Consciousness and Cognition 19 (2010) 235–250

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Contingency learning and unlearning in the blink of an eye: A resource dependent process

James R. Schmidt [a,*], Jan De Houwer [a], Derek Besner [b]

[a] *Department of Psychology, Ghent University, Belgium*
[b] *Department of Psychology, University of Waterloo, Canada*

ARTICLE INFO

Article history:
Received 25 July 2009
Available online 29 January 2010

Keywords:
Contingency learning
Covariation learning
Instance memory
Event files
Connectionist models
Implicit learning
Memory resources
Storage
Retrieval
Unlearning

ABSTRACT

Recent studies show that when words are correlated with the colours they are printed in (e.g., MOVE is presented 75% of the time in blue), colour identification is faster when the word is presented in its correlated colour (MOVE in blue) than in an uncorrelated colour (MOVE in green). The present series of experiments explored the possible mechanisms involved in this colour-word contingency learning effect. Experiment 1 demonstrated that the effect is already present after 18 learning trials. During subsequent *unlearning*, the effect extinguished equally rapidly. Two reanalyses of data from Schmidt, Crump, Cheesman, and Besner (2007) ruled out an account of the effect in terms of stimulus repetitions. Experiment 2 demonstrated that participants who carry a memory load do not show a contingency effect, supporting the hypothesis that limited-capacity resources are required for learning. Experiment 3 demonstrated that memory resources are required for both storage and retrieval processes.

1. Introduction

The ability of humans to learn about contingencies between events in the world has recently re-appeared as a major topic in experimental psychology (e.g., Allan, 2005; Beckers, De Houwer, & Matute, 2007; Mitchell, De Houwer, & Lovibond, 2009). Most often, contingency learning in humans is studied using paradigms in which participants see a series of situations in which stimuli or responses co-occur and are afterward asked to judge the strength of the contingency between the stimuli or responses. Other paradigms allow one to assess learning without asking participants to explicitly judge the strength of contingencies. One version of this is the colour-word contingency learning paradigm (Schmidt, Crump, Cheesman, & Besner, 2007; see also Schmidt & Besner, 2008; Musen & Squire, 1993). For instance, Schmidt and colleagues presented four arbitrary words in four different display colours in a colour identification task using a key press response. Each word was presented in all colours, but more often in a particular colour (e.g., MOVE was presented 75% of the time in blue, SENT 75% of the time in green, etc.). Participants responded faster and made fewer errors on high contingency trials (where the word is presented in its correlated colour; e.g., $\text{MOVE}_{\text{blue}}$) than on low contingency trials (where the word is presented in any other colour; e.g., $\text{MOVE}_{\text{green}}$). To date, little is known about *how* contingency information is actually learned in this paradigm. The present paper briefly reviews past work, discusses several competing accounts, and reports three new experiments and two reanalyses of old data that provide new insights into the mechanisms underlying the form of contingency learning in this paradigm.

* Corresponding author. Department of Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium.
E-mail address: james.schmidt@ugent.be (J.R. Schmidt).

1053-8100/$ - see front matter
doi:10.1016/j.concog.2009.12.016

There are several possible explanations for how contingency relations are learned, but there are a few findings that narrow the field of potential explanations. For instance, awareness of contingency information in the paradigm used here does not seem to be crucial. Very few participants are aware of the contingency manipulation and the size of the colour-word contingency effect is unaffected by a participant's level of awareness (Schmidt et al., 2007, Experiment 3). Thus, awareness of contingencies does not seem to "buy" participants anything; the effect is the same size regardless. This suggests that, independent of the participant's awareness of the task manipulation, learning is implicit. A similar argument has been made from results of a flanker task in which flanking cues were predictive of the response (Carlson & Flowers, 1996), sequence learning (Song, Howard, & Howard, 2007), and other paradigms (e.g., Lewicki, Hill, & Czyzewska, 1992). However, the role of awareness in contingency learning is a highly controversial issue. In particular, there is little consensus on the proper way of assessing awareness and proponents of objective measures of awareness often argue for a small amount of awareness of learned information (e.g., see Fu, Fu, & Dienes, 2007 for a detailed discussion of these issues). We simply note that, at the very least, the results of Schmidt and colleagues are difficult to reconcile with rule-based accounts that demand a role for conscious intention (although such rule-based processes may well play a role in unspeeded judgment tasks; e.g., see De Houwer, 2009; Mitchell et al., 2009 for discussions of propositional accounts of associative learning). As a result, in the rest of the present paper we narrow our focus to implicit learning accounts.

Another important finding of Schmidt and colleagues (2007, Experiment 4) is that the colour-word contingency effect does not simply reduce to stimulus–stimulus association or stimulus familiarity. In the critical experiment, two colours were assigned to the left key (e.g., blue and green) and two others were assigned to the right key (e.g., yellow and orange). If MOVE was presented most often in blue (i.e., $\text{MOVE}_{\text{blue}}$), then participants were faster to make the correct left key response to $\text{MOVE}_{\text{blue}}$ *and* $\text{MOVE}_{\text{green}}$ than they were to make right key responses to $\text{MOVE}_{\text{yellow}}$ and $\text{MOVE}_{\text{orange}}$. Schmidt and colleagues observed no difference in responses to $\text{MOVE}_{\text{blue}}$ and $\text{MOVE}_{\text{green}}$. Thus, it is not the case that MOVE is *associated* with the colour blue (or else $\text{MOVE}_{\text{green}}$ would not have been speeded), nor is it critical that participants saw the stimulus $\text{MOVE}_{\text{blue}}$ more often than the stimulus $\text{MOVE}_{\text{green}}$. Rather, it is critical that MOVE is associated with a left key *response*. When the correct response matches this associated response (for blue or green print), responding is facilitated. These results inform us that the learning mechanism is picking up on the contingencies between the distracting word and the *response*, not the contingencies between the distracting word and the target *colour* (however, it should be noted that effects of stimulus–stimulus associations have been observed in other paradigms; e.g., Colzato, Raffone, & Hommel, 2006). Thus, we narrow our focus here to accounts that posit a relationship between the distracter and the response.

There are a number of accounts that could potentially explain the colour-word contingency effect. The simplest of these can be termed the *repetition account,* which explains the colour-word contingency effect in terms of transient memory effects. There are a few subtle variations of this. In one version, high contingency trials are speeded by the residual activation of the memory of recently encountered matching trials (Bertelson, 1961). For instance, high contingency trials such as $\text{MOVE}_{\text{blue}}$ would often be speeded because $\text{MOVE}_{\text{blue}}$ was recently encountered and the memory of this event is still active, allowing for a quicker response. In contrast, a low contingency trial such as $\text{MOVE}_{\text{green}}$ will rarely be speeded, because the probability of two instances of $\text{MOVE}_{\text{green}}$ occurring close temporally is much less likely. According to a slightly different version of the repetition account, when a stimulus and response occur together the association between them is temporarily strengthened for a period of time. If the same stimulus and response are presented together shortly after this, responding will be facilitated (see Hommel, 1998). Again, high contingency trials are much more likely to have been recently preceded by the same word-response pairing (e.g., $\text{MOVE}_{\text{blue}}$ before $\text{MOVE}_{\text{blue}}$) than are low contingency trials (e.g., $\text{MOVE}_{\text{green}}$ before $\text{MOVE}_{\text{green}}$).

Connectionist accounts such as the simple recurrent network (SRN; Cleeremans & McClelland, 1991; Kinder & Shanks, 2001, 2003) could explain the colour-word contingency effect in terms of a highly interconnected arrangement of nodes in which each trial causes the connection weights between nodes to change. For instance, presentation of $\text{MOVE}_{\text{blue}}$ would lead to an increase in the connection strength between MOVE and the blue response (via a layer of hidden units) and a weakening of other connections (e.g., MOVE to the green response). Unlike what is proposed by the repetition account, this change does not dissipate quickly but is assumed to be relatively permanent until new relevant information is encountered. The idea is that the system uses each trial to update the associations between stimuli and responses to gradually optimise performance by adapting to the statistical regularities in the task. Depending on the learning rates of the model, this process could happen relatively slowly or rapidly.

Finally, we consider a similar but conceptually distinct account based on the storage and retrieval of instances (see Logan, 1988), here termed the *instance account*. According to this instance account, participants in an experiment store a memory of each encountered trial (instance). Each of these instance memories includes information about the stimuli presented along with the response that was executed. Early processing of the word leads to retrieval of a set of the most recently encountered (i.e., most accessible) instances associated with this word (e.g., MOVE leads to retrieval of instances containing MOVE) and from these a response expectancy can be generated. As a result, high contingency trials will tend to be speeded because the system will be able to detect the contingencies in the task and prepare for the high contingency response. Note that the difference between the instance and repetition accounts is that the repetition account purports that individual recently encountered stimuli bias responding, whereas the instance account purports that several recently encountered instances are retrieved and used to determine the likely outcome of the current trial.

