UNIVERSITY of CALIFORNIA PRESS | Collabra: Psychology

Schmidt, J. R., et al. (2020). Learning Habits: Does Overtraining Lead to Resistance to New Learning? *Collabra: Psychology*, 6(1): 21. DOI: https://doi.org/10.1525/collabra.320

ORIGINAL RESEARCH REPORT

Learning Habits: Does Overtraining Lead to Resistance to New Learning?

James R. Schmidt[*,†], Jan De Houwer[†] and Agnes Moors[†,‡,§]

We explore the development of habitual responding within the colour-word contingency learning paradigm, in which participants respond to the colour of neutral words. Each word is most often presented in one colour. Learning is indicated by faster responses to the colour when the word is presented in the expected rather than in the unexpected colour. In Experiment 1, participants took part in two sessions, separated by one day. Critically, one set of words was trained across both days, and other new sets of words were introduced at various time points. Overall performance was faster on trials with overtrained words. Additionally, contingency effects were larger for overtrained words than for words introduced on Day 2. Removing the contingency had a similar impact on the learning effect for overtrained and new words. However, during a counterconditioning phase, where the words were made predictive of new colours, the previous contingency continued to influence performance for overtrained words but not for more recently introduced words. Relatedly, the new contingency was not acquired for the overtrained words. The reverse pattern was observed for recently-introduced words, with the newly-introduced contingency rapidly acquired and the influence of the old contingency quickly extinguished. In Experiments 2 and 3, however, both new and old learning effects were observed for both overtrained and recently-acquired contingencies. The net results suggest that while contingency learning effects are highly pliable during initial and subsequent learning, early-acquired contingency knowledge is maintained after removal of the contingency. Implications for models of learning are discussed.

Keywords: contingency learning; training; habits; overlearning; unlearning; counterconditioning

Introduction

For any intelligent organism to be able to understand and interact effectively with its environment, it is necessary to learn the regularities between events and outcomes (Allan, 2005; Beckers, De Houwer, & Matute, 2007; Shanks, 2010). Knowledge of the meaning of words, the tastes of foods, and the likely results of our actions are all built on this contingency learning backbone. These regularities in the environment shape the behavioural repertoire of the organism. The present paper explores how contingency learning also helps to shape automatized or default responding to a stimulus. In particular, we ask how quickly this responding becomes stable and resistant to changes in contingencies, and we discuss potential relations to habit formation.

One useful paradigm for studying contingency learning is the *colour-word contingency learning paradigm* (Schmidt, Crump, Cheesman, & Besner, 2007; for related paradigms, see Carlson & Flowers, 1996; Levin & Tzelgov, 2016; Lewicki, 1985, 1986; Miller, 1987; Musen & Squire, 1993; Schmidt & De Houwer, 2012b; for a review, see MacLeod, 2019). In the typical version of this paradigm, participants respond to the print colour of words (or the reverse; Forrin & MacLeod, 2017) with a key press (for verbal variants, see Atalay & Misirlisoy, 2012; Forrin & MacLeod, 2017). Each word is presented most often in one colour (e.g., "find" most often in purple, "help" most often in orange, etc.). Learning of the word-response contingencies is indicated by faster and more accurate responses to *high contingency* trials, where the word is presented in its most frequent colour (e.g., "find" in purple), relative to *low contingency* trials, where the word is presented in an infrequent colour (e.g., "find" in orange). One useful feature of this paradigm is the robustness of the effect, with nearly 100% of participants showing a positive contingency effect with very short experiments (e.g., 5–10 minutes).

In previous work, it has been observed that the colour-word contingency learning effect appears almost instantly after the start of the task (Schmidt et al., 2007; Schmidt & De Houwer, 2012c, 2016b; Schmidt, De Houwer, &

* LEAD-CNRS UMR 5022, Université Bourgogne Franche-Comté (UBFC), FR

† Department of Experimental Clinical and Health Psychology, Ghent University, BE

‡ Research Group of Quantitative Psychology and Individual Differences, KU Leuven, BE

§ Centre for Social and Cultural Psychology, KU Leuven, BE

Corresponding author: James R. Schmidt (james.schmidt@ubfc.fr)

Besner, 2010; O. Y.-H. Lin & MacLeod, 2018). Within the very first block of trials, the contingency effect is typically already robust, even with blocks as small as 18 trials. Similarly quick acquisition is observed in other, related implicit learning paradigms (e.g., Lewicki, 1985; Nissen & Bullemer, 1987; Schmidt & De Houwer, 2012a). There is, in addition, a small but significant gradual increase in the contingency effect with increasing practice (O. Y.-H. Lin & MacLeod, 2018; Schmidt & De Houwer, 2016b), indicating that there is a cumulative learning effect (i.e., continued strengthening of learning over time).

Interestingly, work has also indicated that when the contingency is removed from the experiment (i.e., after some blocks of experiencing the contingency), the contingency effect very rapidly diminishes (O. Y.-H. Lin & MacLeod, 2018; Schmidt & De Houwer, 2016b; Schmidt et al., 2010). Though not necessarily eliminated entirely (and even persisting in much reduced form for hundreds of trials; Schmidt & De Houwer, 2016b), the effect does approach zero almost as quickly as the initial contingency learning effect appeared. This suggests that the contingency effect is heavily influenced by very recently-encountered events. The fact that contingency learning effects are so pliable in the colour-word contingency learning paradigm is intriguing. Though persistent effects of previously experienced contingencies are still observed after the contingencies are removed, and the contingency effect cannot be exclusively explained by word-colour conjunctions one to five trials back (Schmidt et al., 2010; but see, Giesen, Schmidt, & Rothermund, 2020; Schmidt, Giesen, & Rothermund, in press), the results so far suggest that, for the most part, recent experience is what matters most.

Contingency learning effects like this are interesting in that they resemble a habit (for a review, see Wood & Rünger, 2016). Organisms tend to repeat behaviours in response to given stimuli, especially when these responses are rewarded (Thorndike, 1911). In the colour-word contingency learning paradigm this is reflected by the bias to repeat the frequently paired response to the (task-irrelevant) word stimulus. The best way to define a habit is not consistently agreed upon (De Houwer, 2019). On the one hand, a habit might be defined in a broad sense as roughly synonymous with "automaticity." In this sense, habitual responding is observed when a learned response is automatically evoked by a stimulus as the so-called "default" response (Evans & Stanovich, 2013). This definition focuses on the conditions under which responding occurs and does not specify the underlying mental mechanisms (whether the response is caused by stimulus representations, attitudes, or goals). A narrower definition of a habit, however, refers exclusively to the automatic priming of a specific response, and specifically excludes the operation of goals or attitudes (Wood & Rünger, 2016). Even more restrictively, Dickinson (1985; Heyes & Dickinson, 1990) defines habits as stimulus-driven responses (mediated by mental S-R associations) that have been installed via overtraining, and contrasts them with goal-directed actions that are driven by representations of the values and expectancies of the outcomes of these responses. By these more restrictive criteria, it can be quite difficult to determine when a behaviour is truly habitual, even in common metrics of habits like stimulus revaluation (e.g., De Houwer, Tanaka, Moors, & Tibboel, 2018; Moors, Boddez, & De Houwer, 2017).

Another way of thinking about contingency learning and habits (which is not necessarily incompatible with the view of Dickinson, 1985) is in terms of memory traces. According to an episodic memory perspective (Schmidt et al., 2010; Schmidt, De Houwer, & Rothermund, 2016), the colour-word contingency learning effect results from the storage and retrieval of episodic memories or exemplars (Logan, 1988). As more and more episodes linking a stimulus to a response are stored, presentation of said stimulus will more strongly bias retrieval in favour of the high contingency response.

In either case, the automatic biasing of the high contingency response to a word in the colour-word contingency learning paradigm might be regarded as a habit, though this might depend both on how one defines a habit and what assumptions are made about the mechanisms producing the effect (De Houwer, 2019). In the present manuscript, our primary interest is in exploring the automatic or default biasing of the high contingency response in the colour-word contingency learning paradigm, though we will return to potential implications for habit formation later.

The fact that contingency effects can be so pliable is interesting, as this suggests that recently-encoded events are particularly potent in their influence on performance, whereas older encoded events have minimal impact or rapidly become less potent (e.g., weakly retrievable). If so, then automatic responding might not be nearly as stable as previously thought. That is, it could be that a large chunk of what we regard as default responding (whether habitual in the restrictive sense or goal-mediated) is actually due to retrieval of only a very limited number of recently-encoded event memories (Giesen et al., 2020; Schmidt et al., in press). Note that persistence of an automatic response over time is not necessarily inconsistent with this notion: Repeatedly responding in a similar way to a stimulus as you did on your recent experiences with the same stimulus will tend to preserve the same behaviours long term.

It should be noted, on the other hand, that previous colour-word contingency learning experiments were relatively short in duration. They may therefore not reflect cases in which, after a considerable amount of practice, a stable representation of the meaningful connection between the stimulus and response may eventually emerge. For instance, over a lifetime one learns that the word "blue" is pronounced "blue." Even after considerable practice in a Stroop task where "blue" is presented equally often in all colours, the word "blue" will continue to interfere with naming of incongruent colours (Ellis & Dulaney, 1991; Gul & Humphreys, 2015; MacLeod, 1998). That is, the relation between the word "blue" and its verbalisation is not quickly forgotten simply because the word is no longer predictive of the "blue" verbal response within the context of the experiment. It could be argued, however, that training studies with Stroop stimuli like this should be interpreted differently. Because the (repeatedly-reinforced) goal is to name the *print colour* congruently throughout the task (e.g., saying "blue" to the print colour blue), the habitual links between *colour word* stimuli and verbal responses may be indirectly reinforced (e.g., the link between the concept "blue" and the verbalization "blue" is reinforced during colour naming, which indirectly keeps the word "blue" linked to a "blue" verbalisation). The colour-word contingency learning paradigm is more neutral in this regard, given the lack of a meaningful, pre-existing relation between the predictive (non-colour word) and target (colour) dimensions.

Also potentially informative in this regard is a series of experiments by MacLeod and Dunbar (1988). In these experiments, participants were trained to name novel shapes as colours (e.g., naming one polygon as "blue," another as "pink," etc.). After practice in shape naming, the task was reversed and participants named the (actual) print colours of the shapes. This allowed assessment of both "congruent" (or high contingency) trials, in which the (now task-irrelevant) shape is presented in the high contingency colour (i.e., the colour the shape was named as during training), and "incongruent" (or low contingency) trials, in which the shape is presented in a low contingency colour. In contrast to the colour-word contingency learning paradigm, a congruency/contingency effect was not observed immediately in this training paradigm. Instead, it took multiple days of training before the shape-colour contingency began to influence performance when naming the colour of the shapes. This may have been due to the change in the task context (i.e., swapping the task-relevant and -irrelevant features) or some other factor. In any case, in their Experiment 2, a congruency/contingency effect was still observed three months after the end of training. Although it is impressive that this contingency effect persisted in the absence of continued training, they did not test to see how this contingency effect was influenced by direct changes in the contingencies. It is possible, for instance, that the contingency effect would be immediately obliterated after introduction of a new contingency (counterconditioning) in which the shapes were briefly retrained with new colour names.

