# Category Learning in the Colour-Word Contingency Learning Paradigm

James R. Schmidt,[1] Maria Augustinova,[2] & Jan De Houwer[1]

[1]Department of Experimental Clinical and Health Psychology, Ghent University

[2]Department of Psychology, Université de Rouen

**Word count:** 3873

## Author Notes

## Abstract

In the typical colour-word contingency learning paradigm, participants respond to the print colour of words where each word is presented most often in one colour. Learning is indicated by faster and more accurate responses when a word is presented in its usual colour, relative to another colour. To eliminate the possibility that this effect is driven exclusively by the familiarity of item-specific word-colour pairings, we examine whether contingency learning effects can be observed also when colours are related to categories of words rather than to individual words. To this end, the reported experiments used three categories of words (animals, verbs, and professions) that were each predictive of one colour. Importantly, each individual word was presented only once, thus eliminating individual color-word contingencies. Nevertheless, for the first time, a category-based contingency effect was observed, with faster and more accurate responses when a category item was presented in the colour in which most of the other items of that category were presented. This finding helps to constrain episodic learning models and sets the stage for new research on category-based contingency learning.

## Introduction

To interact successfully with the world around us, it is necessary to be able to learn the contingent regularities between events (Allan, 2005; Beckers, De Houwer, & Matute, 2007; De Houwer & Beckers, 2002). Whether learning to speak a language, to master a sport, or to take one's first steps, learning how one event produces another is critical for progression. Contingency learning may often be stimulus specific (e.g., learning that your dog hates the family cat), but also often involves categories of stimuli (e.g., learning that most dogs hate all cats). Indeed, learning (particularly human) is often based on abstract information (Brady & Oliva, 2008; Emberson & Rubinstein, 2016). Using language, for instance, we can learn about conceptual relations without necessarily referring to specific stimuli. In the present report, we explore the influence of this latter type of contingency (i.e., regularities based on categories) on performance in the *colour-word contingency learning paradigm* (Schmidt, Crump, Cheesman, & Besner, 2007).

This paradigm is frequently used to study human contingency learning (Atalay & Misirlisoy, 2012; Forrin & MacLeod, 2017, in press; Lin & MacLeod, 2018; Schmidt, Crump, Cheesman, & Besner, 2007; Schmidt & De Houwer, 2012a, 2012d, 2016a, 2016b; Schmidt, De Houwer, & Besner, 2010). In the typical preparation, participants are presented words in colours and are asked to respond to the print colour as quickly and accurately as possible. Critically, as illustrated in Table 1, each of the task-irrelevant words is presented most often in one colour (e.g., "look" most often in red). Responding quickly becomes faster and more accurate to *high contingency* trials (e.g., "look" in red), where the word is presented in the expected colour, relative to *low contingency* trials (e.g., "look" in green), where the word is presented in an unexpected colour.

| Colour | Word | | |
|---|---|---|---|
| | look | find | make |
| red | 8 | 1 | 1 |
| green | 1 | 8 | 1 |
| yellow | 1 | 1 | 8 |

*Table 1.* Typical contingency learning design.

*Note:* An example mapping. High contingency word-colour pairs are counterbalanced across participants.

This contingency effect not only indicates that the contingent regularities are learned, but also that these regularities automatically shape performance (i.e., response times and errors). Indeed, contingency learning effects within this sort of procedure are particularly interesting because they are primarily implicit in nature, with past events having automatic effects on performance (e.g., Schmidt & De Houwer, 2012a, 2012d), rather than being based on conscious judgments. These effects also appear very rapidly. They are already visible within a very small number of trials (Lin & MacLeod, 2018; Schmidt & De Houwer, 2016b; Schmidt et al., 2010), which is also true of related performance-based learning procedures (e.g., Lewicki, 1985; Nissen & Bullemer, 1987).

Even though in past reports many different stimulus dimensions have been used for both the task-irrelevant distracter (e.g., shapes, words, nonwords, colours) and task-relevant target (e.g., colours, colour words, neutral words, positive/negatively valenced words) stimuli (Forrin & MacLeod, 2017; Levin & Tzelgov, 2016; Schmidt & De Houwer, 2012b, 2012c), it was always the case that single, frequently repeated stimuli were the predictive stimuli. In the colour-word contingency learning paradigm depicted above, a very small set of irrelevant stimuli (e.g., three words in the example in Table 1) are presented repeatedly, with each of these distracting stimuli presented highly frequently in one colour (e.g., "look" printed in red multiple times) and less frequently in other colours (e.g., "look" in green or yellow only

occasionally).[1] Therefore, the extent to which this paradigm is actually useful for

understanding learning processes that operate outside of the laboratory remains an open issue.