As can be seen, there are a number of candidate explanations for the colour-word contingency effect. A number of important questions remain to be answered before the best account can be specified. For instance, we still do not have information about basic issues such as: (1) the number of trials needed to obtain the effect (i.e., acquisition speed), (2) whether and how fast the effect disappears when the contingencies are removed (i.e., extinction speed), and (3) whether contingency effects can be found only when sufficient memory resources are available. Just like studies of acquisition, extinction, and the effect of memory resources were crucial in developing theories about other forms of human contingency learning (e.g., De Houwer, 2009; De Houwer, Thomas, & Baeyens, 2001; Shanks, 2007), examining these three issues in the context of the current contingency learning paradigm should provide important information about the processes underlying this effect. Experiment 1 addresses the first two questions and Experiments 2 and 3 address the final question.

2. Experiment 1

The rate of acquisition of contingency information is an important issue. For instance, if contingency information is both learned and unlearned rapidly, then this would pose a problem for a connectionist model with a low learning rate. It is already known that the colour-word contingency effect appears relatively early on in the course of an experiment. In a block analysis, Schmidt et al. (2007) found that the contingency effect was already significant in the very first block of 48 trials. The first goal of Experiment 1 is to increase the resolution of the block analysis by using smaller blocks of 18 trials. One possible outcome is that a contingency effect occurs very early on, perhaps in the first block of 18 trials, indicating that very few trials need to be experienced before contingency information can be extracted. Such a finding would be consistent with any model that is able to alter responding based on a limited sample of trials. This includes the repetition account, which explains the effect in terms of transient connections or activations and for connectionist accounts with a high learning rate. According to low learning rate connectionist accounts, however, acquisition should be slower and participants would need to accumulate experience with several blocks of trials before the effect emerges.

In an instance framework, understanding how fast a contingency is learned does not necessarily provide us with much information on how much data the system can take into account. For instance, imagine an instance account in which the system calculates the most likely response based on the identity of the word using the last 100 trials (a relatively large window) that it has encoded. Presumably, the system does not actually need 100 trials before it can start calculating; it can use whatever information it has accumulated so far (e.g., 12 instances if it is on trial 13). The system can use *up to* 100 trials, but does not necessarily need that many. In this sense, a rapid learning rate is not particularly diagnostic in discriminating between accounts stating that the system can handle, for instance, 100 versus just 10 trials of information. As explained below, the unlearning manipulation reported here is much more informative in this respect.

The second goal of Experiment 1, therefore, is to investigate the rate of *unlearning*. Partway through the experiment, contingencies were suddenly and without notice switched from 67% (in a three-choice task) to 33% (chance; i.e., each word is presented equally often in all colours). The questions being investigated are whether the colour-word contingency effect is eliminated, and if so, how fast? One possibility is that participants discover the statistical regularities early on in the task and stop searching for contingencies. If so, then the contingency effect should not be extinguished by changing the probabilities. More likely, the effect will extinguish, but the rate at which this happens is diagnostic for some of the competing accounts.

The repetition account assumes that the effect results from recent exposure to other similar trials and thus predicts rapid unlearning. Similarly, a high learning rate connectionist account predicts, by definition, a high learning rate and fast extinction. In contrast, a low learning rate connectionist account predicts, by definition, a low learning rate and slow extinction, which would be reflected by a gradual decrease in the size of the contingency effect across several unlearning blocks.

For the instance account, if the window of trials that participants retrieve for response prediction is large (e.g., the last 100 trials), then the contingency effect should very slowly extinguish as participants are exposed to more and more uncorrelated trials. This is because it will take a great deal of unlearning before the average contingency of the last 100 trials is substantially reduced (e.g., on the 21st trial of unlearning, 80% of the trials the system is using are still from the learning phase). This slow unlearning would be reflected by a gradual decrease in the size of the contingency effect across several unlearning blocks (just like the low learning rate connectionist prediction).

For an instance account that posits that the system relies on a limited number of the most recently encountered trials, the effect should extinguishing very rapidly, perhaps in the first block of changed probabilities. For instance, if the system makes its calculations based on just the last 10 trials, then by trial 11 the participant is not using a single trial from the learning phase to generate response expectancies. Thus, for the instance account, both a large window and small window version can accommodate fast learning, but only the small window account predicts fast extinction when unlearning.

In summary, Experiment 1 investigates the rate of initial learning of contingency information and subsequent unlearning. The experiment begins with three short blocks of 18 trials in which there is a 67% contingency. Learning across blocks is analysed to assess acquisition speed. Directly following these three learning blocks were nine unlearning blocks of 18 trials each in which the contingencies were dropped to chance (33%, three choice). The decrease in the size of the contingency effect across unlearning blocks is assessed to determine extinction speed.

The repetition account predicts rapid learning and unlearning. For connectionist accounts, if the learning rate is high, then the contingency effect should emerge rapidly in learning and extinguish rapidly in unlearning. If the learning rate is low, then the contingency effect should emerge gradually in learning and extinguish gradually in unlearning. Finally, for instance accounts, learning could possibly occur rapidly regardless of window size. Unlearning speed will depend on the number of trials the system is able to use to generate response expectancies.

2.1. Method

2.1.1. Participants

Ninety-eight University of Waterloo undergraduates participated in Experiment 1 in exchange for course credit.

2.1.2. Apparatus

Stimulus and response timing were controlled by E-Prime (Experimental Software Tools, 2002). Participants pressed the "j" key for blue, the "k" key for red, and the "l" key for green with the first three fingers of their right hand.

2.1.3. Materials and design

Participants sat approximately 60 cm from the screen and viewed stimuli on a black screen. There were four stimulus words (LOCK, WIDE, REST, CRAM), but any given participant only saw three of these.[1] There were three display colours (blue, red, green). The experiment began with three learning blocks of 18 trials each. In each learning block, each of the three words was presented four out of six times (67%) in a randomly assigned colour (e.g., LOCK in blue, WIDE in red, REST in green) and once in each of the remaining colours (e.g., LOCK would be presented four times in blue, once in red, and once in green). Directly following these three learning blocks there were nine unlearning blocks, again of 18 trials each. In each unlearning block, each of the three words was now presented equally often (two out of six times) in each of the three colours. Participants were not notified of or told to expect the switch from learning to unlearning. Stimuli were presented in lowercase, bold, 18 pt. Courier New font. Stimuli within blocks were presented in random order.

2.1.4. Procedure

At the beginning of each trial participants saw a white fixation cross for 250 ms, followed by a blank screen for 250 ms, followed by the coloured word for 2000 ms or until a response was made. A blank screen was presented for 300 ms following a correct response, and the message "Incorrect" or "No response" was presented in red for 1000 ms following an incorrect or missed response, respectively.

2.2. Results

Trials in which participants failed to respond were deleted from analyses (less than 1% of the data). For response latencies, only correct responses were analysed. For each participant in each cell, response latencies that were more than 2.5 standard deviations above or below the mean were excluded from analysis (approximately 1% of the data). Other than reducing noise, these exclusion criteria do not affect the pattern of the results.[2]

2.2.1. Response latencies

A 2 (contingency; high, low) × 12 (block) ANOVA for response latencies yielded a significant main effect of contingency, $F(1, 97) = 6.794$, $MSE = 6112$, $p = .011$, a main effect of block, $F(11, 1067) = 3.179$, $MSE = 11788$, $p < .001$, and an interaction between these two factors, $F(11, 1067) = 4.736$, $MSE = 5647$, $p < .001$. Planned comparisons were conducted to determine which blocks yielded a significant contingency effect. The data and statistics are presented in Table 1. Comparisons revealed significant and relatively consistent contingency effects for all three learning blocks (Blocks 1–3). There was also a significant (but small) contingency effect in the first unlearning block immediately following learning (Block 4). For the following seven blocks (Blocks 5–11), there were no significant contingency effects and the differences were all close to zero. Unexpectedly, high contingency trials were significantly *slower* than low contingency trials in the final block (Block 12). However, given the number of statistical tests conducted and the fact that this difference is in the wrong direction for a contingency effect, this finding is likely a Type I error. Indeed, this effect is no longer significant after a Bonferroni correction (Block 4 falls below significance with this correction as well).