Given the above considerations, it is not clear whether the automatic response tendencies developed during implicit learning procedures, such as the colour-word contingency learning paradigm, are stable. More generally, if such a training procedure does produce stable responding, then how long might one need to practice before the contingency effect is stable enough to, for instance, be resistant to *unlearning* (i.e., removal of the contingency) or *counterconditioning* (i.e., introduction of a new, conflicting contingency)? That is, how much training is needed for acquired contingency knowledge to form a stable automatic response that is strong enough to override the *currently-experienced* stimulus-response contingencies?

In our previous research, we already studied unlearning with the colour-word contingency learning paradigm, but the contingency was always removed after a relatively short training period (O. Y.-H. Lin & MacLeod, 2018; Schmidt & De Houwer, 2016b; Schmidt et al., 2010). In the present series of three experiments, we investigated whether a contingency effect for heavily overtrained stimuli diminished rapidly after the contingency is removed (*unlearning*) and whether the effect of the original contingency was either further maintained or rapidly changed when a new, different contingency was introduced (*counterconditioning*). These three types of procedures (implemented in different phases of the experiments) are illustrated in **Figure 1**. We also investigated to what extent a new contingency was quickly acquired for stimuli that were previously associated with other colours. In particular, one set of

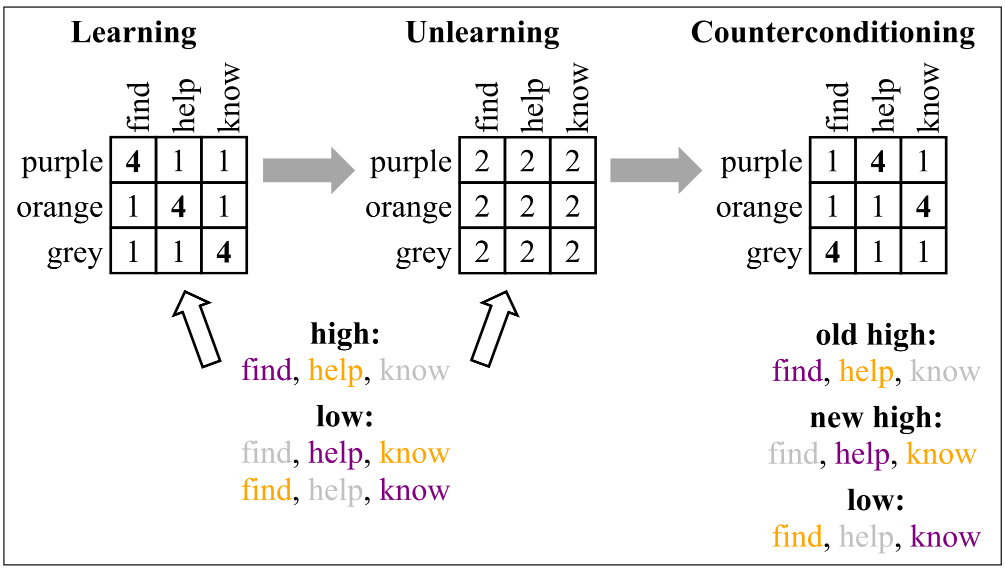

Figure 1: Three types of testing phases in the experiments with the stimuli that correspond to each condition. In the initial learning phase, a contingency was introduced. During the unlearning phase, the contingency was removed. During the counterconditioning phase, a different contingency for the same words was introduced. For overtrained words, the original learning phase comprised many blocks spread over either two days (Experiments 1 and 3) or one long session (Experiment 2). For other words, the original learning phase comprised only five small sub-blocks.

words was trained first for an extended period of time (e.g., for two days in Experiments 1 and 3 or in one long session in Experiment 2). After this training, the contingency was removed for the second to last phase, and finally an alternative contingency was introduced in the final phase. These effects for heavily overtrained stimuli were compared to the same effects for recently-introduced contingencies (i.e., with stimuli that were introduced only briefly before the unlearning and counterconditioning phases).

Two possible results could occur in this setup. The first possibility we term the *recent-events-matter-most* scenario. In this scenario, older events have minimal impact on performance, and performance is primarily determined by the stimulus-response bindings in the recently encountered events. That is, the learning mechanism is strongly "myopic" to events that were *just* experienced. This scenario is radically different from that predicted by traditional views of automatic responding discussed above, which assume that associations progressively strengthen over time, implying that frequency is more important than recency instead (of course, no one would argue that recent events do not influence behaviour at all). If the recent-events-matter most scenario obtains, then the contingency effect will disappear rapidly when the contingency is removed, even for the heavily trained stimuli. For instance, participants will stop responding faster to "find" in purple very shortly after "find" is changed to be presented equally often in all colours. It should similarly be expected that a newly introduced contingency for the same stimuli (i.e., counterconditioning) is rapidly learned. For instance, if "find" is now presented most often in orange (instead of purple), then participants should rapidly begin responding faster to "find" in orange, even after considerable training with "find" in purple. At the same time, the "old" high contingency (e.g., "find" in purple) should no longer influence performance after an extended unlearning phase and subsequent introduction of a "new" high contingency (e.g., "find" in orange).

The second possibility we term the *eventually-stable-habit* scenario. In this scenario, while *recently* acquired contingency knowledge may be more pliable early on (i.e., as the recent-events-matter-most scenario suggests), the memory bias for an *overtrained* contingency is more stable after sufficient training. In other words, sufficiently repeated encoding of a stimulus-response binding into memory eventually makes it difficult for new bindings to "break" the overtrained habit. Thus, for the overtrained stimulus set, unlearning should be less rapid. In other words, we might expect the original contingency to "stubbornly persist" after the contingency is removed for overtrained stimuli (i.e., no unlearning). Similarly, acquisition of a new contingency during counterconditioning should be reduced when the original contingency was overtrained. For instance, if "find" is changed to be presented most often in orange (rather than purple), then speeded responses to "find" in orange should not emerge quickly (perhaps not at all). This would indicate that the find-purple habit is too strongly ingrained to be quickly overcome. At the same time, the "old" high contingency should continue to influence performance, that is, participants should continue to respond quickly to "find" in purple long after the find-purple contingency has been replaced by the find-orange contingency.

It should be noted in advance that the recent-events-matter-most and eventually-stable-habit scenarios are deliberately presented as extremes, the former proposing very myopic learning and the latter proposing stubborn habit persistence. The truth may equally well lie somewhere in between these two extremes. That is, there might be both some continued influence of older experiences (e.g., from overtraining) in addition to influences of recently-acquired information. This would imply that we adapt quickly to newly-experienced events, but do not "catastrophically forget" everything that came before. To foreshadow our results, exactly this sort of mixed influence of both the new and old experiences was observed.

Experiment 1

Method

Participants

Fifty Ghent University undergraduates participated in the study on two separate 30 minute sessions one day apart in exchange for €10. Our sample size was determined a priori, but partially subjectively. In particular, as we had never previously studied counterconditioning with this procedure, we did not know how large of an effect to expect (i.e., for a priori power calculation). The current sample size seemed more than reasonable based on our prior experiences with the procedure on related topics. Two participants, however, did not show up for the second session and were therefore removed from the sample. Another participant had 16% incorrect responses in the main part of the experiment (i.e., excluding the practice block). This was over 2.5 standard deviations above the mean sample error rate. This participant was also removed from the sample. This participant contributed some notable noise to the sample, but inclusion of this participant did not influence the most critical results. Another participant had an empty cell (i.e., no correct response times) and was also removed.

Apparatus

The experiment utilized a standard PC. Stimulus and response timing were controlled with E-Prime 2 software (Psychology Software Tools, Pittsburgh, PA). The experiment files, along with the raw data, participant averaged data, and R scripts are available on the Open Science Framework (https://osf.io/7fwae/). Participants responded to purple, orange, and grey stimuli with the "J," "K," and "L" keys, respectively, on an AZERTY keyboard. Although we did not enforce specific fingers for the three keys, all participants in this and related studies defaulted to the standard keyboard resting position (i.e., right index on the J-key, right middle finger for the K-key, and right ring finger for the L-key).

Design

The structure of the experiment is presented in **Figure 2**. In each of the two testing days, participants were exposed to four larger mega-blocks, each with 5 sub-

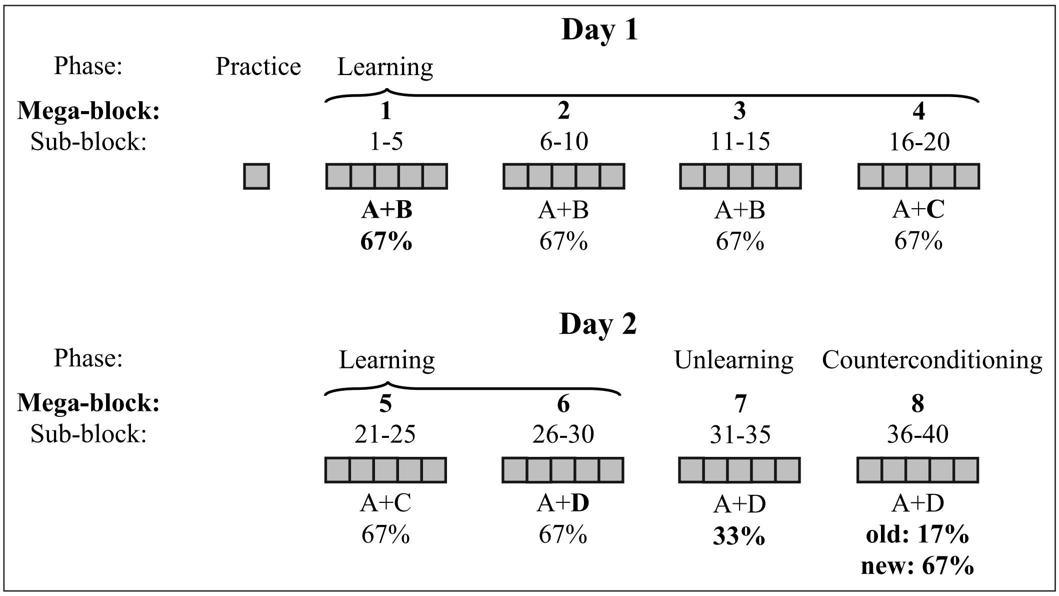

Figure 2: Composition of the phases, mega-blocks, and sub-blocks in the two testing days, with the sub-blocks of each mega-block indicated as separate squares. The two stimulus sets (A–D) in a given mega-block are indicated along with the contingency (in percentage) for the high contingency response. Changes in stimulus sets and contingency percentages from the prior mega-block are presented in bold. Note that in the final mega-block the "old" contingency is weakened and a "new" contingency is introduced.

Table 1: Dutch stimuli and English translations with example set assignments.