Indeed, as with most contingency learning paradigms, the highly restricted stimulus

set that is used in the typical preparation of this paradigm is somewhat artificial. Even more

important, it leaves open the possibility that the performance-based contingency learning

effect observed in this paradigm is exclusively *item-specific*. It may be the case that faster and

more accurate responding to *high* (e.g., "look" in red) as compared to *low* (e.g., "look" in

green) contingency trials is simply due to the fact that participants became highly familiarized

with the compound stimulus "look" in red. Participants may have even kept track of the small

set of repeated stimuli in working memory (Lemercier, 2009). Thus the present report was

aimed at examining the extent to which learning effects observed in the colour-word

contingency learning paradigm occur with contingent regularities other than those associated

to a small set of repeatedly-presented words.

To this end, the colour-word contingency learning paradigm used in this report

comprised a large set of words, each presented only once. Critically, each word belonged to

one of three categories (animals, verbs, or professions). Each *category* was presented most

often in one colour (e.g., professions most often in grey), counterbalanced across

participants.[2] Within the context of the present report, then, high contingency trials are those

in which an item of a category is presented in the color in which most of the other items of

that category are presented (e.g., "lawyer" in grey), whereas low contingency trials are those

in which the item is presented in another colour (e.g., "doctor" in orange). If colour-word

---

[1] This is also generally true of related learning paradigms, such as the flanker contingency paradigm (Carlson & Flowers, 1996; Miller, 1987; Mordkoff, 1996; Mordkoff & Halterman, 2008), where flanking letters are predictive of a central target letter, and other related paradigms (e.g., Musen & Squire, 1993).
[2] Because the high contingency category for each colour was fully counterbalanced across participants (e.g., professions most often in blue for 1/3 of the participants, most often in orange for another 1/3, and most often in grey for the remaining 1/3), all word-colour pairs served equally often as both high and low contingency trials (across participants). As such, a difference between high and low contingency trials can only be due to learning, and not to, for instance, pre-existing associations between a given category and colour.

contingency effects occur even when colours are related to categories of words (i.e., beyond specific items), then faster and more accurate responses will be observed for high as opposed to low contingency trials.

Recall that the latter pattern of results would indicate not only that learning of the contingent regularities based on categories occurs, but also and importantly that this category-based contingency learning automatically shapes performance. As already mentioned, we know that people are capable of category-level learning. For instance, while performing an unrelated picture repetition detection task, a predictable sequence of picture categories (e.g., fish followed by dogs, flowers followed by birds, etc.) can be learned, even when the individual pictures do not repeat (Brady & Oliva, 2008; Emberson & Rubinstein, 2016). However, learning in these past studies was indexed by conscious contingency judgments. Whether this same sort of semantic-level category learning will have more indirect influences on response speed and accuracy in the colour-word contingency learning paradigm is less clear.

There are, however, some indications in the existing literature that category-level learning can have more indirect consequences. More specifically, in binding procedures, where the influence of one co-occurrence (or a small number of co-occurrences) of a stimulus and response on performance is investigated (i.e., rather than repeated presentations of high and low contingency stimulus pairs), results do not seem to be entirely item-specific. For instance, Horner and Henson (2011) found the response decision linked to pictures (e.g., of a dog) during prime trials influenced identification of the related words (e.g., "dog") on probe trials occurring afterward. Similar results have been observed with other changes in stimulus features (Allenmark, Moutsopoulou, & Waszak, 2015; Biederman & Cooper, 1991; Biederman & Gerhardstein, 1993). Given our recent proposal that binding effects likely result from the same memory processes that produce contingency learning (Schmidt & De Houwer,

2016a; Schmidt, De Houwer, & Rothermund, 2016), we expected the colour-word contingency effect based on abstract-level categories to occur as readily as the category-based binding effects just outlined. Accordingly, we expected responding to become faster and more accurate to *high contingency* trials where an item of a category (e.g., professions) is presented in the expected color (e.g., "lawyer" in grey), relative to *low contingency* trials where an item of the same category is presented in the unexpected color (e.g., "doctor" in orange). The two experiments reported here were designed to test this prediction.

## Experiment 1

**Method**

   **Participants.** Fifty-eight Dutch-speaking undergraduates of Ghent University participated in the study in exchange for €5.

   **Apparatus.** The experiment was programmed in E-Prime 2 (Psychology Software Tools, Pittsburgh, PA) and run on a standard PC. Responses were made on an AZERTY keyboard with the "J," "K," and "L" keys for purple, orange, and grey targets, respectively.

   **Design.** For the main part of the experiment, there were three categories of Dutch words: animals, first-person singular verbs, and professions. These words were selected from a Dutch word frequency list on Wiktionary.com. In particular, there were 64 exemplars of each category (192 total). The full set of stimuli is presented in Table A1 in the Appendix. Words were selected randomly without replacement in one large block of 192 trials. Each word was presented only once in one of three display colours (purple, orange, or grey). For each word, the display colour was determined randomly on a participant-by-participant basis, but not evenly. In particular, for one category of words (e.g., animals) there was an 80% chance for each word to be presented in purple, and only a 10% chance for orange or grey. Another category was mostly predictive of orange, and the third category mostly predictive of

grey (i.e., with the same contingencies). This is illustrated in Table 2. Which category was

predictive of which colour was completely counterbalanced across participants (i.e., six

counterbalancing orders). For instance, whether an animal in purple was high or low

contingency depended on the counterbalanced order of the participant. Before the main part of

the experiment, there was a practice phase, in which the neutral stimulus "@@@@@" was

presented in place of the word. This was presented 16 times in each colour (48 trials total).