2.2.2. Error percentages

The error data are presented in Table 2. A 2 (contingency) × 12 (block) ANOVA revealed a significant main effect of block, $F(11, 1067) = 1.857$, $MSE = 62$, $p = .041$, but no main effect of contingency, $F(1, 97) = 2.561$, $MSE = 65$, $p = .113$, nor an interaction, $F(11, 1067) = 1.433$, $MSE = 66$, $p = .152$. These data are not discussed further.

---

[1] This was a programming error. Four words (rather than three) were randomly assigned to an array of size four for each participant. The program only needed three words and only referenced the first three positions of this array. Thus, whichever word was assigned to the forth position of the array for a given participant was simply never referenced and never presented to the participant. Note that this in no way confounds our results.

[2] Unlike Experiment 2 to follow, immediate repetition trials were not trimmed in this experiment (i.e., trials in which the preceding trial had the same word and/or colour). The reason that this is a particularly important trimming procedure is because complete repetition trials (i.e., trials in which both the word and the colour are repeated) are responded to very quickly and these trials are disproportionately represented in the high contingency condition. In fact, due to the blocked structure of the task, the only way it is possible to have a complete repetition in the low contingency condition is for the last trial of one block to match the first trial of the next block. We opted not to perform this trimming procedure in Experiment 1 for two reasons. First, there were already so few observations per cell (in fact, only 10 of the 98 participants had an observation left in every cell after this trim). Second, sequential effects do not confound analyses in the unlearning blocks, given that complete repetitions are no longer disproportionately represented in the high contingency cells. Moreover, analyses with repetition trials removed yield similar (howbeit substantially noisier) results. The same is true of Experiment 3.

Table 1
Experiment 1 response latencies (in milliseconds) and statistical comparisons for block and contingency.

	Contingency			Statistic
	High	Low	Effect	
Learning				
Block 1	593	638	45	$t(97) = 3.697$, $SE_{diff} = 12$, $p < .001$**
Block 2	567	604	37	$t(97) = 3.004$, $SE_{diff} = 12$, $p = .003$**
Block 3	540	585	45	$t(97) = 4.524$, $SE_{diff} = 10$, $p < .001$**
Unlearning				
Block 4	563	586	23	$t(97) = 2.186$, $SE_{diff} = 11$, $p = .031$*
Block 5	579	571	−8	$t(97) = .721$, $SE_{diff} = 11$, $p = .473$
Block 6	578	578	0	$t(97) = .039$, $SE_{diff} = 11$, $p = .969$
Block 7	566	569	3	$t(97) = .336$, $SE_{diff} = 9$, $p = .715$
Block 8	590	583	−7	$t(97) = .658$, $SE_{diff} = 10$, $p = .512$
Block 9	584	585	1	$t(97) = .118$, $SE_{diff} = 12$, $p = .906$
Block 10	579	580	1	$t(97) = .105$, $SE_{diff} = 10$, $p = .916$
Block 11	606	588	−18	$t(97) = 1.455$, $SE_{diff} = 12$, $p = .149$
Block 12	601	578	−23	$t(97) = 2.425$, $SE_{diff} = 9$, $p = .017$*

* $p < .05$.
** $p < .004$ (Bonferroni correction).

Table 2
Experiment 1 percentage errors for block and contingency.

	Contingency		
	High	Low	Effect
Learning			
Block 1	5.8	9.1	3.3
Block 2	5.3	6.6	1.3
Block 3	4.7	5.9	1.2
Unlearning			
Block 4	4.4	6.4	2.0
Block 5	3.6	4.9	1.3
Block 6	5.2	5.2	0.1
Block 7	5.9	5.6	−0.3
Block 8	6.3	4.4	−2.0
Block 9	4.7	5.3	0.7
Block 10	5.7	5.1	−0.6
Block 11	5.4	4.8	−0.4
Block 12	6.1	5.9	−0.2

2.3. Discussion

The results of this experiment clearly demonstrate that both learning and unlearning occur extremely rapidly. Initial contingency learning was significant in the very first block of 18 trials. Unlearning seems to occur just as rapidly. There was only a very small carryover from the learning blocks into the first unlearning block, and then the effect disappeared in the following unlearning blocks. Thus, it is clear that the learning mechanism is highly responsive to the actual contingencies. This rules out a few of the accounts considered in the Introduction. The data are consistent with connectionist accounts, but only if a high learning rate is assumed. With a low learning rate, it would take the system much longer to accrue enough information to learn contingencies in the learning phase and it would take substantially more unlearning for the effect to extinguish. Similarly, the finding of rapid extinction rules out an instance account in which it is assumed that the system draws on a relatively large sample of trial memories. Fast learning and unlearning, however, is consistent with a small window instance account. Finally, the repetition account posits that the colour-word contingency effect results from transient repetition effects and is thus consistent with the observed rate of learning and unlearning.

3. Reanalysis 1

The repetition account of the colour-word contingency effect, as noted above, attributes the effect to either residual activation or temporary SR associations occurring more often for high contingency trials than low contingency trials. The earliest experiments we conducted using the colour-word contingency paradigm had constraints on presentation order such that no colour could be repeated from one trial to the next, thus making it impossible for such complete repetitions (e.g., $\text{MOVE}_{\text{blue}}$

could never directly follow $MOVE_{blue}$; Schmidt et al., 2007) and we have also been careful to control for $n-1$ sequence effects wherever we had enough data per cell to do so (i.e., by deleting trials in which the colour repeats, thus eliminating complete repetitions, which are faster than other trials). Thus, we can already rule out an account that holds that colour-word contingency learning results from trial $n-1$ repetition effects. However, these controls have not ruled out sequence effects beyond trial $n-1$. For instance, it may be the case that complete repetitions on trial $n-2$, $n-3$, or perhaps even more distant lags also produce a speeding of responses. Thus, the contingency effect could simply be the result of the combination of benefits from various lags. We therefore conducted a reanalysis of data from Schmidt et al. (2007, Experiment 2) to test for $n-2$ through $n-5$ repetition effects. The critical test condition is *complete repetitions*, where both the word and colour repeat. We also coded for *word repetitions*, where the word but not the colour repeats; *colour repetitions*, where the colour but not the word repeats; and *alternations*, where neither the word nor colour repeats. The reason for selecting this particular experiment for our reanalysis is that the contingency manipulation was small enough (50% in a four choice task, where chance is 25%) to allow sufficient observations in all cells (e.g., in experiments with larger contingency manipulations where each low contingency pairing only occurs once per block, as was the case in Experiment 1 here, the only way to get a complete repetition is for the last trial in one block to match the first trial in the next block).

The predictions of connectionist and instance accounts are less clear than those of the repetition account. One might argue that a connectionist model with a high enough learning rate should predict a larger influence of more recent trials (given that each new trial needs to be able to have a significant influence on connection weights). In that sense, repetition effects might be expected. However, even with high learning rates there is still an accumulation over several trials. For the instance account there is no a priori reason to expect that the most recent events should (or should not) have a greater influence on responding than later trials within the window of trials that the system takes into account. It is our suggestion that it is *not* the case that each individual retrieved instance biases its associated response; rather, the idea is that participants are retrieving a number of associated instance memories and determining the likely (i.e., most frequent) response based on these.

3.1. Method

A full description of the methodology for the experiment used in this reanalysis can be found in the original article (Schmidt et al., 2007, Experiment 2). The study was very similar to Experiment 1 here. Participants were 16 University of Saskatchewan undergraduates. The task was four choice rather than three. In each block, each of four words was presented 6 out of 12 times (50%) in a randomly assigned colour and twice in the remaining colours in each of eight blocks. There was a constraint on presentation order such that the display colour could not repeat from one trial to the next. Trials were recoded for both contingency and for repetition type at four lags ($n-2$, $n-3$, $n-4$, and $n-5$). Complete repetitions were trials in which both the word and colour repeated. Word repetitions were trials in which only the word repeated. Colour repetitions were trials in which only the colour repeated. Finally, alternations were trials in which neither the word nor the colour repeated.

3.2. Results

There were very few errors in the experiment used for this and the following reanalysis (in fact, the average participant made about seven errors total, less than the number of conditions used in the following analyses). We therefore restrict our analyses to response latencies. Trials on which participants failed to respond (less than 1% of the data) and incorrect responses (less than 4% of the data) were deleted. These trimming procedures do not alter the basic pattern of data reported below. The data are presented in Table 3.