Set A		Set B		Set C		Set D	
Dutch	English	Dutch	English	Dutch	English	Dutch	English
vind	find	denk	think	zend	send	kies	choose
help	help	roep	call	lach	laugh	maak	make
weet	know	geef	give	neem	take	trek	pull

Note: This is only one example counterbalancing of words to sets.

blocks of 36 trials (180 trials in each mega-block; 1440 trials total). In addition, on Day 1 participants also began with a practice block of 36 trials to familiarize participants with the colour-to-key mapping. To this end, trials in the practice block consisted of the stimulus "@@@@" presented in each of three colours (purple, orange, and grey) 12 times each and participants were instructed to respond as quickly and accurately as possible. For the four experimental mega-blocks of each day, four sets of three words (Sets A–D) were created from a list of 12 four letter, first person Dutch verbs. The verbs are presented in **Table 1**. Note that which words were a part of which set and which words were presented most often in which colour were counterbalanced by assigning words, in the order listed in **Table 1**, to lists offset by participant number (e.g., "vind," "help," and "weet" were the Set A words for Participants 1, 13, 25, etc., "help," "weet," and "denk" were the Set A words for Participants 2, 14, 26, etc.). The first three experimental mega-blocks of Day 1 constituted the training phase in which Set A (e.g., "vind," "help," and "weet") and Set B words (e.g., "denk," "roep," and "geef") were presented. In each of five sub-blocks of each mega-block, one word from each set was presented most often (i.e., four of six times: 67%) in purple (e.g., "vind" and "denk"), a second most often in orange (e.g., "help" and "roep"), and a third most often in grey (e.g., "weet" and "geef"). Each word was presented once (17%) in each of the remaining two colours per sub-block. Note that Set A stimuli were the "overtrained" words that were used throughout the entire experiment. Set B words were filler words included in the first training phase to keep the task similar throughout (i.e., with six words in three colours). Because they were not of interest for the main analyses, we do not report the analyses on Set B stimuli, though we note that comparisons with Set A stimuli revealed nothing problematic. Sets C and D served as the non-overtrained stimuli that appeared later in the experiment. In the forth and final mega-block of Day 1, Set A stimuli remained, but Set B words were replaced by Set C words, which had a word-colour contingency manipulation identical to that for Sets A and B. In this final mega-block, it is possible to compare the magnitude of the learning effect of the heavily-overtrained Set A words with the newly-experienced Set C words.

On Day 2, the first mega-block was identical to the last one experienced on Day 1 (i.e., with Set A and Set C words), allowing a comparison of Set A and Set C again, but after a night of consolidation. The next mega-block again maintained the Set A words, but replaced the Set C words with Set D words (i.e., another newly trained set). In the third mega-block to follow, the same Set A and Set D words were again presented, but the words were no longer predictive of the colour response. In particular, each of the words was presented two out of six times (33%) in each of the three colours. Thus, this mega-block constitutes the unlearning phase as it allows comparing the rate of unlearning for the heavily-overtrained Set A words with the newly-learned Set D words. In a final mega-block, which constitutes the counterconditioning phase, the same Set A and Set D words were again presented, but the contingencies were now changed. In particular, the word that used to predict purple was now presented most often in orange, the word that used to predict orange was presented most often in grey, and the word that used to be presented most often in grey was presented most often in purple. Thus, this mega-block allowed us to assess to what extent the new contingency is learned and whether the heavily-overtrained Set A contingencies are more resistant to a change in contingency than the newly-experienced Set D contingencies. For each of these sets (A and D), we can compare *old high contingency* items, which are pairings that previously had a high contingency but not anymore, to *new high contingency* items, which are pairings that currently have a high contingency but did not before. Both of these can further be compared to *low contingency* items, which are pairings that never had high contingency.

Note that for most of the above-mentioned contrasts we might not only expect larger contingency effects for overtrained stimuli, but also potential main effects of word type. This is because both the high and the low contingency pairings with overtrained stimuli have been experienced more frequently than the high and low contingency pairings with recently-acquired contingencies (for further discussion of the frequency versus proportion distinction, see Schmidt & De Houwer, 2016a). In other words, we might expect that high contingency Set A items will be responded to faster than high contingency Set D items and, similarly, that low contingency Set A items will be responded to faster than low contingency Set D items.

Procedure

Stimuli were presented in the center of a black (0,0,0) screen, and presented in bold, 18 pt. Courier New font. On each trial, the participant was first presented a white (255,255,255) fixation "+" for 150 ms. This was followed by the word (or @'s during practice) in a neutral brown (255,183,113) for 150 ms, which was then colourized in one of the three target colours: purple (128,0,128), orange (255,165,0), or a light grey (192,192,192), which correspond to "purple," "orange," and "silver" in the standard E-Prime/HTML colour palette. This word preview was used because it is known to boost contingency effects (Schmidt & De Houwer, 2016b), likely because the word has more time to influence colour identification. The stimulus remained on the screen until either a response was made or 1500 ms elapsed. Following correct responses, the next trial began immediately. Following an incorrect response or 1500 ms without a response "XXX" was presented in white for 1000 ms before the next trial. Participants were instructed to try to respond as quickly and accurately as possible.

Results

Analyses focused on mean correct response times during the main phases of the experiment (i.e., practice phase excluded). Trials for which participants did not respond before the 1500 ms deadline were excluded, but no other response time trims were performed (as has been our standard practice with this paradigm). Error data are not reported here given that they were far too noisy to produce anything meaningful and that the general length of the reported analyses was already long, but there were no speed-accuracy trade-offs and the error data are available for download (along with the response time data and R scripts for both dependent measures) on the Open Science Framework (https://osf.io/7fwae/). In all analyses, sub-block was treated as a linear factor, which is more sensible than treating the sub-block as an unordered factor. Indeed, a linear factor allows for inferences about increases or decreases across sub-blocks (whether for the main effect or interactions involving sub-block), whereas the same is not true when treating sub-block as a categorical factor (e.g., a significant sub-block effect could hypothetically emerge due to abnormally fast or slow responses in one of the middle blocks). Linear factors should generally be used for any interval or scale factor. In the interest of brevity, only the theoretically interesting contrasts are reported (whether significant or non-significant). We do not report the less interesting contrasts, unless $p < .1$ (thus, the reader can correctly assume that any non-reported factor or interaction is not significant). Note that we did not preregister, but all data analyses for this and the following two experiments were planned in advance and similar to those in our past reports with this task. The only exception was the addition of some Bayes tests, which we added in response to editor feedback. All Bayes tests were conducted with the BayesFactor package in r with the default Cauchy prior and 100,000 recomputes to increase precision. The contingency effect as a function of the mega-blocks is presented in **Figure 3** and the sub-block means and standard errors are presented in Table A1 for both response times and errors (see also the R scripts).

Day 1, Sets A and C

First, we compared Set A (which had already been trained for 15 sub-blocks) with the newly added Set C using a sub-block (16–20) by contingency (high vs. low) by set (A vs. C) ANOVA to test for potential differences between the more heavily trained Set A stimuli over the new Set C stimuli. The contingency effect was significant, $F(1,45) = 10.233$, $MSE = 5520$, $p = .002$, $\eta_p^2 = .19$, and this contingency effect increased across blocks, $F(1,45) = 10.645$, $MSE = 4321$, $p = .002$, $\eta_p^2 = .19$. Interestingly, there was a main effect of set, $F(1,45) = 8.987$, $MSE = 3331$, $p = .004$, $\eta_p^2 = .17$, indicating faster overall performance for the overtrained Set A stimuli. In particular, Set A high contingency trials (mean = 612 ms, $SE = 11$) were marginally faster than Set C high contingency trials (mean = 622 ms, $SE = 11$), $F(1,45) = 3.493$, $MSE = 2986$, $p = .068$, $\eta_p^2 = .07$, and Set A low contingency trials (mean = 626 ms, $SE = 12$) were significantly faster than Set C low contingency trials (mean = 639 ms, $SE = 13$), $F(1,45) = 4.852$, $MSE = 4189$, $p = .033$, $\eta_p^2 = .10$. Thus, an overall advantage was evident

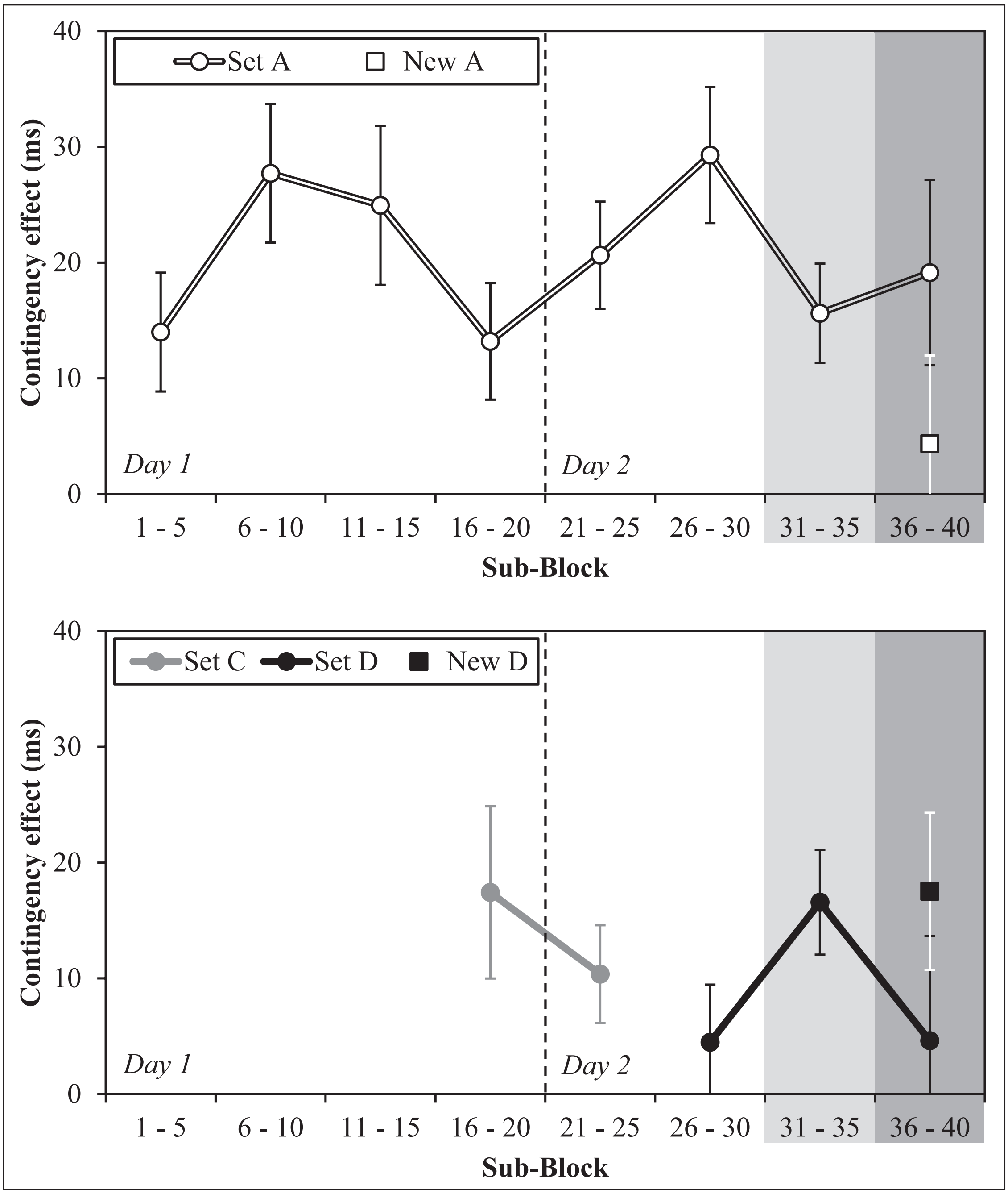

Figure 3: Experiment 1 contingency effect (low minus high contingency) as a function of block for the overtrained stimuli (top) and other stimuli (bottom) with standard errors. The sub-block numbers are indicated on the *x*-axis, but data are aggregated over the larger mega-blocks for presentation purposes. Light grey indicates the unlearning phase and dark grey indicates the counterconditioning phase. The single squares in the last block represent the low minus new high contingency contrast.

for the overtrained stimuli, albeit only significantly so for the infrequent pairings. Globally, this is consistent with the notion that the frequency of co-occurrences between a stimulus and response is important, not just the proportion (Schmidt & De Houwer, 2016a). However, the contingency effect (i.e., the difference between high and low contingency trials) was not significantly different between Set A and C stimuli, $F(1,45) = 0.213$, $MSE = 3843$, $p = .545$, $\eta_p^2 < .01$, $BF_{01} = 8.95$.