*Table 2.* Contingency manipulation.

| Colour | Category | | |
|--------|----------|-------|-------------|
|        | Animals  | Verbs | Professions |
| purple | 8        | 1     | 1           |
| orange | 1        | 8     | 1           |
| grey   | 1        | 1     | 8           |

*Note:* An example mapping. High contingency
category-colour pairs were counterbalanced across
participants.

**Procedure.** Stimuli were presented in bold, 18 pt. Courier New font in the center of a

black (0,0,0) screen. Each trial began with a white (255,255,255) fixation "+" for 200 ms,

followed by a blank screen for 200 ms. Next, the word was presented in purple (128,0,128),

orange (255,165,0), or light grey (192,192,192). The stimulus remained on the screen until

either a response was made or 1500 ms elapsed. The next trial began immediately after a

correct response. If an error was made or the participant failed to respond in 1500 ms,

however, "XXX" was presented in white for 1000 ms.

**Data Analysis.** We report analyses of both mean correct response times and error

percentages. All participants had reasonably acceptable error rates (highest: 16%). Trimming

some of the (relatively) more error-prone participants from the sample had no influence on the

significance of results reported.

**Results and Discussion**

First, we investigated mean response times. As expected, mean correct response times

were significantly faster to high contingency trials (561 ms) than to low contingency trials

(572 ms), $t(57) = 2.241$, $SE_{diff} = 5$, $p = .029$, $\eta^2 = .08$. Thus, an 11 ms contingency learning

effect was produced with abstract-level categories. Because learning of contingent regularities

associated to a small set of repeatedly-presented words is already visible within a very small

number of trials (see above), as a supplementary analysis the aforementioned contingency

learning effect was further investigated by comparing the first versus second half of the

experiment (96 trials each). These data are presented in the left panel of Figure 1. In the first

half of the experiment, there was no significant difference between high (561 ms) and low

contingency (566 ms) trials (effect: 5 ms), $t(57) = .754$, $SE_{diff} = 6$, $p = .454$, $\eta^2 < .01$. In the

second half of the experiment, however, the contingency effect was significant (high: 561 ms;

low: 577 ms; effect 17 ms), $t(57) = 2.613$, $SE_{diff} = 6$, $p = .011$, $\eta^2 = .11$. However, this 12 ms

increase should be interpreted with caution, as experiment half and contingency did not

significantly interact, $F(1,57) = 2.309$, $SE_{diff} = 901$, $p = .134$, $\eta_p^2 = .04$.
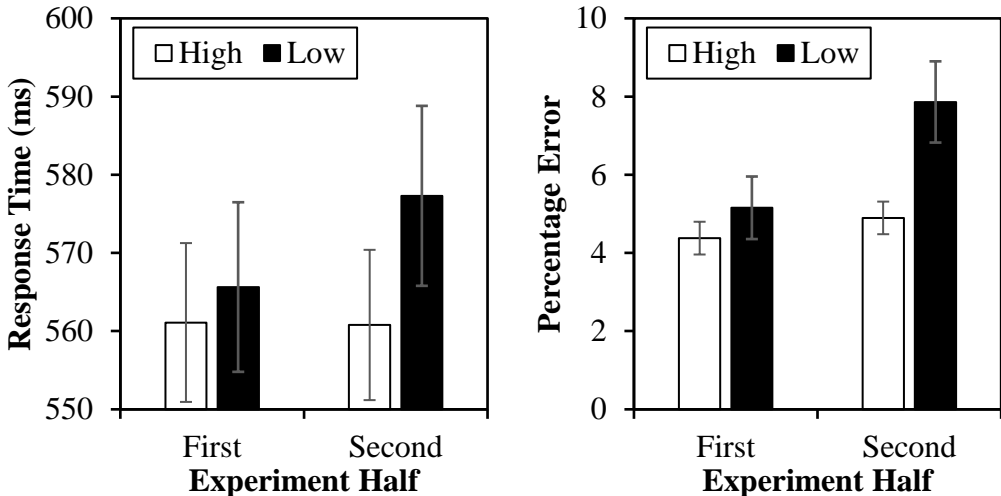


*Figure 1.* Mean response times (left) and percentage errors (right) with standard errors by
experiment half and contingency.