3.2.1. Trial n-2

A 2 (contingency; high, low) × 4 (repetition type; complete repetition, word repetition, colour repetition, alternation) ANOVA for response latencies revealed a marginal main effect for contingency, $F(1, 15) = 3.178$, $MSE = 2587$, $p = .095$. Critically, there was no main effect of repetition type, $F(3, 45) = 1.871$, $MSE = 2383$, $p = .148$, nor an interaction, $F(3, 45) = 1.453$, $MSE = 1453$, $p = .240$.

3.2.2. Trial n-3

A 2 (contingency; high, low) × 4 (repetition type; complete repetition, word repetition, colour repetition, alternation) ANOVA for response latencies revealed a significant main effect for contingency, $F(1, 15) = 8.624$, $MSE = 3813$, $p = .010$. Again, there was no main effect of repetition type, $F(3, 45) = .465$, $MSE = 2905$, $p = .708$, nor an interaction, $F(3, 45) = .375$, $MSE = 4504$, $p = .772$.

3.2.3. Trial n-4

A 2 (contingency; high, low) × 4 (repetition type; complete repetition, word repetition, colour repetition, alternation) ANOVA for response latencies revealed a marginal main effect for contingency, $F(1, 15) = 3.180$, $MSE = 7190$, $p = .095$. There was no main effect of repetition type, $F(3, 45) = .006$, $MSE = 6370$, $p = .999$, nor an interaction, $F(3, 45) = .510$, $MSE = 5669$, $p = .677$.

Table 3
Reanalysis 1 response latencies (in milliseconds) for lag, repetition type, and contingency.

	Contingency	
	High	Low
Trial n − 2		
Complete repetition	713	719
Word repetition	705	738
Colour repetition	741	740
Alternation	703	729
Trial n − 3		
Complete repetition	708	759
Word repetition	709	732
Colour repetition	711	747
Alternation	711	730
Trial n − 4		
Complete repetition	701	750
Word repetition	722	725
Colour repetition	709	740
Alternation	711	735
Trial n − 5		
Complete repetition	690	709
Word repetition	712	727
Colour repetition	709	727
Alternation	715	741

3.2.4. Trial n-5

A 2 (contingency; high, low) × 4 (repetition type; complete repetition, word repetition, colour repetition, alternation) ANOVA for response latencies revealed a significant main effect for contingency, $F(1, 15) = 5.128$, $MSE = 2324$, $p = .039$. There was no main effect of repetition type, $F(3, 45) = 1.868$, $MSE = 2499$, $p = .149$, nor an interaction, $F(3, 45) = .070$, $MSE = 2089$, $p = .976$.

3.3. Discussion

The results of Reanalysis 1 show no evidence for repetition effects at lags of two to five trials. For each of these four lags, no effect of repetition type emerged. These null findings are problematic for the repetition account, which purports to explain the contingency effect solely by the influence of these transient repetition effects. Of course, interpreting the null is always difficult. One might argue that we merely lacked statistical power to detect these lag effects. However, there is a way to demonstrate that, in fact, lag effects do not explain the contingency effect. For this we turn to Reanalysis 2.

4. Reanalysis 2

Reanalysis 1 indicated no evidence for $n - 2$ through $n - 5$ repetition effects. Rather than simply failing to reject this null, we conduct a further analysis to demonstrate that these (absent) lag effects do not account for the contingency effect. Recall that the repetition account purports to *fully explain* the contingency effect in terms of these short-lived associations or activations. Thus, the argument is not *only* that there should be observable lag repetition effects, but also that these repetition effects should explain the variance attributed to the contingency effect. In other words, after accounting for the variance attributed to these repetition effects, there should be no variance left over for the contingency manipulation to explain (i.e., because repetition effects *are* the contingency effect in this conceptualisation). Thus, if the repetition variables are entered into the first step of a regression analysis and then contingency is added to the regression analysis in a second step, then the new regression model with contingency included should *not* explain more variance. If more variance *is* explained by contingency, then this verifies that our initial analyses were not simply the result of poor statistical power. Instead, transient repetition effects do not fully explain the contingency effect. The reader is again reminded that $n - 1$ repetition effects were controlled by design (i.e., colour repetitions were impossible), so only lags $n - 2$ and beyond need to be entered into the regression.

4.1. Method

The same data set used for Reanalysis 1 was used for Reanalysis 2. For this analysis, the full raw data set was dummy coded for participant, contingency, and the repetition type at each lag. That is, each individual trial for each participant was included as an observation in the regression and then participant number was included as a predictor in the regression along with contingency, repetition type, and lag (for an explanation of how to do regression with repeated observations per participants see Bland & Altman, 1995).

4.2. Results

Null and incorrect responses were trimmed (as in the previous analysis). These trimming procedures do not alter the basic pattern of data reported below.

4.2.1. Step 1 – participant, repetition type, and lag

In Step 1 of the regression, the dummy coded variables for participants and for repetition trial types at the various lags were entered as predictors and response latency was entered as the outcome variable. Unsurprisingly, this model explained a significant amount of variance, $R^2 = .256$, $F(27, 5896) = 75.262$, $p < .001$. Note that this model explains the variance *between* participants (i.e., the multiple observations per participant were coded for participant number and instead of removing this variance, as in a traditional regression, between-participant variance was included as a predictor).

4.2.2. Step 2 – adding contingency

In Step 2 of the regression, all of the variables in Step 1 were included plus the new variable for contingency (high, low). The test for a change in the amount of variance explained was significant, R^2 *Change* = .001, *F Change* (1, 5895) = 11.018, $p = .001$. Note that the reason for the small value of the R^2 Change is that the between-participant differences account for an enormous chunk of the variation (accounted for in Step 1 of the regression). Within the full model, contingency accounts for 19 ms of variance.

4.3. Discussion

The results of Reanalysis 2 corroborate the findings of Reanalysis 1 by showing that (the non-existent) repetition effects at lags of two to five trials do not explain the contingency effect. After putting all of these repetition variables into the first step of a regression to account for what variation they could, contingency continued to explain variance in the second step of the regression. Note again that this experiment, by design, rules out $n - 1$ repetition effects due to the constraint on presentation order (i.e., colour repetitions were impossible). As a result of this analysis, it is safe to conclude that the colour-word contingency effect reflects more than simple priming by transient activations or SR associations as posited by the repetition account, at least as far out as five trials.

The implication of these two reanalyses for connectionist and instance accounts is less certain. One might have expected some repetition effects at recent lags for a high learning rate connectionist account, but the argument probably cannot be made that such lag repetition effects should have completely accounted for the contingency effect. No strong prediction was made for the instance account.

5. Experiment 2

Given how rapid learning and unlearning were in Experiment 1, it is clear that the "window" of trials that participants take into account when calculating their response prediction is remarkably small. This led us to the notion that participants may be using limited-capacity memory resources to retrieve a small number of recently encountered trial memories in preparing a response. This is consistent with the finding from the sequence learning literature that carrying a memory load impairs learning (Nissen & Bullemer, 1987), though it is not clear that learning between trials is necessarily always the same as learning within trials (see the General Discussion for a discussion of the similarities and differences between the colour-word contingency paradigm and several other paradigms).

Experiment 2 tests this memory resource hypothesis by examining the impact of memory load on the colour-word contingency effect. Participants in one condition were given a set of five digits to remember at the beginning of each trial and were tested for their recognition at the end of each trial. Forcing participants to remember five digits should create a high load on memory, which leaves little or no memory resources to retrieve trial information that can be used to learn contingencies. Other participants were given two digits to remember. Thus, there is a low load on memory, which ought to enable participants to use their remaining memory resources for learning contingencies. Thus, a contingency effect is expected in the low load condition, where a smaller (or possibly null) effect is expected in the high load condition.

These predictions are largely derived from our preferred instance account of the colour-word contingency effect in which it is assumed that the system needs to perform memory retrieval functions on a trial-by-trial basis to facilitate responding to high contingency trials. It is probably the case that connectionist models can be modified to allow a role for limited-capacity memory resources, as well. It is less clear, however, why a memory load should interfere with residual activation or temporary SR bindings, so memory load effects would seem inconsistent with a repetition account.

5.1. Method

5.1.1. Participants

Sixty University of Waterloo undergraduates participated in Experiment 2 in exchange for course credit. None had participated in the previous experiment. Two participants were deleted from the high load condition and two from the low load condition for having less than 70% accuracy on the memory task, leaving 28 participants in each of the high and low load conditions.

5.1.2. Apparatus

The apparatus for Experiment 2 was identical to Experiment 1 with one exception. In addition to the "j," "k," and "l" keys that were pressed with the right hand to respond to colours, participants used their left hand to press the "y" key for "yes" responses and the "n" key for "no" responses in regard to the load manipulation.