Day 2, Sets A and C

Sets A and C were again compared on Day 2 with a sub-block (21–25) by contingency (high vs. low) by set (A vs. C) ANOVA. The contingency effect was significant, $F(1,45) = 30.091$, $MSE = 1824$, $p < .001$, $\eta_p^2 = .40$. There was again a main effect of block, $F(1,45) = 13.312$, $MSE = 6216$, $p < .001$, $\eta_p^2 = .23$, perhaps again hinting at fatigue. As with the previous block, the contingency effect was not significantly different between the two sets, $F(1,45) = 2.277$, $MSE = 2713$, $p = .138$, $\eta_p^2 = .05$, $BF_{01} = 2.68$, though the trend was for a larger effect for the overtrained Set A stimuli (mean = 550 ms, $SE = 8$ and mean = 571 ms, $SE = 8$, respectively, for high and low contingency) than for Set D stimuli (mean = 560 ms, $SE = 8$ and mean = 570 ms, $SE = 10$, respectively).

Day 2, Sets A and D

Next, we compared Set A with a newly-introduced Set D using a sub-block (26–30) by contingency (high vs. low) by set (A vs. D) ANOVA, again to test for any potential advantages for the overlearned Set A stimuli over the new Set D stimuli. The main effect of contingency was significant, $F(1,45) = 21.608$, $MSE = 2915$, $p < .001$, $\eta_p^2 = .32$. This contingency effect interacted with set, $F(1,45) = 8.817$, $MSE = 3996$, $p = .004$, $\eta_p^2 = .16$, indicating larger contingency effects for the overtrained Set A stimuli than for the new Set D stimuli. Indeed, the contingency effect was significant for Set A, $F(1,45) = 24.466$, $MSE = 3932$, $p < .001$, $\eta_p^2 = .35$, but not for Set D, $F(1,45) = 0.672$, $MSE = 2978$, $p = .417$, $\eta_p^2 = .01$. Again, there was a main effect of set, $F(1,45) = 5.295$, $MSE = 1589$, $p = .026$, $\eta_p^2 = .11$, indicating faster responses to the overtrained Set A stimuli. In particular, Set A high contingency trials (mean = 563 ms, SE = 9) were significantly faster than Set D high contingency trials (mean = 582 ms, SE = 10), $F(1,45) = 19.160$, $MSE = 2037$, $p < .001$, $\eta_p^2 = .30$, but there was no difference between Set A (mean = 592 ms, SE = 11) and D (mean = 586 ms, SE = 8) low contingency trials, $F(1,45) = 1.299$, $MSE = 3547$, $p = .260$, $\eta_p^2 = .03$.

Day 2, unlearning

Next, we compared Sets A and D during unlearning using a sub-block (31–35) by contingency (high vs. low) by set (A vs. D) ANOVA (i.e., where the contingency factor codes for the previously-applicable regularity). The contingency effect across sets was still significant during unlearning, $F(1,45) = 26.748$, $MSE = 2178$, $p < .001$, $\eta_p^2 = .37$. However, there was no difference in the contingency effect between sets, $F(1,45) = 0.034$, $MSE = 2382$, $p = .855$, $\eta_p^2 < .01$, $BF_{01} = 9.48$, with the difference between high (mean = 579 ms, SE = 11) and low contingency trials (mean = 594 ms, SE = 11) in Set A being similar to that in Set D (mean = 578 ms, SE = 11 and mean = 595 ms, SE = 10, respectively). This finding is not as the eventually-stable-habit hypothesis would predict. In addition, the contingency effect was marginally smaller for Set A stimuli in the unlearning phase relative to the preceding acquisition phase, $F(1,45) = 4.029$, $MSE = 528$, $p = 0.051$, $\eta_p^2 = 0.08$. The fact that the contingency effect was still robust in the absence of a contingency is also not consistent with the recent-events-matter-most hypothesis. There was also a marginal main effect of sub-block, $F(1,45) = 3.483$, $MSE = 3324$, $p = .069$, $\eta_p^2 = .07$, again hinting at fatigue.

Day 2, counterconditioning

Finally and most critically, we considered the counterconditioning blocks where we directly pitted an old (recent or overtrained) high contingency against a newly-introduced inconsistent high contingency. Most importantly, we began by considering whether reduction in the effect of the old high contingency and learning of the new high contingency was faster with the new Set D stimuli than with the overtrained Set A stimuli. For this, we began by comparing the trials in which the word was presented with the colour that was high contingency during initial training (old high contingency) with trials in which the word was presented with the colour that is currently high contingency (new high contingency). Trials in which the word was presented in a colour that was low contingency in all phases will be considered afterwards. Thus, we first conducted a sub-block (36–40) by contingency (old high vs. new high) by set (A vs. D) ANOVA, and followed this with the relevant contrasts. There was no main effect of contingency, $F(1,45) = 0.002$, $MSE = 3438$, $p = .969$, $\eta_p^2 < .01$. Note, of course, that this is not a test of the contingency effect per se, but of the comparison between the new versus old contingency. Thus, this should not be interpreted as no evidence of learning. There was a robust crossover interaction between contingency and set, $F(1,45) = 14.334$, $MSE = 4009$, $p < .001$, $\eta_p^2 = .24$, as illustrated in **Figure 4**. Exploring this interaction further, old high contingency items (mean = 579 ms, SE = 10) were responded to significantly faster than new high contingency items (mean = 593 ms, SE = 11) in Set A, $F(1,45) = 8.780$, $MSE = 3336$, $p = .005$, $\eta_p^2 = .16$, but significantly *slower* (mean = 595 ms, SE = 11 vs. mean = 582 ms, SE = 10, respectively) in Set D, $F(1,45) = 6.856$, $MSE = 4111$, $p = .012$, $\eta_p^2 = .13$. Furthermore, in ANOVAs comparing high to low contingency, responses were significantly faster on old high contingency relative to low contingency trials (mean = 598 ms, SE = 12) for Set A, $F(1,45) = 6.890$, $MSE = 7019$, $p = .012$, $\eta_p^2 = .13$, indicating a preservation of the original learning, but there was no significant difference between new high contingency and low contingency trials, $F(1,45) = 0.351$, $MSE = 6774$, $p = .557$, $\eta_p^2 < .01$, $BF_{01} = 7.96$, suggesting a failure to acquire the new contingency (albeit with a numerical trend in the correct direction). In contrast, for Set D stimuli there was no difference between old high and low contingency trials (mean = 599 ms, SE = 11), $F(1,45) = 0.356$, $MSE = 10570$, $p = .553$, $\eta_p^2 < .01$, $BF_{01} = 8.84$, suggesting abolition of the original contingency learning effect (though, again the means were in the correct direction), but a significant difference between new high and low contingency items, $F(1,45) = 8.742$, $MSE = 6013$, $p = .005$, $\eta_p^2 = .16$, indicating acquisition of the new contingency.

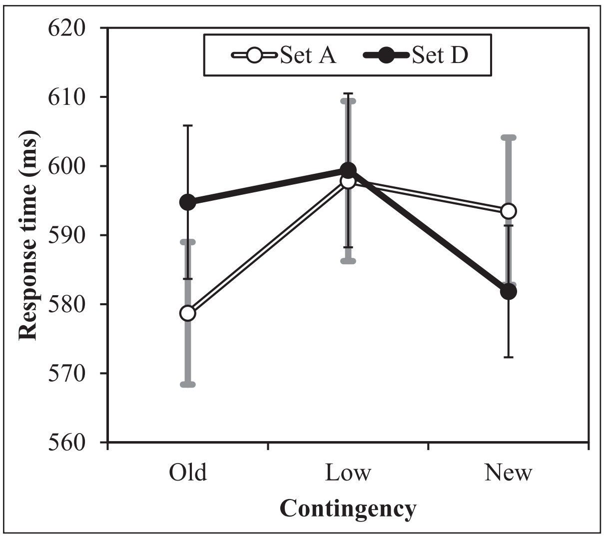

Figure 4: Experiment 1 mean response times with standard errors for old, new, and low contingency items (collapsed across sub-block) for the overtrained (Set A) and recently-introduced (Set D) stimuli during counterconditioning.

Discussion

Experiment 1 produced results that did not overwhelmingly support either of the two strong views mentioned in the Introduction. For some contrasts, there was no robust evidence of a difference between overtrained and newly-acquired contingencies. For instance, no differences were observed between Set A and Set D stimuli in the unlearning phase, where the contingency effect for overtrained Set A stimuli marginally decreased (i.e., did not "stubbornly persist" in an unchanged magnitude). Findings such as this could be considered as evidence for the recent-events-matter-most scenario, which predicts no differences between heavily-overtrained and newly-learned contingencies. On the other hand, overall response times were faster to the overtrained Set A stimuli relative to newly-added stimuli during the acquisition phases for Sets C and D (but only on Day 1 for the former). That is, there was a general speedup of responses to the overtrained stimuli, both high and low contingency, relative to recently-introduced stimuli. This is consistent with the notion that practice with the co-occurrence of stimuli (even if task irrelevant) and responses benefits performance (Lemercier, 2009; Schmidt & De Houwer, 2016a; Schmidt et al., 2016). That is, even though the proportions of high versus low contingency pairings are equivalent for overtrained and recently-introduced contingencies (i.e., 4/6 high contingency pairings, and 1/6 for each low contingency pairing), participants have more frequently observed each compound stimulus for the overtrained Set A stimuli (Y.-H. Lin, 2015). For instance, by the end of the acquisition phase for Set D, participants had seen each high contingency pairing 120 times for Set A, but only 20 times for Set D. Similarly, they had seen each low contingency pairing 30 times for Set A, but only 5 times for Set D. However, the contingency effect (i.e., difference between high and low contingency trials) was only (robustly) larger in the Set A versus Set D comparison. For Set C, it is possible that the overall advantage (i.e., main effect) for the overtrained stimuli worked against the contingency effect, in line with previous findings that effects tend to scale up with mean response time (Stevens et al., 2002; Urry, Burns, & Baetu, 2015; Schmidt & De Houwer, 2016b; Schmidt et al., 2016). That is, even though contingency *knowledge* may be stronger for the overlearned stimuli, there is less time for this contingency knowledge to be *expressed* (i.e., to influence colour decisions) as overall responding was quicker to overlearned stimuli. Stated differently, even though the contingency knowledge might be stronger for Set A, the difference between high and low contingency trials might not be notably larger in this set because the fast overall response speed to Set A stimuli precluded the influence of the contingency knowledge on the response times.