The percentage error data revealed a similar pattern of results, only slightly more

robust. Overall, errors were less frequent to high contingency trials (4.6%) than to low

contingency trials (6.4%), $t(57) = 2.798$, $SE_{diff} = 0.6$, $p = .007$, $\eta^2 = .12$, a 1.8% category-based

contingency effect. Similar to the response times, the contingency effect was not significant in

the first half of the experiment (high: 4.4%; low: 5.2%; effect: 0.8%), $t(57) = 1.003$, $SE_{diff} =$

0.8, $p = .320$, $\eta^2 = .02$, but was in the second half (high: 4.9%; low: 7.9%: effect: 3.0%), $t(57)$

$= 3.157$, $SE_{diff} = 0.9$, $p = .003$, $\eta^2 = .15$, and the experiment half by contingency interaction

was marginally significant, $F(1,57) = 3.755$, $SE_{diff} = 18.5$, $p = .058$, $\eta_p^2 = .06$.

In sum, Experiment 1 revealed evidence for learning of the contingencies between

categories of words and colour responses (i.e., a main effect of contingency). Most

interestingly, this was observed in a paradigm in which the individual words were presented

only once each, contrasting with the typical preparation in which a small set of words are

repeatedly presented throughout the experiment. This contingency effect was observed both in

the response times and in the error data. Moreover, there were some hints, albeit not robust,

suggesting that the categorical contingency effect took some time to develop, contrasting with

results in the standard colour-word contingency learning paradigm where the effect emerges

very quickly.

## Experiment 2

Experiment 2 served two purposes. First, we wanted to be certain that the category

contingency effects observed in Experiment 1 were not a Type 1 error. For this reason, we

conducted a (pre-registered) replication. Second, the exploratory block analyses reported in

Experiment 1 suggested that categorical contingency effects might take some time to develop.

However, this contingency by block interaction was only marginal in errors and non-

significant in response times, so the current experiment aimed to investigate this potential

effect further. As such, we lengthened the category lists and tested a larger sample of

participants in Experiment 2.

**Method**

**Participants.** In accordance with our pre-registration, we started with 58 participants

(same as previous study) and continued testing in testing days until we either

reached/exceeded our maximum sample size (92; based on initial power calculation) or had

strong Bayesian evidence for or against (a) a main effect of contingency and (b) a block by

contingency interaction. The resulting sample consisted of 94 participants,[3] recruited from the

same pool and with the same rewards as the previous experiment. One extra participant who

misunderstood the instructions (did not respond to grey stimuli) was replaced.

   **Apparatus, design, procedure, and data analysis.** The entire setup of the experiment

was identical to Experiment 1 with a few exceptions. First, we expanded the list of category

exemplars from 64 words per category to 100, resulting in 300 experimental trials total. The

extra stimuli are also presented in the Appendix. Given the increase in the number of trials,

block analyses were performed on three blocks of 100 trials each using a linear contrast. In all

other respects, the two experiments were identical.[4] Which categories were high contingency

in which colours was again fully counterbalanced.

**Results**

   As in Experiment 1, participants responded faster and more accurately to high

contingency trials (579 ms and 5.1% errors) than to low contingency trials (581 ms and 5.8%

errors), but this difference was only significant in errors, $t(93) = 2.303$, $SE_{diff} = 0.3$, $p = .024$,

$\eta^2 = .05$, and not in response times, $t(93) = .864$, $SE_{diff} = 3$, $p = .390$, $\eta^2 < .01$. The interaction

between block and contingency, presented in Figure 2, was not significant for response times,

$F(1,93) = .444$, $MSE = 918$, $p = .507$, $\eta_p^2 < .01$, or errors, $F(1,93) = .079$, $MSE = 19.4$, $p =$

$.779$, $\eta_p^2 < .01$.

---

[3] Note that this exceeds our maximum sample size because the final day of testing reached and exceeded the target sample size, consistent with how we described our stopping criteria in pre-registration.
[4] After completion of the study, participants were also asked for subjective category, subjective contingency, and objective contingency awareness. Due to space concerns, we retain these data for later analysis (as pre-registered).
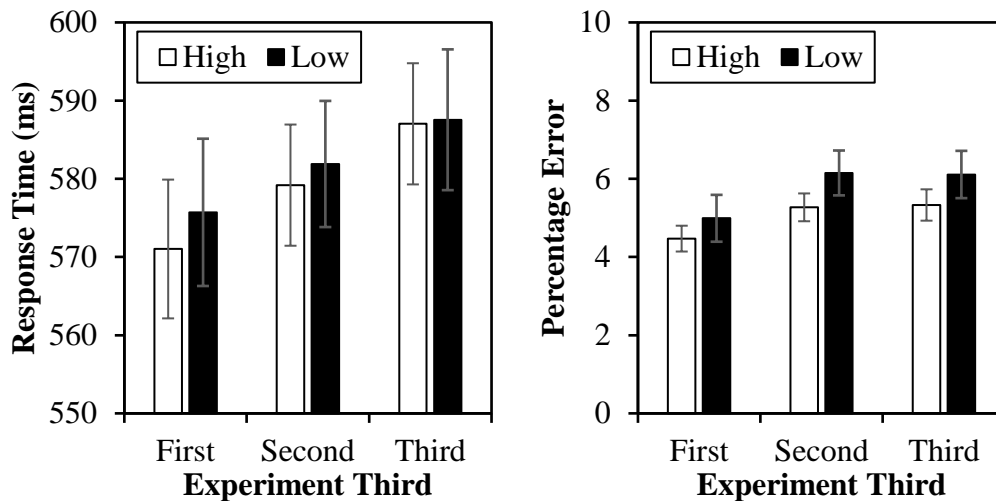
*Figure 2.* Mean response times (left) and percentage errors (right) with standard errors by experiment third and contingency.