5.1.3. Materials and design

The materials and design for Experiment 2 were identical to Experiment 1 with the following exceptions. There were only three stimulus words (LOCK, WIDE, REST). At the beginning of each trial, participants were presented with either five (high load) or two (low load) random digits (0–9) horizontally presented with three spaces between each digit. Following a response to the target colour on each trial, participants were presented with a second set of digits. For both groups of participants, there were two blocks of 60 trials each. In each block, a randomly selected digit in the memory set was changed to a new random digit on half of the trials and none of the digits changed on the other half of the trials. Orthogonal to this, each of the three words was presented eight out of ten times (80%) in an assigned colour and once in each of the remaining colours (e.g., LOCK 80% in blue). It is expected that with five digits to remember memory will be too highly loaded to learn contingencies, whereas with only two digits to remember memory is not heavily loaded and will have leftover resources for storing trial information to learn contingencies. As a result, a contingency effect is expected in the low load condition, but not the high load condition.

5.1.4. Procedure

At the beginning of each trial participants saw a white fixation cross for 250 ms, followed by a digit memory set for 2000 ms. Participants were instructed to remember these digits in order. Next, there was a blank screen for 250 ms, followed by the coloured word for 2000 ms or until a response was made. The message "Correct," "Incorrect," or "No response" was presented in white for 500 ms following correct, incorrect, and null responses, respectively. A second set of digits was then presented until participants decided whether one of the digits had changed by pressing the "y" key (for "yes") or the "n" key (for "no"). This was followed by a second feedback screen, which was identical to the first (except that null responses were impossible).

5.2. Results

Null responses were deleted (less than 3% of the data), as were trials in which participants failed on the memory test (about 11% and 8% of the data in the high and low load conditions, respectively). Because we were interested in trial n contingency effects and not sequential effects all trials where the word or colour was the same as that on the preceding trial were deleted. For response latencies, only correct responses were analysed. In addition, for each participant in each cell, response latencies that were more than 2.5 standard deviations above or below the mean were excluded from analysis (approximately 2% of the data). These trimming procedures do not alter the basic pattern of data reported below.

5.2.1. Response latencies

The response latencies for Experiment 2 are presented in Table 4. A 2 (contingency; high, low) × 2 (memory load; high, low) ANOVA for response latencies yielded a significant main effect of contingency, $F(1, 54) = 16.921$, $MSE = 7611$, $p < .001$, and an interaction, $F(1, 54) = 5.667$, $MSE = 7611$, $p = .021$, in which there was a larger contingency effect for the low relative to the high load group. The main effect of memory load was not significant, $F(1, 54) = .453$, $MSE = 47878$, $p = .504$. Planned comparisons revealed that participants in the low load group responded faster to high contingency trials (779 ms) than to low contingency trials (886 ms), $t(27) = 4.055$, $SE_{diff} = 26$, $p < .001$. In contrast, participants in the high load group did not respond significantly faster to high contingency trials (846 ms) than to low contingency trials (874 ms), $t(27) = 1.446$, $SE_{diff} = 20$, $p = .160$.

5.2.2. Error percentages

Percentage errors for Experiment 2 are presented in Table 5. A 2 (contingency) × 2 (memory load) ANOVA for error percentages was conducted. The main effect of contingency, $F(1, 54) = .219$, $MSE = 44$, $p = .642$, the main effect of memory load,

Table 4
Experiment 2 response latencies (in milliseconds) for contingency and load.

	Contingency		
	High	Low	Effect
Low load	779	886	107
High load	846	874	28

Table 5
Experiment 2 percentage errors for contingency and load.

	Contingency		
	High	Low	Effect
Low load	4.5	4.1	−0.4
High load	5.4	6.9	1.5

$F(1, 54) = 1.263$, $MSE = 72$, $p = .266$, and the interaction, $F(1, 54) = .565$, $MSE = 44$, $p = .455$, were not significant. Planned comparisons revealed no significant differences in errors between high and low contingency trials for participants in the low load group (4.5% and 4.1%, respectively), $t(27) = 0.216$, $SE_{diff} = 1.7$, $p = .830$, or in the high load group (5.4 and 6.9%), $t(27) = .808$, $SE_{diff} = 1.9$, $p = .426$.

5.3. Discussion

The results of Experiment 2 demonstrate quite dramatically that participants in the high memory load condition did *not* show a contingency effect (or at least the effect was significantly attenuated), whereas those participants in the low memory load condition *did* show a contingency effect. This is consistent with the idea that limited-capacity memory resources of some sort are required for colour-word contingency learning. Specifically, the argument is that when memory resources are taxed with a secondary task, there are no (or less) resources left over to store and/or retrieve instances that can be used to learn contingencies. The system requires memory resources to be free in order for instances to be stored and contingency information to be learned. We have nothing to say about the nature of these resources at the present time. Future research will be needed to answer such as questions as whether these resources are domain general or domain specific. We again note that connectionist models can likely be modified to allow a role for limited-capacity resources, as well. It is less clear how the repetition account could accommodate these findings.

These results are also consistent with the finding of Nissen and Bullemer (1987), where load was shown to prevent sequential learning. Some authors, however, have argued that the apparent memory load effects observed in sequential learning may actually be due to an integration of the secondary task stimuli into the memory of the sequence (Heuer & Schmidtke, 1996; Schmidtke & Heuer, 1997). For instance, if a participant has to monitor tones presented between target stimuli, then these tones could, according to this account, become part of the sequence (i.e., *Stimulus* 1 → *random tone* → *Stimulus* 2 instead of *Stimulus* 1 → *Stimulus* 2), which interferes with learning by separating stimuli in the sequence rather than by load per se. Whether this is the full explanation of load effects in sequence learning is up for debate, but in our paradigm this is less of an issue. Learning in the colour-word contingency learning paradigm by definition occurs within trials rather than between. Thus, even if the secondary task information is somehow being encoded along with the target, distracter, and response, failure to infer the distracter–response contingency is still an issue of load (i.e., too much information loaded on the system to extract the important contingency).

6. Experiment 3

The results of Experiment 2 leave several unanswered questions about the *specific* role of memory resources in contingency learning. One possibility is that memory resources are required for the binding of features and responses into instances. We term this the *encoding hypothesis*. That is to say, participants need memory resources in order to initially *make and store* instances. Thus, if memory resources are taxed by a difficult enough secondary task, then instances will not be recorded and there will, resultantly, be no instances (or perhaps incomplete instances) to retrieve to use to determine the high contingency response. If this view is correct, then it is *not* simply the case that participants are not showing a contingency effect while under load; rather, participants have not learned anything about the contingencies in the task.

A second possibility is that participants *are* able to create and store instances while under a memory load, but they are *unable* to retrieve these instances while under load. We term this the *retrieval hypothesis*. In this sense, participants put under memory load are learning contingency information, but are simply unable to use this learning in the presence of the secondary task.

A third possibility is that participants require memory resources *both* for the creation of instances *and* for the subsequent retrieval of these instances. We term this the *resource hypothesis*. According to this hypothesis, memory resources are needed more broadly to carry out the various memory functions required for contingency learning. Thus, memory load, according to this hypothesis, impairs both encoding *and* subsequent retrieval processes.

To test these various accounts, two groups of participants were included in Experiment 3. Both groups underwent an initial Learning Block (36 trials) in which contingencies were introduced, followed by a Transfer Block (36 trials) in which contingencies were removed. The critical test block in Experiment 3, as discussed below, is the Transfer Block. Note that although unlearning is rapid when contingencies are removed, transfer was observed in the initial unlearning block in

Experiment 1. For Group 1, memory load was high for learning and low for transfer. For Group 2, memory load was low for learning and high for transfer. As described below, a control experiment was also run that was identical except that memory load was low for both learning *and* transfer.

Participants in Group 1 were put under high load in the Learning Block and low load in the Transfer Block. If the *encoding* hypothesis is correct (i.e., memory resources are required for the *creation* of instances), then participants will not learn contingencies while loaded in the Learning Block, leading to no transfer in the subsequent Transfer Block when the memory load is removed. Alternatively, if the *retrieval* hypothesis is correct and participants *are* storing instances while under load in the Learning Block (but are simply not able to retrieve and use them while under load), then there should be a transfer of learning in the Transfer Block when the load is removed (i.e., a significant contingency effect). If memory resources are required for encoding *and* retrieval (the resource hypothesis), then no transfer should be observed.