During counterconditioning, there was a trend for faster responses to both the old and new high contingency trials relative to low contingency trials, though only the old contingency effect was robust for the overtrained Set A and only the new contingency effect was robust for newly added Set D items. The interaction between set (A vs. D) and contingency (new vs. old) was robust. This suggests that the old contingency persists to a greater extent with overtraining, and the new contingency is learned more quickly with newly-acquired contingencies. For all comparisons between overtrained and recently-acquired contingencies, no results pointed in the reverse direction than the eventually-stable-habits account would suggest (e.g., smaller effects of the old contingency for Set A), though an overwhelming difference between overtrained and recently-acquired contingencies does not seem apparent (i.e., for many contrasts there were no differences, and for others only small differences). Together, the results might suggest some carryover influence of overtraining, but a remaining potent influence of recent events (i.e., even with overtraining).

Experiment 2

Results from Experiment 1 suggest that neither of the two extreme views we discussed in the Introduction are correct. That is, it was not the case that only recent events had an influence on performance (which would have predicted no differences at all between Set A and Sets C and D), and it was also not the case that overlearning completely prevented any new learning (which would have predicted, for instance, no unlearning at all for Set A). Instead, the results supported an intermediate view, with some findings suggesting a lasting (though perhaps subtle) influence of overtraining, but with a robust influence of very recent experiences. Experiment 2 aimed to provide a conceptual replication of Experiment 1. The most important change was that we dropped the two-day design and instead used a longer single day design (with 6 mega-blocks in one day, instead of 4 per day). We also dropped the Set C stimuli. Put differently, Experiment 2 was identical to Experiment 1 in all respects except that sub-blocks 16–25 (the fourth and fifth mega-blocks in **Figure 2**) were dropped and the remaining 30 sub-blocks (6 mega-blocks) were tested in one day. This does imply, however, that the initial training duration for Set A is shorter in Experiment 2 and does not include a night of sleep consolidation. Also note that we continue to label the last-introduced words as "Set D" for consistency with the prior experiment even though there was no longer a Set C.

Method

Participants

Sixty-one Ghent University undergraduates participated in the study in one 30 minute session in exchange for €5. A slightly larger sample was collected because we supposed that any potential differences between overtrained and recently-learned contingencies might be smaller with a shorter training period. Using the same exclusion criteria as the prior experiment, only one participant was excluded due to an empty cell.

Apparatus, Design, and Procedure

The apparatus, design, and procedure of the current experiment were identical to Experiment 1 with the following exceptions. The two mega-blocks with Set C stimuli from Experiment 1 (see **Figure 2**) were dropped and the remaining six mega-blocks were run in one day (i.e., 3 A+B learning mega-blocks, followed by A+D learning, unlearning, and counterconditioning blocks; 1080 trials in total).

Results

Data were analysed in the same manner as in the prior experiment. The contingency effect as a function of the mega-blocks is presented in **Figure 5**. The means and standard errors for the response times and errors are presented in Table A2 in the Appendix (see also the R scripts).

Sets A and D

First, we compared Set A (which had already been trained for 15 sub-blocks) with a newly-introduced Set D using a sub-block (16–20) by contingency (high vs. low) by set (A vs. D) ANOVA. The main effect of contingency was significant, $F(1,59) = 28.503$, $MSE = 5422$, $p < .001$, $\eta_p^2 = .33$. This contingency effect interacted with set, $F(1,59) = 11.529$, $MSE = 3099$, $p = .001$, $\eta_p^2 = .16$, indicating larger contingency effects for the overtrained Set A stimuli than for the new Set D stimuli. The contingency effect was significant for Set A (mean = 595 ms, $SE = 8$ and mean = 629 ms, $SE = 8$ for high and long contingency trials, respectively), $F(1,59) = 39.518$, $MSE = 4288$, $p < .001$, $\eta_p^2 = .40$, and for Set D (mean = 610 ms, $SE = 8$ and mean = 622 ms, $SE = 8$, respectively), $F(1,59) = 4.919$, $MSE = 4233$, $p = .030$, $\eta_p^2 = .08$. There was also a marginal main effect of sub-block, $F(1,59) = 3.256$, $MSE = 6562$, $p = .076$, $\eta_p^2 = .05$.

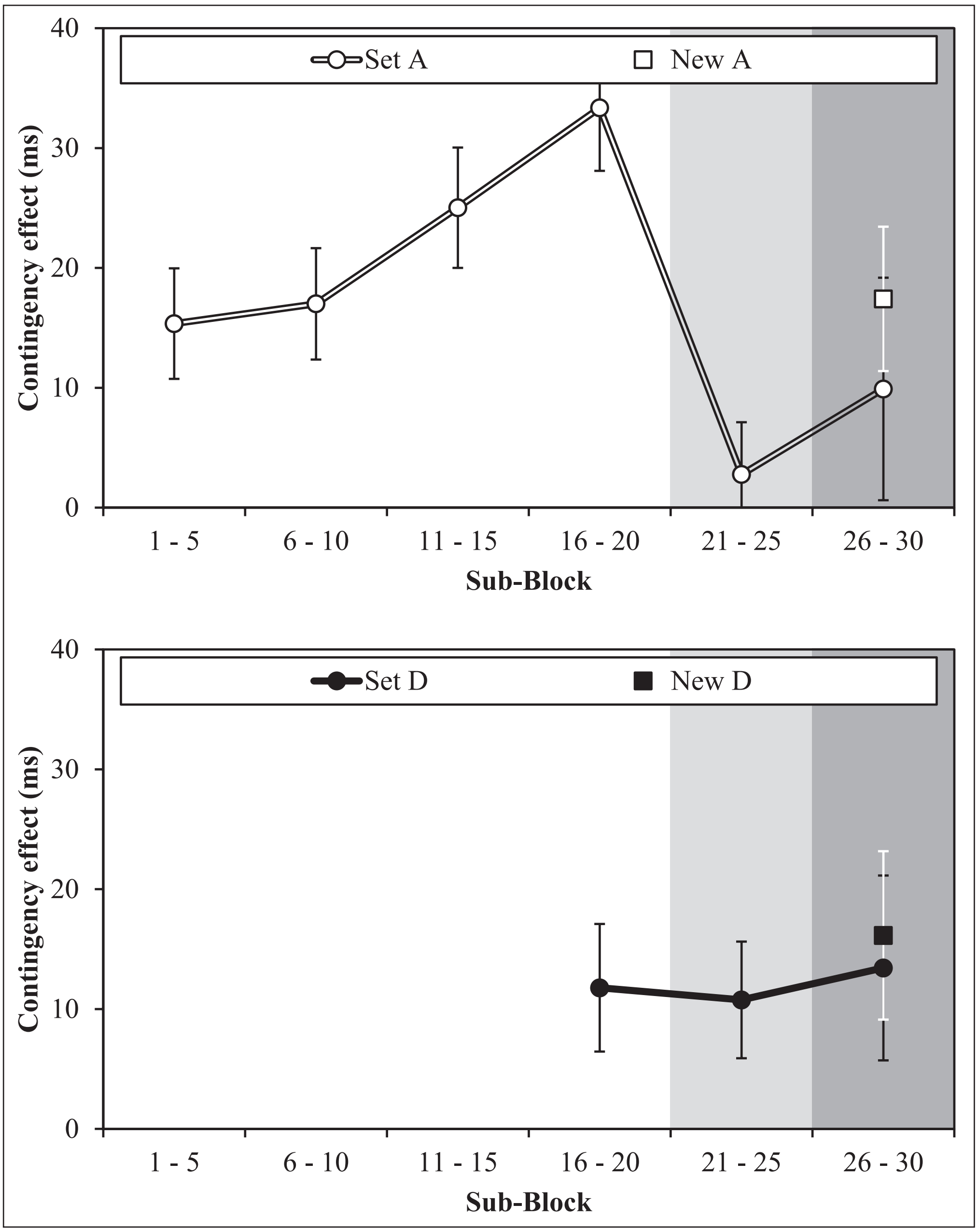

Figure 5: Experiment 2 contingency effect (low minus high contingency) as a function of mega-block (collapsed across sub-blocks, indicated in *x*-axis) with standard errors. Data are aggregated over the larger blocks. Light grey indicates the unlearning phase and dark grey indicates the counterconditioning phase. The single squares in the last block represent the low minus new contingency contrast.

Unlearning

Next, we compared Sets A and D during unlearning using a sub-block (21–25) by contingency (high vs. low) by set (A vs. D) ANOVA. The contingency effect was only marginal during unlearning, $F(1,59) = 3.103$, $MSE = 4003$, $p = .083$, $\eta_p^2 = .05$. There was a main effect of sub-block, $F(1,59) = 6.113$, $MSE = 6745$, $p = .016$, $\eta_p^2 = .09$, hinting at fatigue. Critically, there was no difference in the contingency effect between sets, $F(1,59) = 2.432$, $MSE = 2350$, $p = .124$, $\eta_p^2 = .04$, $BF_{01} = 4.72$, with the difference between high (mean = 610 ms, *SE* = 8) and low contingency trials (mean = 612 ms, *SE* = 7) in Set A similar to that in Set D (mean = 606 ms, *SE* = 9 and mean = 617 ms, *SE* = 8, respectively). In addition, the contingency effect was robustly smaller for Set A stimuli during unlearning relative to the preceding acquisition phase, $F(1,59) = 22.060$, $MSE = 676$, $p < .001$, $\eta_p^2 = .27$. These latter two findings are again inconsistent with the eventually-stable-habit view.

Counterconditioning

For the counterconditioning phase, we again begin with a block (26–30) by contingency (old vs. new) by set (A vs. D) ANOVA. Again, there was no main effect of contingency, $F(1,59) = 0.780$, $MSE = 5259$, $p = .375$, $\eta_p^2 = .01$ (note again that this is not a test of learning, but rather a comparison of old vs. new learning). In contrast to Experiment 1, however, the contingency by set interaction was not significant, $F(1,59) = 0.347$, $MSE = 3756$, $p = .558$, $\eta_p^2 < .01$, $BF_{01} = 9.64$, as illustrated in **Figure 6**. Relative to low contingency trials (mean = 629, *SE* = 10 for Set A and mean = 623, *SE* = 10 for Set D), the contingency effect (across sets) was significant for the old contingency (mean = 619, *SE* = 10 for Set A and mean = 610, *SE* = 10 for Set D) trials (i.e., old high vs. low), $F(1,59) = 4.501$, $MSE = 8898$, $p = .038$, $\eta_p^2 = .07$, and for the new contingency (mean = 612, *SE* = 8 for Set A and mean = 607, *SE* = 8 for Set D) trials (i.e., new high vs. low), $F(1,59) = 12.257$, $MSE = 5729$, $p < .001$, $\eta_p^2 = .17$, indicating both some preservation of the old contingency and acquisition of the new one.