## Discussion

Experiment 2 replicated a category-level contingency learning effect, albeit only robustly in errors. Why the effect in Experiment 2 was smaller than that in Experiment 1 is uncertain,[5] but the error results are consistent with the first experiment. Block analyses did not reveal any increase in the contingency effect over time. This seems to disconfirm the hint of an increasing effect in Experiment 1.

## General Discussion

The current report explored whether the contingency learning effect in the colour-word contingency learning paradigm also occurs with regularities based on abstract-level categories, rather than with a small set of repeated stimuli. As anticipated, even without repeated stimulus words, a contingency learning effect was observed both in response times and in errors in Experiment 1 and in errors in Experiment 2. This result thus precludes the

---

[5] Despite Experiment 2 being the larger experiment, the data were more atypically distributed (see supplementary data). In particular, Experiment 2 had a similar range of difference scores to that of Experiment 1, but fewer slightly above and more slightly below average scores than is typical of a normally-distributed effect. Aside from Type 1 or 2 error, it is alternatively possible that the added words in Experiment 2 were suboptimal or that the time of testing (beginning of academic year in Experiment 2) influenced effect magnitudes.

possibility that the colour-word contingency effects in this paradigm are driven exclusively by the familiarity of item-specific word-colour pairings (Lemercier, 2009).

As one limitation, the effect sizes were notably small in the current report. In the standard colour-word contingency learning paradigm (i.e., with a small set of repeated words), practically all participants show a contingency effect in the correct direction, which was clearly not the case in the current experiments. We did not, however, directly compare item-specific and category learning within one procedure, which might be an interesting avenue for future research. Future work might also aim to boost the size of the category-based contingency learning effect. For instance, pre-exposing the distracting word on each trial (i.e., positive stimulus onset asynchrony between word and colour) has been shown to boost the normal colour-word contingency learning effect (Schmidt & De Houwer, 2016b), which might also prove fruitful with the current preparation. Alternatively, participants might be informed about the categories (but not necessarily contingencies) in advance.

Despite this limitation, the present report shows that category-level (not just item-specific) contingencies can be learned and can subsequently influence performance in the colour-word contingency learning paradigm. This finding rules out the possibility that performance-based contingency learning effects observed in this paradigm are exclusively *item-specific*. In this respect, performance-based contingency learning effects are similar to those reported in the stimulus-response binding literature. This latter observation seems rather incompatible with the idea that binding effects result from temporary links between stimuli and responses formed in a short-term memory store, whereas contingency learning effects result from encoded regularities in a longer-term store (e.g., Colzato, Raffone, & Hommel, 2006). Instead, this similarity points to the viability of our proposal that binding and contingency learning effects result from the exact same memory processes (Schmidt & De Houwer, 2016a; Schmidt et al., 2016), namely episodic storage and retrieval (Logan, 1988).

Indeed, using the one memory retrieval mechanism, the Parallel Episodic Processing (PEP) model (Schmidt et al., 2016) was able to simulate performance (i.e., response times and errors) as typically observed in these distinct speeded-response paradigms (see also, Schmidt, in press).

The PEP model would struggle with the current data, however, as it uses localist inputs (i.e., one input node for each stimulus that could be presented to the model). Because of this, it is less clear how, for instance, one animal word (e.g., "cow") would lead to retrieval of other animals (e.g., dolphin, crow, etc.), because these correspond to different input nodes. However, the localist inputs to the PEP model are strictly for simplicity. In other (non-performance) models of episodic memory, retrieval from memory is determined by similarity in an *array* of stimulus features (e.g., Hintzman, 1984, 1986, 1988; Nosofsky, 1988a; Medin & Schaffer, 1978; Nosofsky, 1988b; Nosofsky & Palmeri, 1997; Nosofsky, Little, Donkin, & Fific, 2011). That is, presentation of a stimulus will lead to very strong retrieval of memories of the exact same stimulus, but also to (weaker) retrieval of memories of stimuli that are similar to the current stimulus. Though there were no clear visual features (Conway & Christiansen, 2005, 2006; Mervis & Rosch, 1981) distinguishing the categories in the current experiment (i.e., all stimuli were words), words within categories shared semantic overlap. For instance, presentation of the word "koe" (cow) will lead to partial retrieval of trial memories in which other animals were presented (dolfijn [dolphin], gans [goose], muis [mouse], etc.), 80% of which will have been paired with the same colour response (e.g., purple, if animals were presented most often in purple). Thus, responding will be biased toward the high contingency response. The reason for the relatively muted contingency effects in the current experiment also follows the standard episodic memory logic: category members may share *some* regularities (allowing for category-level learning), but are not perfect retrieval cues for memories of trials with other category members. Along this vein, the current

work suggests that future neural network research might aim to combine the performance capabilities of models like the PEP with more traditional episodic models, such as MINERVA 2, to simulate the category-level performance learning effects observed here.