Participants in Group 2 were put under low load in the Learning Block and high load in the Transfer Block. If the retrieval hypothesis is correct, then contingency knowledge can only be used when sufficient resources are available to retrieve instances. Thus, participants will successfully learn contingencies in the Learning Block under a low load, but will not show an effect in the Transfer Block when a high load is introduced. Alternatively, if the *encoding* hypothesis is correct and memory resources are needed for initial encoding of instances, then participants should learn contingencies in the Learning Block with low load and continue to show a contingency effect in the Transfer Block after a high load has been introduced. In other words, according to the encoding hypothesis it does *not* matter if memory is *currently* loaded, so long as contingency information has been learned. Lastly, if memory resources are required for both encoding *and* retrieval (the resource hypothesis), then no transfer should be observed.

To summarise, the encoding hypothesis predicts that contingency effects will be observed when participants are not highly loaded while learning, thus predicting transfer in Group 2, but not in Group 1. The retrieval hypothesis predicts that contingency effects will be observed when participants are *currently* not highly loaded (i.e., when they are able to retrieve instances), thus predicting transfer in Group 1 but not in Group 2. Finally, the resource hypothesis predicts that both encoding *and* retrieval cannot be accomplished under load, thus predicting no transfer in either of the two groups. Given the latter possibility, a control experiment was also conducted to ensure that transfer can occur within the specific parameters used in this experiment. The control experiment was identical to the main experiment save for the fact that memory load was low in both the learning and transfer blocks. It is unclear what a connectionist account should predict, as it would first have to be determined how such an account would include a role for memory resources.

6.1. Method

6.1.1. Participants

Eighty University of Waterloo undergraduates participated in Experiment 3 in exchange for course credit, with 40 in each of the two groups. Seven participants in Group 1 and seven participants in Group 2 were deleted due to less than 70% accuracy on the memory task, leaving 33 participants per group. Another 33 participants from the same participant pool were in the control experiment. One participant was deleted due to less than 70% accuracy on the memory task, leaving 32 participants. None of the participants had participated in any of the previous experiments.

6.1.2. Apparatus

The apparatus for Experiment 3 was identical to Experiment 2.

6.1.3. Materials and design

The materials and design for Experiment 3 were identical to Experiment 2 with the following exceptions. For both groups of participants, there were two blocks of 36 trials each. In the initial Learning Block, each of the three words was presented 8 out of 12 times (67%) in an assigned colour and once in each of the remaining colours. In the subsequent Transfer Block, each of the three words was presented 4 out of 12 times in each colour (33%, chance). Orthogonal to this, a randomly selected digit in the memory set was changed to a new random digit on half of the trials and none of the digits changed on the other half of the trials. For one group of participants (Group 1), load was high (five items) in the Learning Block and low (two items) in the Transfer Block. For the other half of the participants (Group 2), load was low in the Learning Block and high in the Transfer Block. Participants were counterbalanced across groups. In the control experiment, load was low for both blocks. The critical question of interest is which groups of participants show transfer.

6.1.4. Procedure

The procedure for Experiment 3 was identical in all respects to Experiment 2.

6.2. Results

Trials on which participants failed to respond (less than 1% of the data) and trials on which participants made an error on the memory task (approximately 14% of the data) were removed. Correct response latencies were trimmed by removing trials for each participant in each cell that were over 2.5 standard deviations from the mean (less than 2% of the data). These trimming procedures do not affect the basic pattern of results described below.

Table 6
Experiment 3 response latencies (in milliseconds) for group, block, and contingency.

	Contingency		
	High	Low	Effect
Control			
Learning Block (low)	891	930	39*
Transfer Block (low)	788	815	27*
Group 1			
Learning Block (High)	1015	1032	17
Transfer Block (low)	860	839	−21
Group 2			
Learning Block (low)	924	983	59*
Transfer Block (high)	900	897	−3

* $p < .05$.

6.2.1. Control: low load learning – low load transfer

Participants in the control experiment were given 67% contingencies to learn under low load in the Learning Block and then were presented with chance 33% contingencies under low load in the Transfer Block in order to ensure transfer was possible in the task.

6.2.1.1. Response latencies. Response latency data for the control experiment are presented in Table 6. A *t*-test on the Learning Block revealed that high contingency trials (891 ms) were responded to significantly faster than low contingency trials (930 ms), $t(31) = 2.759$, $SE_{diff} = 14$, $p = .010$. Critically, a *t*-test on the Transfer Block revealed a significant transfer effect; high contingency trials (788 ms) were responded to significantly faster than low contingency trials (815 ms), $t(31) = 2.393$, $SE_{diff} = 11$, $p = .023$. Thus, transfer can be observed in this version of the paradigm.

6.2.1.2. Percentage error. Percentage error data for the control experiment are presented in Table 7. A *t*-test on the Learning Block control data revealed that high contingency trials (3.8%) did not generate significantly different errors than low contingency trials (3.1%), $t(31) = .532$, $SE_{diff} = 1.1$, $p = .599$. Additionally, a *t*-test on the Transfer Block revealed no significant difference between high contingency trials (3.1%) and low contingency trials (4.3%), $t(31) = .847$, $SE_{diff} = 1.4$, $p = .403$.

6.2.2. Group 1: high load learning – low load transfer

The first group of participants were given 67% contingencies to learn under high load in the Learning Block and then were presented with chance (33%) contingencies under low load in the Transfer Block.

6.2.2.1. Response latencies. Response latencies for Group 1 are presented in Table 6. A *t*-test on the Learning Block revealed that high contingency trials (1015 ms) were not responded to significantly faster than low contingency trials (1032 ms), $t(32) = .801$, $SE_{diff} = 23$, $p = .429$. Critically, a *t*-test on the Transfer Block revealed no significant transfer effect; high contingency trials (860 ms) were not responded to faster than low contingency trials (839 ms), $t(32) = 1.131$, $SE_{diff} = 19$, $p = .267$. Note that the numbers were numerically in the wrong direction. Additionally, this null effect was significantly smaller than the transfer effect in the control experiment, $F(1, 63) = 4.726$, $MSE = 3928$, $p = .033$. Thus, there was no evidence for the hypothesis that participants can learn under load.

6.2.2.2. Percentage error. The percentage error data for Group 1 are presented in Table 7. A *t*-test on the Learning Block revealed that high contingency trials (3.6%) did not generate significantly different errors than low contingency trials (5.4%),

Table 7
Experiment 3 percentage errors for group, block, and contingency.

	Contingency		
	High	Low	Effect
Control			
Learning Block (low)	3.8	3.1	−0.7
Transfer Block (low)	3.1	4.3	1.2
Group 1			
Learning Block (high)	3.6	5.4	1.8
Transfer Block (low)	2.2	2.8	0.6
Group 2			
Learning Block (low)	2.6	7.4	4.8*
Transfer Block (high)	3.5	3.0	−0.5

* $p < .05$.

$t(32) = 1.034$, $SE_{diff} = 1.7$, $p = .309$. Additionally, a t-test on the Transfer Block revealed no significant transfer effect; high contingency trials (2.2%) did not generate significantly different errors than low contingency trials (2.8%), $t(32) = .717$, $SE_{diff} = 0.9$, $p = .479$.

6.2.3. Group 2: low load learning – high load transfer

The second group of participants were given 67% contingencies to learn under low load in the Learning Block and then were presented with chance (33%) contingencies under high load in the Transfer Block.

6.2.3.1. Response latencies. Response latencies for Group 2 are presented in Table 6. A t-test on the Learning Block revealed that high contingency trials (924 ms) were responded to significantly faster than low contingency trials (983 ms), $t(32) = 3.013$, $SE_{diff} = 20$, $p = .005$. Critically, a t-test on the Transfer Block revealed no significant transfer effect; high contingency trials (900 ms) were not responded to faster than low contingency trials (897 ms), $t(32) = .159$, $SE_{diff} = 18$, $p = .875$. Although this null effect was not significantly smaller than the transfer effect in the control experiment, $F(1, 63) = 1.964$, $MSE = 3564$, $p = .166$, note that the numerical difference was again in the wrong direction. Thus, there was no evidence for the hypothesis that participants can retrieve and apply learning while under load.

6.2.3.2. Percentage error. The percentage error data are presented in Table 7. A t-test on the Learning Block revealed that high contingency trials (2.6%) generated significantly less errors than low contingency trials (7.4%), $t(32) = 2.916$, $SE_{diff} = 1.6$, $p = .006$. Additionally, a t-test on the Transfer Block revealed no significant transfer effect; high contingency trials (3.5%) did not generate significantly different errors than low contingency trials (3.0%), $t(32) = .390$, $SE_{diff} = 1.2$, $p = .699$.