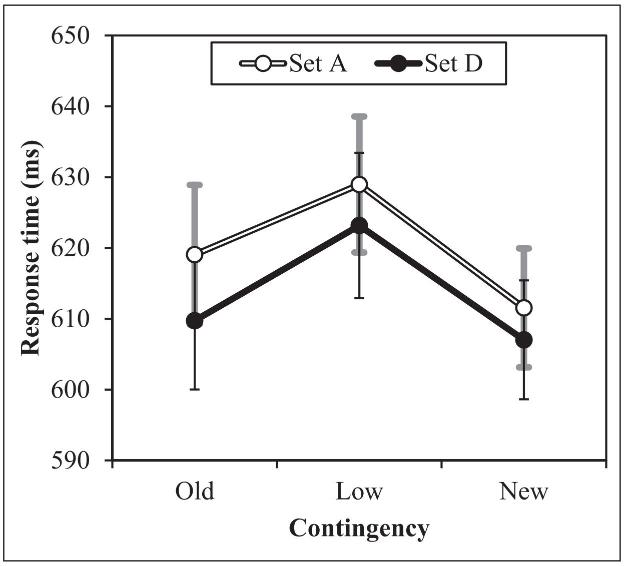

Figure 6: Experiment 2 mean response times (collapsed across sub-block) with standard errors for old, new, and low contingency items for the overtrained (Set A) and recently-introduced (Set D) stimuli during counterconditioning.

Discussion

Like Experiment 1, Experiment 2 produced some results indicating an overtraining advantage for Set A stimuli, but this time only in the comparison of the contingency effect for the just-introduced Set D stimuli relative to the overtrained Set A stimuli. In particular, there was a larger contingency effect for Set A. However, there was again no difference between the two sets during unlearning. Further, there was evidence of both a preservation of the old contingency and acquisition of the new contingency during the counterconditioning phase. Unlike Experiment 1, however, this did not seem to be influenced markedly by set. The old contingency effect does, on its own, indicate a persisting influence of older experiences: Even after 180 trials of unlearning and a subsequent introduction of an opposing contingency during counterconditioning, the originally-trained regularity continued to influence behaviour. However, even extensive overtraining (i.e., for Set A) did not seem to altogether prevent acquisition of a new contingency and this is not consistent with the strong view that heavy overtraining prevents acquisition of new knowledge.

Experiment 3

Results from Experiments 1 and 2 both provide some hints of an effect of overtraining, but not evidence for an overwhelming effect. For instance, the original contingency for Set A stimuli did not "stubbornly persist" through unlearning, instead reducing substantially as with the Set D stimuli. Further, Experiment 2 did not replicate the interaction between set (A vs. D) and contingency (old high vs. new high), instead showing significant effects for both new and old high contingencies (relative to low contingency) across sets, indicating both persistence of old contingency knowledge (inconsistent with the recent-events-matter-most scenario) and acquisition of new contingency knowledge (inconsistent with the eventually-stable-habits scenario). This might indicate that the amount of initial training is less important for persistence of the old contingency through counterconditioning than Experiment 1 suggested. Alternatively, it might be that the shorter overtraining phase was responsible for the lack of an interaction in Experiment 2. We therefore decided to run a third experiment as a conceptual replication of Experiment 1, that is, with a session of two separate days, but with an even longer training phase.

Method

Participants

Fifty Ghent University undergraduates participated in the study in two sessions in exchange for €10, as in Experiment 1. Using the same exclusion criteria as the prior experiments, no participants were excluded.

Apparatus, Design, and Procedure

The apparatus, design, and procedure of the current experiment were identical to Experiment 1 with the following exceptions. There were six mega-blocks per day instead of four, but we again excluded Set C stimuli and ran nine mega-blocks of A+B learning (i.e., 6 on Day 1 and 3 on Day 2), followed by the same A+D learning, unlearning, and counterconditioning mega-blocks as in the prior two experiments (2160 trials total).

Results

Data were analysed in the same manner as in the prior experiments. The contingency effect as a function of the larger blocks is presented in **Figure 7**. The means and standard errors for response times and errors are presented in Table A3 in the Appendix (see also the R scripts).

Sets A and D

First, we compared the overtrained Set A with a newly-introduced Set D using a sub-block (46–50) by contingency (high vs. low) by set (A vs. D) ANOVA. The main effect of contingency was significant, $F(1,49) = 28.982$, $MSE = 4675$, $p < .001$, $\eta_p^2 = .37$. This contingency effect interacted with set, $F(1,49) = 7.061$, $MSE = 3667$, $p = .011$, $\eta_p^2 = .13$, again indicating larger contingency effects for the overtrained Set A stimuli than for the new Set D stimuli. The contingency effect was significant for Set A (mean = 560 ms, SE = 8 and mean = 593 ms, SE = 9 for

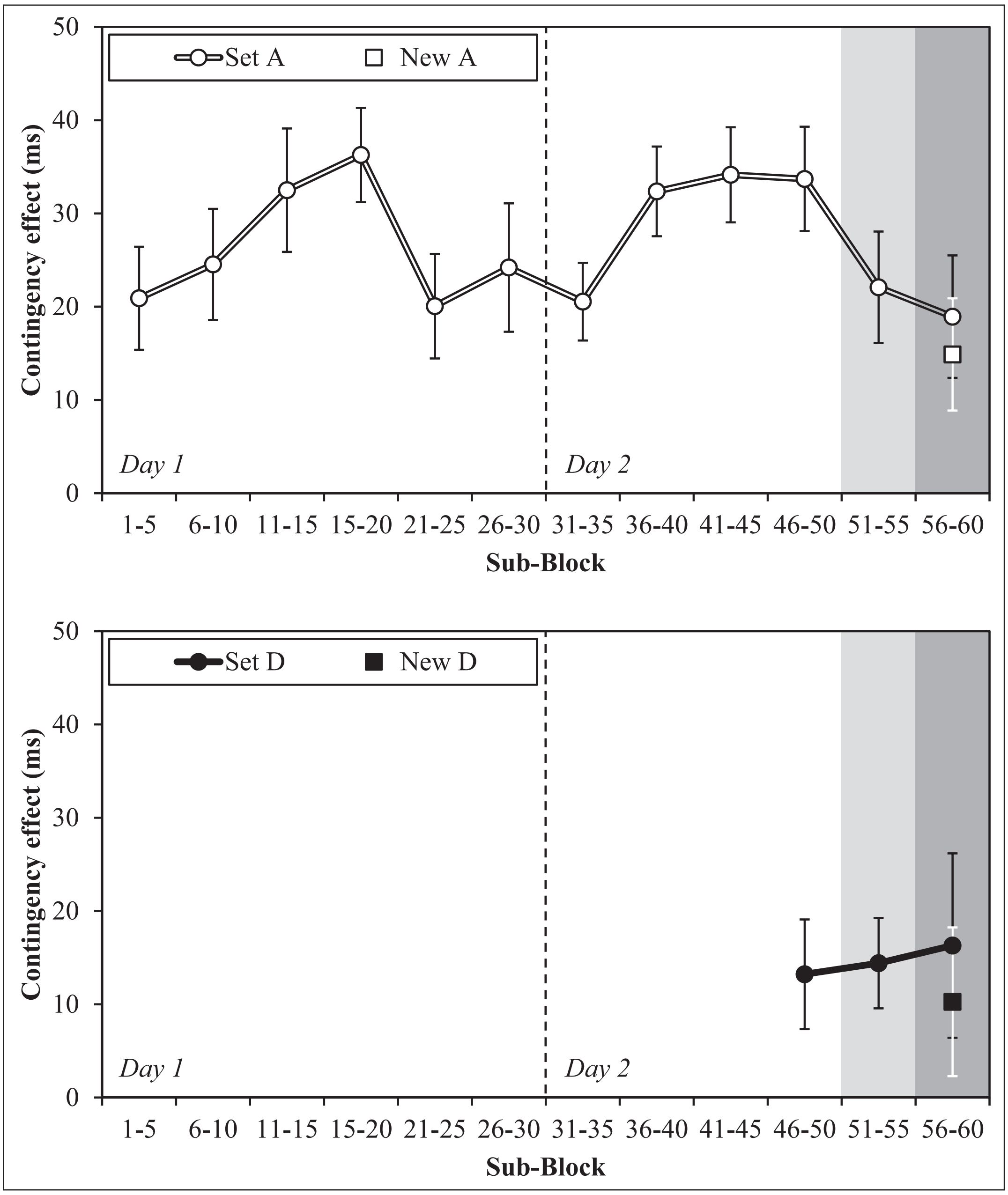

Figure 7: Experiment 3 contingency effect (low minus high contingency) as a function of block (sub-blocks indicated in *x*-axis) with standard errors. Data are aggregated over the mega-blocks. Light grey indicates the unlearning phase and dark grey indicates the counterconditioning phase. The single squares in the last sub-block represent the low minus new high contingency contrast.

high and low contingency, respectively), $F(1,49) = 33.698$, $MSE = 4152$, $p < .001$, $\eta_p^2 = .41$, and for Set D (mean = 574 ms, $SE = 8$ and mean = 587 ms, $SE = 8$, respectively), $F(1,49) = 5.121$, $MSE = 4190$, $p = .028$, $\eta_p^2 = .09$.

Unlearning

Next, we compared Sets A and D during unlearning using a block (51–55) by contingency (high vs. low) by set (A vs. D) ANOVA. The contingency effect was significant, $F(1,49) = 19.746$, $MSE = 4235$, $p < .001$, $\eta_p^2 = .29$, indicating persistence of the initial contingency during unlearning. There was a marginal main effect of set, $F(1,49) = 3.925$, $MSE = 2862$, $p = .053$, $\eta_p^2 = .07$, indicating overall faster responses for Set A stimuli, consistent with an overall benefit for overtrained stimuli. However, there was no difference in the contingency effect between sets, $F(1,49) = 0.846$, $MSE = 3177$, $p = .362$, $\eta_p^2 = .02$, $BF_{01} = 7.27$, though with a numerical trend in the direction that the eventually-stable-habit account would predict (mean = 565 ms, $SE = 8$ and mean = 587 ms, $SE = 7$ for high and low contingency in Set A, and mean = 575 ms, $SE = 8$ and mean = 590 ms, $SE = 7$ in Set D). The contingency effect for Set A was only non-significantly reduced in the unlearning phase relative to the preceding acquisition phase, $F(1,49) = 1.983$, $MSE = 891$, $p = .165$, $\eta_p^2 = .04$.