Future research might also focus on comparing other aspects of the item-specific and category-level contingency paradigms. For instance, awareness of the contingency manipulation seems to be weakly related (but not completely unrelated) to the magnitude of the observed contingency effect in the typical preparation (Schmidt et al., 2007; Schmidt & De Houwer, 2012a, 2012d; Schmidt et al., 2010). The same may or may not be true for the categorical version of the paradigm presented here. On the one hand, simply detecting that there *is* a contingency manipulation in the current procedure may be substantially more difficult than in the more artificial case in which a very, very small set of stimulus words are repeatedly presented in the same colours. In other words, it is possible that *none* of the participants will become contingency aware with non-repeating words (meaning, of course, that contingency awareness did not play a role in the currently-observed effect). Alternatively, it could be the case that the categorical contingency effect observed here is *more* dependent on contingency awareness. For instance, it might be proposed that learning of lower-level stimulus feature to response bindings can be implicitly acquired and implemented, but that learning of regularities between abstract concepts requires conscious intervention. Future research might thus aim at studying the role of contingency awareness within the categorical contingency learning procedure novelly introduced in the current report. More broadly, the fact that learning effects successfully develop in both the item-specific and category-level color-word contingency learning paradigms points to perhaps the most important lesson that can be learned from this report: the colour-word contingency learning paradigm does serve as a viable means of studying the acquisition of knowledge about the relation between abstract features and responses.

# References

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

Allenmark, F., Moutsopoulou, K., & Waszak, F. (2015). A new look on S-R associations: How S and R link. *Acta Psychologica, 160*, 161-169.

Atalay, N. B., & Misirlisoy, M. (2012). Can contingency learning alone account for item-specific control? Evidence from within- and between-language ISPC effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 1578-1590.

Beckers, T., De Houwer, J., & Matute, H. (2007). Editorial: Human contingency learning. *Quarterly Journal of Experimental Psychology, 60*, 289-290.

Biederman, I., & Cooper, E. E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology, 23*, 393-419.

Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 1162-1182.

Brady, T. F., & Oliva, A. (2008). Statistical learning using real-world scenes: Extracting categorical regularities without conscious intent. *Psychological Science, 19*, 678-685.

Carlson, K. A., & Flowers, J. H. (1996). Intentional versus unintentional use of contingencies between perceptual events. *Perception & Psychophysics, 58*, 460-470.

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.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 24-39.

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between

    modalities: Pitting abstract against stimulus-specific representations. *Psychological*

    *Science, 17*, 905-912.

De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and

    theories on human contingency learning. *Quarterly Journal of Experimental*

    *Psychology, 55B*, 289-310.

Emberson, L. L., & Rubinstein, D. Y. (2016). Statistical learning is constrained to less

    abstract patterns in complex sensory input (but not the least). *Cognition, 153*, 63-78.

Forrin, N. D., & MacLeod, C. M. (2017). Relative speed of processing determines color-word

    contingency learning. *Memory & Cognition, 45*, 1206–1222.

Forrin, N. D., & MacLeod, C. M. (in press). The influence of contingency proportion on

    contingency learning. *Attention, Perception, & Psychophysics*.

Hintzman, D. L. (1984). Minerva 2: A simulation model of human memory. *Behavior*

    *Research Methods Instruments & Computers, 16*, 96-101.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model.

    *Psychological Review, 93*, 411-428.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace

    memory model. *Psychological Review, 95*, 528-551.

Horner, A. J., & Henson, R. N. (2011). Stimulus-response bindings code both abstract and

    specific representations of stimuli: Evidence from a classification priming design that

    reverses multiple levels of response representation. *Memory & Cognition, 39*, 1457-

    1471.

Lemercier, C. (2009). Incidental learning of incongruent items in a manual version of the

    Stroop task. *Perceptual and Motor Skills, 108*, 705-716.

Levin, Y., & Tzelgov, J. (2016). Contingency learning is not affected by conflict experience: Evidence from a task conflict-free, item-specific Stroop paradigm. *Acta Psychologica, 164*, 39-45.

Lewicki, P. (1985). Nonconscious biasing effects of single instances on subsequent judgments. *Journal of Personality and Social Psychology, 48*, 563-574.

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

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492-527.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207-238.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 89-115.

