and Evolution, 1*, 114–117. https://doi.org/10.1111/j.2041-210X.2010.00014.x

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. https://doi.org/10.3758/s13423-012-0255-5

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

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

History
Received April 24, 2019
Published online July 2, 2019

James R. Schmidt
Université Bourgogne Franche-Comté (UBFC)
LEAD-CNRS UMR 5022 - Pôle AAFE
11 Esplanade Erasme
21000 Dijon
France
james.schmidt@ubfc.fr

Jan De Houwer
Ghent University
Henri Dunantlaan 2
9000 Ghent
Belgium
jan.dehouwer@ugent.be