Counterconditioning

For the counterconditioning phase, we again begin with a sub-block (56–60) by contingency (old high vs. new high) by set (A vs. D) ANOVA. Again, there was no main effect of contingency, $F(1,49) = 0.329$, $MSE = 5955$, $p = .569$, $\eta_p^2 < .01$ (note again that this is not a test of the contingency effect proper, but the new versus old contingency). Again, the contingency by set interaction was not significant, $F(1,49) = 0.322$, $MSE = 3860$, $p = .573$, $\eta_p^2 < .01$, as illustrated in **Figure 8**, $BF_{01} = 9.30$. There was a significant sub-block by contingency interaction, $F(1,49) = 6.134$, $MSE = 6842$, $p = .017$, $\eta_p^2 = .11$, indicating a tendency for the old contingency effect to decrease and the new contingency effect to increase across sub-blocks. Relative to low contingency trials (mean = 591 ms, $SE = 9$ in Set A and mean = 589 ms, $SE = 9$ in Set D), the contingency effect (across sets) was significant for the old contingency (mean = 572 ms, $SE = 9$ in Set A and mean = 573 ms, $SE = 9$ in Set D) trials (i.e., old vs. low), $F(1,49) = 6.087$, $MSE = 10300$, $p = .017$, $\eta_p^2 = .11$, and for the new contingency (mean = 576 ms, $SE = 8$ in Set A and mean = 579 ms, $SE = 7$ in Set D) trials (i.e., new vs. low), $F(1,49) = 8.001$, $MSE = 5310$, $p = .007$, $\eta_p^2 = .14$.

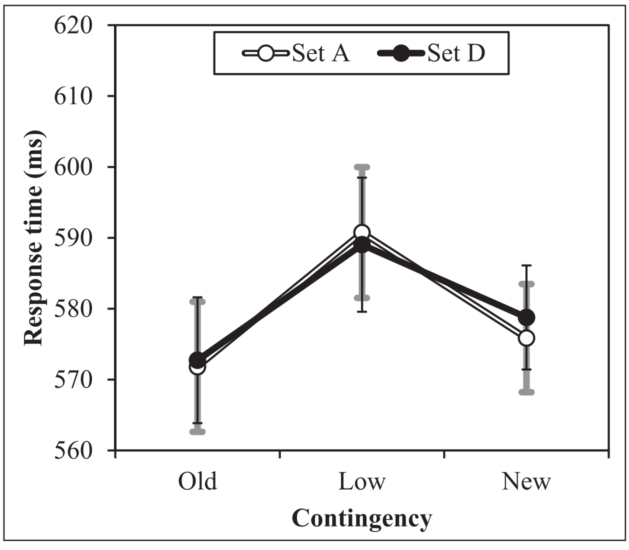

Figure 8: Experiment 3 mean response times (aggregated across sub-blocks) with standard errors for old high, new high, and low contingency items for the overtrained (Set A) and recently-introduced (Set D) stimuli during counterconditioning.

Discussion

As in the prior two experiments, Experiment 3 produced some results consistent with a benefit from overtraining. In particular, contingency effects were larger for Set A stimuli relative to Set D during initial acquisition of Set D contingencies. Similarly, response times were overall faster to Set A stimuli during unlearning. However, learning of the new contingency was observed for both sets of stimuli during counterconditioning. There was, nevertheless, again a persistence of the old contingency in the counterconditioning phase, which did not differ notably between the overtrained and recently-acquired stimuli. Together, the results again suggest a very strong influence of recent experiences, including for overtrained stimuli, as the recent-events-matter-most hypothesis would suggest. However, the results again suggest that learning is not *completely* “myopic” to only very recent experiences.

General Discussion

In the present series of experiments, we asked whether overlearning of contingencies either over two days (Experiments 1 and 3) or a single longer session (Experiment 2) would lead to a more stable learning effect, resistant to unlearning or counterconditioning, which we referred to as the eventually-stable-habit view. Alternatively, we considered the idea that learning might be rather “myopic” to recent events, whereby “habits” are maintained merely due to the continued repetition of recently-executed behaviours. This alternative view would suggest no observable differences between overtrained and recently-acquired contingencies. Our data are consistent with a more intermediate view: both lasting influences of older experiences and marked sensitivity to recent ones.

Some influences of older events were clearly observed. For instance, in all three experiments there was a larger contingency effect and/or an overall main effect speeding for Set A (overtrained) stimuli relative to newly-introduced Set B or C stimuli. This is also consistent with past findings that contingency effects, while acquired quite quickly, do tend to slowly increase with further training (Schmidt & De Houwer, 2016b). Similarly, contingency effects remained significant within the 180-trial unlearning phase, indicating persistence of a contingency when the regularity no longer applied. On the other hand, the contingency was not "stubbornly resistant" to the extent that no reductions in the learning effect were observed.

Perhaps most interesting was the newly-introduced counterconditioning phase. In all three experiments, the trend was for persistence of the old contingency effect, albeit in a reduced form, despite the preceding 180-trial unlearning phase and the introduction of a new, competing contingency. There was also a trend for acquisition of the *new* contingency effect in the counterconditioning phase, however. In Experiment 1, there was a significant interaction between the old and new contingency effects. In particular, while the new contingency effect was not robust (though numerically trending) and the new contingency effect was significant for the recently-introduced stimuli (Set D), the reverse was true for overlearned stimuli: The old contingency effect persisted and the newly-changed contingency was non-significant (though again trending in the correct direction). This pattern did not replicate in Experiments 2 and 3, however, where both the old and new contingencies influenced performance regardless of set. The reason for this discrepancy is uncertain. The significant interaction in Experiment 1 could have been a Type 1 error. Alternatively, a true but very small effect might exist that was detected only in one experiment. Indeed, as a general trend "hints" in favour of some overtraining effect were rather systematically observed in our experiments, but only some findings were significant and no overtraining effects were overwhelmingly large. Globally, then, the results are again not consistent with the strong idea that overtrained stimuli are inflexibly resistant to new learning: Some lasting influence of older events is observed, but new contingencies are picked up even for overtrained stimuli.

It is further noteworthy that there were a number of contrasts between overtrained and recently-acquired stimuli that were performed (i.e., two main ones per phase if we consider both the main effect of set and the interaction between set and contingency). Only some of these came out as significant and most of the observed differences were not substantial. We also note that we did not make corrections for multiple comparisons and there was clearly some noise and inconsistencies across the many comparisons, so more targeted replications of specific findings seem warranted. However, we do note that all significant cross-set comparisons were in the direction predicted by the eventually-stable-habit view. Still, any lasting influences from overtraining seem rather underwhelming. It is perhaps important to stress that these findings were not, however, unclear, as the results revealed significant effects that are inconsistent with each of the "extreme" positions we discussed in the Introduction. For instance, the eventually-stable-habits view clearly should have predicted, for overtrained stimuli: (a) no reduction in contingency effects with unlearning (but this was observed), (b) no acquisition of new contingencies during counterconditioning (but this was also observed). Similarly, the results also are not consistent with the extreme view that learning is myopic to only very recently occurring events, as the persistence of the old contingency through unlearning and counterconditioning clearly demonstrates. Thus, robust effects argue against both extreme views and therefore suggest that a more moderate view is necessary.

Collectively, the results of the present three experiments are consistent with the idea that both frequency and recency influence the quality of representations (in this case, representations of contingency; see Moors, 2016). Our results might be coherently explained by one high learning rate memory mechanism (e.g., Logan, 1988; Schmidt et al., 2016). According to this view, the individual impact of a given experience on current behaviour is related to how long ago the past experience occurred. Recently-experienced events have a particularly potent influence on behaviour, whereas older and older memory traces have increasingly smaller (but non-zero) influences on current behaviour. This notion has often been referred to as the *power law of practice* (Logan, 1988; Newell & Rosenbloom, 1981). The actual form of acquisition may be exponential (Heathcote, Brown, & Mewhort, 2000; Myung, Kim, & Pitt, 2000), which, averaged across participants, appears more like a power function. In any case, the rough notion is illustrated in **Figure 9**. If the participant is, for instance, currently partway through a counterconditioning phase, then learning of a "new" contingency can be explained by the potent influence of recent events. However, there will also be continued influence of older events from the prior unlearning and acquisition phases, which can explain the persistence of an old contingency. Differences between overtrained (Set A) and recently-acquired (Set D) stimuli can be explained by the extra traces encountered for the former stimuli. The reason for relatively weak overtraining effects can similarly be explained by the decay of much older memory traces. Similar notions have been forwarded to explain skill acquisition (Logan, 1988) and repetition priming (Grant & Logan, 1993; Logan, 1990).

This theoretical account may also explain why more notable differences were observed between Set A and Set D stimuli during acquisition, but less clearly during unlearning and counterconditioning. In particular, the extra learning experiences for Set A stimuli were more recent during acquisition, but more distant in time when the unlearning and subsequent counterconditioning phases eventually began. In that vein, somewhat larger effects of overlearning might be observable during counterconditioning if the counterconditioning directly follows acquisition.

Related to the above discussion, the present results might prove informative in constraining conceptual and modelling accounts of learning. For instance, the Parallel Episodic Processing (PEP) model of Schmidt and colleagues (2016) has been used to simulate a range of findings from the colour-word contingency learning paradigm, along with a number of other binding, timing, attentional, and control phenomena. Similar to what we described above, the PEP stores traces of individual events and similarity-based retrieval of these "exemplars" produces learning effects. It is already the case in the PEP model that recently-encoded events are more strongly retrieved from memory that older events. This allows the model to learn quickly and simulate (with the same mechanism) more transitory influences on behaviour, such as distracter-response binding effects (Frings, Rothermund, & Wentura,

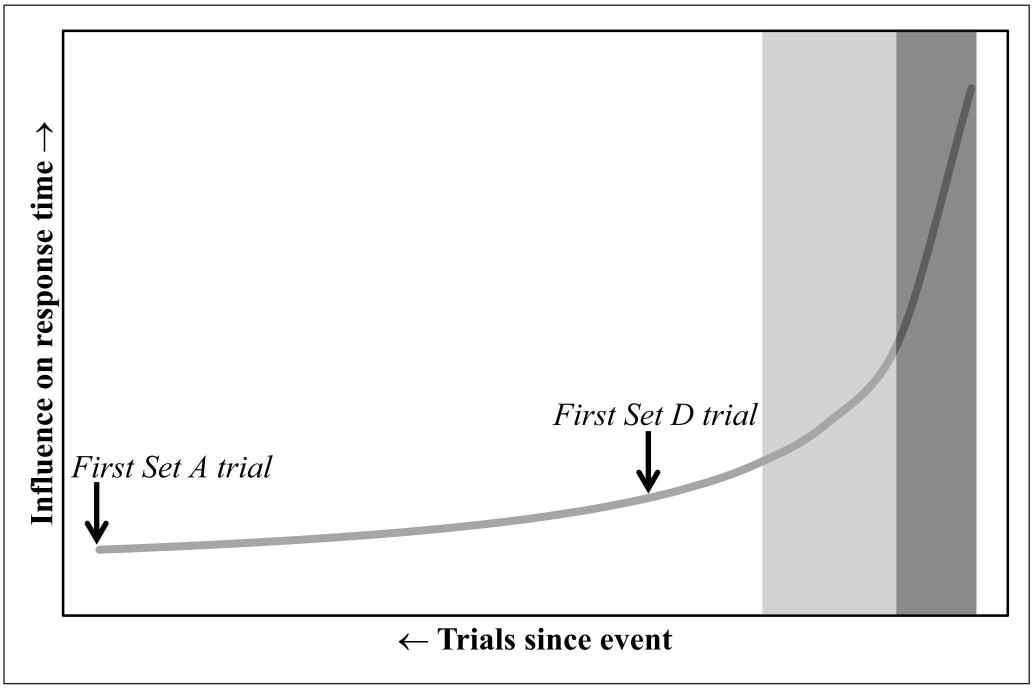

Figure 9: An illustration of a power law influence on behaviour. Recently-encoded events have a particularly strong influence on behaviour, whereas older and older memories have ever diminishing influences. In this example, the counterconditioning phase is marked in dark grey, the unlearning phase in light grey, and acquisition in white.