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

Mordkoff, J. T. (1996). Selective attention and internal constraints: There is more to the flanker effect than biased contingencies. In A. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 483–502). Washington, DC: APA.

Mordkoff, J. T., & Halterman, R. (2008). Feature integration without visual attention: Evidence from the correlated flankers task. *Psychonomic Bulletin & Review, 15*, 385-389.

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

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

Nosofsky, R. M. (1988a). Exemplar-based accounts of relations between classification,
    recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory,
    and Cognition, 14*, 700-708.

Nosofsky, R. M. (1988b). Similarity, frequency, and category representations. *Journal of
    Experimental Psychology: Learning, Memory, and Cognition, 14*, 54-65.

Nosofsky, R. M., Little, D. R., Donkin, C., & Fific, M. (2011). Short-term memory scanning
    viewed as exemplar-based categorization. *Psychological Review, 118*, 280-315.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded
    classification. *Psychological Review, 104*, 266-300.

Schmidt, J. R. (in press). Best not to bet on the horserace: A comment on Forrin and MacLeod
    (2017) and a relevant stimulus-response compatibility view of colour-word
    contingency learning asymmetries. *Memory & Cognition*.

Schmidt, J. R., Crump, M. J. C., Cheesman, J., & Besner, D. (2007). Contingency learning
    without awareness: Evidence for implicit control. *Consciousness and Cognition, 16*,
    421-435.

Schmidt, J. R., & De Houwer, J. (2012a). Adding the goal to learn strengthens learning in an
    unintentional learning task. *Psychonomic Bulletin & Review, 19*, 723-728.

Schmidt, J. R., & De Houwer, J. (2012b). Contingency learning with evaluative stimuli:
    Testing the generality of contingency learning in a performance paradigm.
    *Experimental Psychology, 59*, 175-182.

Schmidt, J. R., & De Houwer, J. (2012c). Does temporal contiguity moderate contingency learning in a speeded performance task? *Quarterly Journal of Experimental Psychology, 65*, 408-425.

Schmidt, J. R., & De Houwer, J. (2012d). Learning, awareness, and instruction: Subjective contingency awareness does matter in the colour-word contingency learning paradigm. *Consciousness and Cognition, 21*, 1754-1768.

Schmidt, J. R., & De Houwer, J. (2016a). Contingency learning tracks with stimulus-response proportion: No evidence of misprediction costs. *Experimental Psychology, 63*, 79-88.

Schmidt, J. R., & De Houwer, J. (2016b). Time course of colour-word contingency learning: Practice curves, pre-exposure benefits, unlearning, and relearning. *Learning and Motivation, 56*, 15-30.

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

Schmidt, J. R., De Houwer, J., & Rothermund, K. (2016). The Parallel Episodic Processing (PEP) Model 2.0: A single computational model of stimulus-response binding, contingency learning, power curves, and mixing costs. *Cognitive Psychology, 91*, 82-108.

# Appendix

*Table A1.* Experiment 1 Dutch stimuli and English translations.

| Animals | | Verbs | | Professions | |
|---|---|---|---|---|---|
| Dutch | English | Dutch | English | Dutch | English |
| aap | monkey | begin | begin | acteur | actor |
| adelaar | eagle | begrijp | understand | advocaat | lawyer |
| alligator | alligator | beloof | promise | agent | agent |
| baviaan | baboon | bescherm | protect | anatoom | anatomist |
| beer | bear | beslis | decide | apotheker | pharmacist |
| dolfijn | dolphin | betaal | pay | architect | architect |
| duif | pigeon/dove | beweeg | move | artiest | artist |
| eekhoorn | squirrel | bijt | bite | arts | physician |
| eend | duck | bind | tie | auditor | auditor |
| egel | hedgehog | blijf | stay | baker | midwife |
| ekster | magpie | brand | burn | bakker | baker |
| eland | moose | breng | bring | barman | bartender |
| fret | ferret | buig | bow | bewaker | security guard |
| gans | goose | controleer | control | biograaf | biographer |
| geit | goat | dans | dance | bioloog | biologist |
| gekko | gecko | denk | think | bloemist | florist |
| giraffe | giraffe | doe | do | bode | messenger |
| gorilla | gorilla | draag | wear | boekhouder | librarian |
| haai | shark | draai | turn | boer | farmer |
| haas | hare | drink | drink | butler | butler |
| hagedis | lizard | droom | dream | chirurg | surgeon |
| hert | deer | gebruik | use | cineast | filmmaker |
| hond | dog | geloof | believe | clown | clown |
| kakkerlak | cockroach | geniet | enjoy | coach | coach |
| kalkoen | turkey | herinner | remember | curator | curator |
| kameel | camel | hoor | hear | decaan | dean |
| kangoeroe | kangaroo | ken | know (person) | dierenarts | veterinarian |
| kat | cat | kies | choose | docent | teacher |
| kikker | frog | krijg | receive | drummer | drummer |
| koala | koala | maak | make | duiker | diver |
| koe | cow | neem | take | econoom | economist |
| konijn | rabbit | noem | name | editor | editor |
| kraai | crow | onderzoek | research | fabrikant | fabricator |
| krab | crab | ontwikkel | discover | fotograaf | photographer |
| kreeft | lobster | overleven | survive | imker | beekeeper |
| krokodil | crocodile | passeer | pass | intern | intern |
| leeuw | lion | praat | converse | kapitein | captain |
| lemur | lemur | probeer | try | kapper | hairdresser |
| lynx | lynx | rijd | drive | klerk | clerk |
| meerval | catfish | roep | call | kok | chef |
| muis | mouse | schaar | cut | komiek | comic |
| mus | sparrow | schrijf | write | laborant | lab worker |
| neushoorn | rhinoceros | sluit | close | manager | manager |
| olifant | elephant | speel | play | masseur | masseuse |