6.3. Discussion

The results of Experiment 3 provide support for the resource hypothesis by showing that memory load interferes with both storage and retrieval. Participants in Group 1 were not able to encode instances under high load, as indicated by the lack of transfer in the Transfer Block when the load was reduced. Further, participants in Group 2 were not able to retrieve stored instances in the Transfer Block when put under high load. Data from the control experiment confirm that transfer is observable in this task setup. Thus, the combined results suggest that memory resources are required for both encoding and retrieval, in support of the resource hypothesis. Although the manipulations in this experiment were largely inspired to test various versions of the instance account, it is likely the case that connectionist accounts could be modified to accommodate these findings, as well.

7. General discussion

The results of past work and the experiments and reanalyses presented here help to narrow the range of potential explanations for colour-word contingency learning. The available data suggest that these contingencies are acquired implicitly (Schmidt et al., 2007), that the critical contingency is between the word and the response (Schmidt et al., 2007), that learning and unlearning of contingencies is extremely rapid (Experiment 1), that the effect does not result solely from repetition effects (Reanalyses 1 and 2), and that contingency learning requires limited-capacity memory resources (Experiments 2 and 3) for both storage and retrieval (Experiment 3). Given these criteria, we can begin to piece together a model of learning in this paradigm.

Our favoured account of colour-word contingency learning assumes that participants use instances to represent contingency information. According to this instance hypothesis, on each trial a representation of the stimuli and response that was made are bound into an instance memory. These instances are then stored in an episodic store. On each trial, after the word is processed a number of matching instances are retrieved and a response expectancy is determined. For instance, as the participant processes the word MOVE, they will retrieve a number of instances that are associated with this (the most recently encountered ones being the most accessible) and use these to determine that blue is the most probable response. If blue is the correct response (i.e., a high contingency trial), then responding will be speeded.

The results of the experiments and reanalyses presented here are completely consistent with the instance account. The rapid learning of contingencies in Experiment 1 is consistent, because it will only take a handful of trials for participants to have been exposed to a number of high contingency pairings, while only seeing one or two low contingency pairings. Thus, right from the start, participants should be able to begin using contingency information to speed high contingency responses. In addition, because memory has a limited capacity and only so many instances can be retrieved during target processing, it will only take a small amount of unlearning before participants are no longer retrieving instances from the preceding learning phase (i.e., because the more-recently encountered unlearning trials are more accessible). As such, the rapid unlearning observed in Experiment 1 is also consistent with the instance account. Finally, the results of Experiments 2 and 3 are consistent with the instance account, because participants should need memory resources to carry out the memory functions required to store and subsequently retrieve instances, and memory load impairs these functions.

The rapid learning and unlearning in Experiment 1 are also consistent with the connectionist account, so long as the learning rate is assumed to be high. Presumably, models such as the SRN can be easily modified to allow a role for limited-capacity resources in storage and retrieval processes (though of course a demonstration to that effect would be welcome). Note that the primary difference between the proposed instance account and connectionist models such as the SRN is the way in which learned information is represented. In the SRN, information is distributed across a network of hidden units. In the instance account, trial information is stored in discrete instances. Further research will need to be conducted to distinguish between these two possibilities.

Finally, we were able to rule out a repetition account in Reanalyses 1 and 2 by demonstrating that there were no lag effects that were able to explain the variance attributed to the contingency manipulation.

7.1. Relation to past research

The colour-word contingency learning paradigm shares obvious similarities with numerous other cognitive paradigms. However, these paradigms also differ in a number of ways from the paradigm used in the present studies, including the type of stimuli and responses that are involved in the task, the speed of judgment, and several other factors. Although commonalities surely exist, it remains to be seen which common processes underlie which effects of contingencies on performance. Until this issue is examined further, care should be taken when generalising the conclusions from these studies to contingency learning in other paradigms (and vice versa). In the following sections we discuss the relation of the current paradigm to three other broad categories of paradigms: conflict paradigms (e.g., Stroop, Eriksen flanker), judgement tasks (e.g., evaluative conditioning, hidden covariation detection), and sequential learning.

7.1.1. Conflict paradigms

The one paradigm that most of our colleagues seem to equate with the colour-word contingency learning paradigm is the Stroop task. Nonetheless, of the three types of paradigms discussed here, conflict paradigms such the Stroop task are arguably the *least* similar to the present contingency paradigm. On the surface, the colour-word contingency task is very similar to a Stroop task: participants are presented with coloured words and are asked to ignore the identity of the word and respond to the print colour. However, aside from this surface similarity, it can be argued that the two tasks are in fact quite different.

Conflict paradigms such as Stroop are based on over-trained relations, are partially semantic in nature (e.g., De Houwer, 2003; Risko, Schmidt, & Besner, 2006; Schmidt & Cheesman, 2005), and are driven almost entirely by interference (see MacLeod, 1991 for a review). In contrast, colour-word contingency learning is based on newly-trained covariations, is non-semantic (Schmidt et al., 2007), and is driven entirely by facilitation (Schmidt & Besner, 2008). Thus, the informativeness of data from conflict paradigms for our contingency learning work is questionable.

However, contingency manipulations have been introduced within the context of conflict paradigms. The best example of this for our purposes is the work by Miller (1987; see also Carlson & Flowers, 1996) using a flanker task in which a centrally presented target letter was flanked by two identical distracting letters. Some of these letters were presented most often with a particular response. Similar to the colour-word contingency paradigm, it was found that targets with high contingency flankers were responded to faster than low contingency flankers. Although Miller's work did not investigate learning speed or the effect of memory load, the structural similarity between the two paradigms is so extensive that it is likely that the effects in both paradigms result from similar processes.

7.1.2. Judgement tasks

The colour-word contingency learning paradigm shares similarities with various judgement tasks. For instance, in the hidden covariation paradigm, participants learn the contingencies between facial characteristics and personality characteristics (Lewicki, 1985, 1986; Lewicki, Hill, & Czyzewska, 1997; but see Hendrickx, De Houwer, Baeyens, Eelen, & Van Avermaet, 1997a, 1997b). Similarly, in the evaluative conditioning paradigm participants' liking of objects is altered by being paired with valenced words (see De Houwer et al., 2001 for a review). However, there are also many important differences. For instance, in these judgement paradigms the contingencies are typically 100% (e.g., in hidden covariation detection, facial characteristic X is *always* presented with personality characteristic Y). In colour-word contingency learning, contingencies are less than chance. There are at least two reasons why this difference is interesting. First, it is interesting that participants are able to detect a regularity in a noisy (i.e., non-100% contingency) dataset. Second, it is not certain whether detecting regularities in a noisy versus noiseless dataset involves identical processes (e.g., the latter case may lend itself more to explicit recognition of contingencies and be more prone to strategic influences).

Also, the colour-word contingency task involves speeded responses as the dependent measure, whereas judgement tasks such as evaluative conditioning most often involve a relatively slower judgment response (e.g., a judgement of the valence of an object). Changes in the *rate* of processing do not necessarily imply that the system will reach a different *response*. That is, just because a contingency may help to make a judgement *faster*, it does not follow that the participant will necessarily be any more likely to make a given response (e.g., Stimulus B may cause a participant to select Response B regardless of whether they select the response quickly or slowly). Additionally, response latencies are sometimes used in these judgement tasks (e.g., Lewicki, 1986), but these judgment responses are overall much slower than rapid identification responses, so it remains unclear whether effects occurring in a few hundred milliseconds are simply a "scaled down" version of the effects occurring at a few thousand milliseconds. In particular, the relatively slower judgement responses may include more explicit (rather than implicit) processes.

7.1.3. Sequential learning

Sequential learning is another paradigm that shares many similarities with the colour-word contingency paradigm. In the typical sequential learning paradigm participants are presented with a series of target stimuli to respond to (no distracters) and the stimuli follow a predictive sequence (Nissen & Bullemer, 1987). Most sequence learning research has participants respond to a sequence that is either random or 100% predictive (i.e., the same series of stimuli keep repeating). Learning is determined as the difference in response times between these two conditions. More similar to the colour-word contingency paradigm, some research with sequence learning has been done using probabilistic sequences (i.e., where the next item in the sequence is predictable, but not perfectly; Jimenez & Mendez, 1999; Song, Howard, & Howard, 2007).