| Dutch | English | Dutch | English | Dutch | English |
|---|---|---|---|---|---|
| ooievaar | stork | teken | sign | mijnwerker | miner |
| otter | otter | vecht | fight | model | model |
| papegaai | parrot | verander | change | naaister | dressmaker |
| raaf | raven | verdenk | suspect | notaris | notary |
| rat | rat | verdien | earn | ober | waiter |
| schaap | sheep | vergeet | forget | opticien | optician |
| schildpad | tortoise | vergelijk | compare | piloot | pilot |
| slang | snake | verstuur | submit | priester | priest |
| tijger | tiger | vertaal | translate | rechter | judge |
| uil | owl | vertrek | leave | slager | butcher |
| varken | pig | vertrouw | trust | smid | smith |
| vleermuis | bat | verwissel | exchange | tandarts | dentist |
| vlinder | butterfly | vind | find | tolk | interpreter |
| vos | fox | vlieg | fly | trainer | trainer |
| walvis | whale | wacht | wait | uitgever | publisher |
| wasbeer | raccoon | wandel | walk | uitvinder | inventor |
| wolf | wolf | weet | know (thing) | vertaler | translator |
| zebra | zebra | werp | throw | violist | violinist |
| zeemeeuw | seagull | zet | put | visser | fisher |
| zwaan | swan | zwem | swim | zakenman | businessperson |

*Table A2.* Experiment 2 extra Dutch stimuli and English translations.

| Animals | | Verbs | | Professions | |
|---|---|---|---|---|---|
| Dutch | English | Dutch | English | Dutch | English |
| aal | eel | accepteer | accept | arbeider | labourer |
| aardvarken | aardvark | bedoel | mean | baas | boss |
| antilope | antelope | bemoei | interfere | barbier | barber |
| bever | beaver | bepaal | determine | bediende | servant |
| buffel | buffalo | bereik | reach | boswachter | forest ranger |
| capybara | capybara | commentaar | comment | chauffeur | driver |
| cavia | guinea pig | contacteer | contact | colporteur | canvasser |
| chimpansee | chimpanzee | ervaar | experience | conducteur | conductor |
| chinchilla | chinchilla | evacueer | evacuate | dagloner | day labourer |
| ezel | donkey | gooi | fling | dekker | roofer |
| gerbil | gerbil | help | help | dienstmeid | maid |
| goudvis | goldfish | huil | cry | doelman | goalkeeper |
| hamster | hamster | identificeer | identify | dokter | doctor |
| hyena | hyena | kruip | creep | houtzager | sawyer |
| jakhals | jackal | lees | read | ingenieur | engineer |
| kariboe | caribou | leid | lead | kassier | cashier |
| kip | chicken | lieg | lie | kastelein | landlord |
| kwal | jellyfish | luister | listen | landmeter | surveyor |
| lama | llama | moet | must | leraar | instructor |
| lemming | lemming | overhaal | persuade | loodgieter | plumber |
| luipaard | sloth | realiseer | realize | molenaar | miller |
| mangoest | mongoose | respecteer | respect | monteur | mechanic |
| muilezel | mule | schiet | shoot | muzikant | musician |
| octopus | octopus | schop | kick | opzichter | overseer |

| orka | killer whale | slaap | sleep | postbode | letter carrier |
|------|--------------|-------|-------|----------|----------------|
| paard | horse | slik | swallow | president | president |
| panda | panda | snijd | slice | procureur | attorney |
| panter | panther | spring | jump | regisseur | movie director |
| salamander | salamander | staan | stand | schilder | painter |
| stinkdier | skunk | stoor | disturb | schrijver | writer |
| tonijn | tuna | strijd | fight | soldaat | soldier |
| walrus | walrus | studeer | study | timmerman | carpenter |
| wezel | weasel | uitzoek | search | uitdrager | broker |
| wombat | wombat | verwerk | process | verkoper | salesperson |
| zalm | salmon | voel | feel | voetballer | soccer player |
| zeester | starfish | zend | send | voorzitter | chairperson |