In many ways, the colour-word contingency learning paradigm may seem redundant with the sequential learning paradigm, because both are speeded reaction time tasks that involve the learning of the relationship between stimuli and subsequent responses. However, the paradigms do differ in fundamental ways that may (or may not) prove significant. For instance, our paradigm involves participants applying contingency information on a trial-by-trial basis (i.e., participants cannot know what response to expect until they have begun to process the word). In sharp contrast to this, in the sequence learning paradigm participants learn a long repeating series of stimuli and responses. This may result in strategic differences in learning and may also affect the rate of learning. Additionally, instead of learning the association of stimuli to responses, in sequence learning participants may be learning the series of responses (which is impossible in the colour-word contingency learning paradigm, because there is no response sequence).

Another fundamental difference between our contingency learning paradigm and the sequence learning paradigm is the type of information being retrieved. For colour-word contingency learning, participants are required to learn the relation between a distracter and the associated response *within* individual trials to determine what response is likely given the current word. In contrast, for sequence learning participants are required to learn the relation between stimuli to responses *across* a number of trials to determine what response is likely to follow. For instance, if presentation of Stimulus A leads to retrieval of Memory X (i.e., an instance that contains Stimulus A), then participants could use Memory X to predict the response in our contingency learning paradigm, but would need to retrieve Memory X + 1 to predict the next item in sequence learning. What differences in learning this will lead to is unclear. More importantly, given these numerous fundamental differences, it cannot simply be assumed that every result found in sequence learning will also be found in colour-word contingency learning, or vice versa.

7.2. Summary

As we have highlighted, the colour-word contingency paradigm shares many similarities with other paradigms used to study contingency learning, but also has some differences. Thus, it appears premature to assume that an effect observed in one paradigm necessarily generalises to the colour-word contingency paradigm (or vice versa). That said, there are some important ways in which the current results parallel findings from other contingency learning paradigms. Experiment 1 demonstrated extremely rapid learning and unlearning of contingency information. Although we are the first to study unlearning, the finding of rapid learning is consistent with what has been found in the hidden covariation detection paradigm, where response biasing has been demonstrated after exposure to as few as one or two consistent trials (Lewicki, 1985, 1986). In the sequence learning task, learning has been shown to take about seven blocks of a ten-trial sequence (Nissen & Bullemer, 1987).

It is fascinating, however, that learning occurs so fast even in the colour-word contingency paradigm where contingencies are not 100%. Rapid learning in a probabilistic task has also been reported by Jacoby and colleagues (2003) using an item specific proportion congruent manipulation (which Schmidt & Besner, 2008 have argued is simply a colour-word contingency effect incidentally observed within the context of a conflict paradigm). Although Jacoby and colleagues did not provide individual *t*-tests for each block, visual inspection of their data suggests that a contingency effect was present in their very first block of 16 trials. Although there are not many studies on the learning rate in contingency learning paradigms (and we are not aware of any work on unlearning), it does appear that, in general, the human cognitive system is capable of very rapid learning (and unlearning) of covariations.

The results of Experiments 2 and 3 produced evidence that contingency learning in the colour-word paradigm is impaired when memory is loaded with a secondary task. Indeed, a similar result has been found in the sequence learning task, where minimal learning was found for participants under load (Nissen & Bullemer, 1987). Although more work is certainly needed, it is interesting that apparently very simple learning processes that are generally reported to occur without awareness (e.g., Lewicki, 1986; Nissen & Bullemer, 1987; Schmidt et al., 2007) seem to be dependent on the availability of memory resources (see Hassin, 2005 for a discussion of implicit working memory). The results of Experiment 3 also add an interesting extension to this research by showing that memory load appears to affect both storage and retrieval processes.

7.3. Conclusions

The colour-word contingency paradigm is a useful tool to study contingency learning. It is very simple, easy to program, and produces highly reliable results. In the three experiments presented here it can be seen that learning and unlearning of contingencies in this paradigm is very rapid and is dependent on memory resources. Two reanalyses of old data ruled out a repetition account of these data. We have suggested that a viable explanation for these (and other findings) is that participants encode and subsequently retrieve a finite set of instances and use these instances to extract contingency information to be used to facilitate high contingency responses. Connectionist accounts such as the SRN can likely also be modified to account for the current results. Whatever account ultimately prevails, the current results constrain viable such accounts to those that are fast and those that require a role for limited-capacity memory resources.

Acknowledgments

This research was supported by Grant A0998 from the Natural Sciences and Engineering Research Council of Canada (NSERC) to Derek Besner, by Grant BOF/GOA2006/001 of Ghent University to Jan De Houwer, and by an NSERC Canada Graduate Scholarship and FWO Visiting Postdoctoral Fellowship to James R. Schmidt.

References

Allan, L. G. (2005). Learning of contingent relationships. *Learning & Behavior, 33*, 127–129.
Beckers, T., De Houwer, J., & Matute, H. (2007). Editorial: Human contingency learning. *Quarterly Journal of Experimental Psychology, 60*, 289–290.
Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Quarterly Journal of Experimental Psychology, 13*, 90–102.
Bland, J. M., & Altman, D. G. (1995). Calculating correlation coefficients with repeated observations: Part 1—correlation within subjects. *BMJ, 310*, 446.
Carlson, K. A., & Flowers, J. H. (1996). Intentional versus unintentional use of contingencies between perceptual events. *Perception & Psychophysics, 58*, 460–470.
Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120*, 235–253.
Colzato, L. S., Raffone, A., & Hommel, B. (2006). What do we learn from binding features? Evidence for multilevel feature integration. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 705–716.
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. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior, 37*, 1–20.
De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127*, 853–869.
Fu, Q., Fu, X., & Dienes, Z. (2007). Implicit sequence learning and conscious awareness. *Consciousness and Cognition, 17*, 185–202.
Hassin, R. R. (2005). Nonconscious control and implicit working memory. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious. Oxford series in social cognition and social neuroscience* (pp. 196–222). New York: Oxford University Press.
Hendrickx, H., De Houwer, J., Baeyens, F., Eelen, P., & Van Avermaet, E. (1997a). Hidden covariation detection might be very hidden indeed. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 201–220.
Hendrickx, H., De Houwer, J., Baeyens, F., Eelen, P., & Van Avermaet, E. (1997b). The hide-and-seek of hidden covariation detection: Reply to Lewicki, Hill, and Czyzewska (1997). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 229–231.
Heuer, H., & Schmidtke, V. (1996). Secondary-task effects on sequence learning. *Psychological Research, 59*, 119–133.
Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus–response episodes. *Visual Cognition, 5*, 183–216.
Jacoby, L. L., Lindsay, D. S., & Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin & Review, 10*, 638–644.
Jimenez, L., & Mendez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory & Cognition, 25*, 236–259.
Kinder, A., & Shanks, D. R. (2001). Amnesia and the declarative/procedural distinction: A recurrent network model of classification, recognition, and repetition priming. *Journal of Cognitive Neuroscience, 13*, 648–669.
Kinder, A., & Shanks, D. R. (2003). Neuropsychological dissociations between priming and recognition: A single-system connectionist account. *Psychological Review, 110*, 728–744.
Lewicki, P. (1985). Nonconscious biasing effects of single instances on subsequent judgements. *Journal of Personality and Social Psychology, 48*, 563–574.
Lewicki, P. (1986). Processing information about covariations that cannot be articulated. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 135–146.
Lewicki, P., Hill, T., & Czyzewska, M. (1992). Nonconscious acquisition of information. *American Psychologist, 47*, 796–801.
Lewicki, P., Hill, T., & Czyzewska, M. (1997). Hidden covariation detection: A fundamental and ubiquitous phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 221–228.
Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492–527.
MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163–203.
Miller, J. (1987). Priming is not necessary for selective-attention failures: Semantic effects of unattended, unprimed letters. *Perception & Psychophysics, 41*, 419–434.
Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32*, 183–198.
Musen, G., & Squire, L. R. (1993). Implicit learning of color-word associations using a Stroop paradigm. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 19*, 789–798.
Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*, 1–32.
Psychology Software Tools. (2002). E-Prime. <http://www.pst-net.com>.
Risko, E. F., Schmidt, J. R., & Besner, D. (2006). Filling a gap in the semantic gradient: Color associates and response set effects in the Stroop task. *Psychonomic Bulletin & Review, 13*, 310–315.
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, & 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.
Schmidtke, V., & Heuer, H. (1997). Task integration as a factor in secondary-task effects on sequence learning. *Psychological Research, 60*, 53–71.
Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology, 60*, 291–309.
Song, S., Howard, J. H., Jr., & Howard, D. V. (2007). Implicit probabilistic sequence learning is independent of explicit awareness. *Learning & Memory, 14*, 167–176.