2007). Future modelling work, however, could aim to see whether the PEP (or other models) is able to simulate the both the lasting and dynamic adaptations to contingencies as observed in the present report.

The current research introduces interesting new avenues for future research. For instance, we observed that an originally-trained contingency that was acquired over two sessions on two days persisted through 180 unlearning trials and even through another 180 counterconditioning trials. But how much longer would this effect persist? Various permutations of the learning and counterconditioning phase lengths might clarify this issue further. Presumably, the new contingency will overwhelm the old contingency eventually, and the 180-trial counterconditioning phase in the current report may have simply been too short. Future work with longer counterconditioning phases might therefore explore whether an old contingency does eventually extinguish entirely. Similarly, it may be the case that the old contingency eventually does become stable and completely prevents acquisition of a new contingency. Future research might aim to extend the initial training over much longer durations (e.g., weeks) before introducing counterconditioning.

More globally, the present results might suggest that much of what we consider to be a habit is driven by recent experiences. If recent experiences are, indeed, more influential on the maintenance of an automatic behaviour than is typically assumed, then this might have interesting implications. Although certainly speculative, to change automatic behaviour (e.g., undesirable habits), it might suffice to force a change in a limited number of recent experiences. That is, the default response to a stimulus could change even if only the most recent experiences are different from many old ones. At the same time, our data show that old memory traces will persist to some degree, which could explain relapse of old automatic behaviours, including undesirable habits. Such relapse might be particularly likely as time elapses. In this case, events inconsistent with the original default response (e.g., counter-habitual behaviours) will lose their advantage of recency so that traces of old automatic behaviours can resurface. This is related to explanations for spontaneous recovery, that is, the re-emergence of a previously-extinguished behaviour (Briggs, 1954). It would thus be interesting to rerun our studies but add a delay between counterconditioning and a subsequent unlearning test phase. It might be the case that the old contingency resurfaces more strongly with delay (Pavlov, 1927).

Relatedly, two of our three experiments included two days of training. This was largely done with the aim of breaking up the lengthy training. However, the two-day training did involve an intermediate night of sleep reconsolidation. Future work might explore the role of consolidation more directly. Some existing work already suggests that consolidation does play some role in the strength of learning. For instance, in Geukes, Gaskell, and Zwitserlood (2015) trained participants with novel words and colour words as pairs and observed Stroop-like (or learning) effects when intermixed with colour word distracters. Without colour word distracters, however, the Stroop effect was only observed on Day 2 (i.e., after consolidation). Similarly, the role of sleep or simply time in consolidation could be explored (Lindsay & Gaskell, 2013).

Another factor of interest that might be interesting to explore in the current design is context dependency. It is likely that, in real life, automatic (e.g., habitual) behaviours have been emitted in many contexts whereas recent (e.g., counter-habitual) behaviours inconsistent with the original learning have been emitted in more restricted contexts (e.g., therapy). Also, from learning psychology we know that first learning is less context dependent than new learning (for a review, see Bouton, 2004). Hence, old memory traces might have more impact in new contexts than in those contexts specifically used for unlearning or counterconditioning. To explore this notion with the present materials, one might therefore imagine introducing contextual changes, for instance, after counterconditioning to see whether this leads to a re-emergence of the old contingency.

We also remind the reader that "habit" is inconsistently defined in the literature (see the Introduction). Thus, depending on how one defines a habit, our lengthy acquisition phase may or may not be considered sufficient to establish a habit. It is similarly not clear whether the incidental learning effects explored in the present report are due to goal-free stimulus-response learning or whether the learning also includes goal-directed actions. Independently of whether the absence of goals should be considered a defining feature of a habit (see De Houwer, 2019, for concerns with this perspective), it remains an interesting open question to explore in future research (e.g., by conducting similar studies with or without deliberate learning goals).

Appendix

Table A1: Experiment 1 response time and error means and standard errors.

Block	Contingency	Response Times				Error Percentages			
		Set A		Sets B–D		Set A		Sets B–D	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
1	High	612	12	607	11	5.0	0.9	5.2	1.2
	Low	609	16	603	13	7.0	1.6	9.6	1.8
2	High	601	12	591	11	5.9	1.0	5.0	1.0
	Low	614	14	606	14	8.1	1.5	5.5	1.4
3	High	598	12	601	13	6.9	1.2	4.6	0.9
	Low	617	15	614	15	6.5	1.6	5.9	1.5
4	High	595	10	597	13	4.3	0.8	4.9	1.1
	Low	621	15	625	16	6.2	1.4	4.8	1.2
5	High	598	12	610	12	4.3	0.9	6.5	1.1
	Low	617	12	621	12	4.4	1.2	9.5	1.9
6	High	592	12	611	11	6.1	1.1	6.1	1.0
	Low	606	14	640	17	5.7	1.5	6.7	1.6
7	High	612	12	617	13	6.1	1.0	4.5	0.9
	Low	637	13	636	17	7.7	1.5	7.0	1.5
8	High	615	14	617	14	6.6	1.0	7.6	1.0
	Low	632	17	634	15	6.6	1.4	7.4	1.9
9	High	631	15	622	15	6.9	1.2	3.6	1.0
	Low	673	17	647	16	7.0	1.3	10.8	1.8
10	High	609	13	619	14	5.9	1.1	5.7	1.1
	Low	648	15	630	14	7.3	1.7	7.4	1.7
11	High	611	14	607	13	7.0	0.9	6.1	1.1
	Low	633	15	628	16	7.0	1.6	7.8	1.5
12	High	617	13	609	11	5.7	1.2	5.8	1.0
	Low	662	16	635	15	7.8	1.6	8.8	2.1
13	High	607	13	619	14	6.0	1.1	5.7	1.1
	Low	631	12	644	14	8.0	1.9	7.4	1.6
14	High	618	13	616	11	6.3	1.1	5.9	1.1
	Low	632	16	637	14	7.7	1.5	7.7	1.6

(Contd.)

Block	Contingency	Response Times				Error Percentages			
		Set A		Sets B–D		Set A		Sets B–D	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
15	High	604	15	607	13	6.0	1.0	4.3	0.8
	Low	622	11	640	17	11.6	1.8	8.5	1.8
16	High	622	12	637	13	6.2	1.2	8.3	1.4
	Low	609	14	624	15	7.4	1.7	8.1	1.7
17	High	610	12	624	14	6.6	1.2	5.3	1.0
	Low	615	15	621	15	6.6	1.5	8.4	1.9
18	High	603	13	620	14	6.3	1.3	5.6	1.0
	Low	636	17	663	18	8.9	1.8	6.3	1.3
19	High	610	14	619	15	5.6	1.2	5.9	1.3
	Low	633	15	651	16	11.6	2.2	9.1	2.0
20	High	617	13	609	12	6.6	1.3	5.9	1.0
	Low	637	18	638	17	7.0	1.7	11.0	1.7
21	High	537	9	555	10	2.7	0.8	2.5	0.8
	Low	553	9	558	10	2.6	0.9	1.8	0.8
22	High	543	7	554	9	3.9	0.8	3.9	0.9
	Low	569	12	573	11	5.2	1.2	4.8	1.4
23	High	552	11	552	10	3.7	0.8	2.8	0.7
	Low	559	9	560	12	4.8	1.3	5.5	1.4
24	High	549	8	561	10	3.4	0.8	5.1	1.1
	Low	580	14	576	13	8.0	2.0	8.1	1.9
25	High	571	12	576	11	5.2	1.1	5.8	1.0
	Low	594	12	584	13	7.3	1.4	6.6	1.8
26	High	559	8	574	10	2.7	0.7	5.2	1.0
	Low	575	11	580	10	6.7	1.4	5.9	1.2
27	High	558	12	586	11	4.1	0.8	3.9	1.0
	Low	592	15	582	12	3.0	1.0	8.4	1.9
28	High	571	12	582	13	5.4	0.9	3.6	0.8
	Low	609	17	576	11	7.0	1.5	6.2	1.6
29	High	561	11	586	12	3.6	0.9	4.8	0.9
	Low	602	14	592	12	8.4	1.8	6.3	1.2
30	High	567	11	581	12	4.5	1.1	4.8	1.0
	Low	583	15	600	13	6.6	1.4	5.2	1.3
31	High	574	15	562	11	5.6	1.2	5.5	1.4
	Low	589	12	593	12	5.9	0.8	5.2	1.1
32	High	568	13	582	14	4.8	1.2	5.9	1.3
	Low	605	12	594	13	4.2	0.8	6.3	1.0
33	High	594	14	570	13	4.0	1.3	8.1	1.5
	Low	589	13	596	11	6.3	1.0	5.4	1.1
34	High	574	12	584	12	4.0	1.4	4.0	1.2
	Low	591	12	595	10	6.4	1.0	8.1	1.4

(Contd.)

Block	Contingency	Response Times				Error Percentages			
		Set A		Sets B–D		Set A		Sets B–D	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
35	High	584	15	593	15	4.4	1.4	5.2	1.3
	Low	597	12	596	13	7.0	1.0	7.0	1.4
36	Old	581	16	612	19	7.2	2.3	5.7	1.9
	Low	609	17	596	18	11.5	2.6	8.6	2.4
	New	604	13	590	11	5.7	1.0	6.2	1.1
37	Old	571	17	586	15	6.5	2.0	5.0	1.8
	Low	591	15	611	18	7.2	2.7	9.4	2.7
	New	599	14	589	12	6.2	1.3	4.7	1.0
38	Old	575	18	611	20	5.0	1.8	5.8	2.1
	Low	590	15	596	18	5.7	1.9	8.6	2.4
	New	588	14	574	12	5.8	0.9	8.5	1.3
39	Old	585	16	590	13	3.6	1.5	2.2	1.2
	Low	593	18	616	24	7.2	2.3	6.5	2.0
	New	588	12	582	10	6.8	1.0	6.3	1.1
40	Old	576	13	588	17	3.6	1.5	9.3	2.5
	Low	607	19	597	14	7.2	2.0	4.3	2.0
	New	588	11	573	11	6.3	1.2	4.5	0.9

Table A2: Experiment 2 response time and error means and standard errors.

Block	Contingency	Response Times				Error Percentages			
		Set A		Sets B–D		Set A		Sets B–D	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
1	High	604	11	599	11	2.3	0.7	4.4	0.8
	Low	618	12	605	11	4.5	1.2	4.0	1.2
2	High	590	9	600	10	2.7	0.5	3.9	0.9
	Low	607	12	609	10	3.1	0.9	3.1	0.9
3	High	598	10	599	11	3.4	0.8	3.0	0.5
	Low	617	12	604	11	3.4	1.0	2.3	0.8
4	High	603	10	597	10	4.0	0.9	3.0	0.7
	Low	625	12	592	10	5.1	1.2	5.1	1.2
5	High	604	10	600	10	4.0	0.8	4.7	0.8
	Low	612	10	619	13	4.0	0.9	7.3	1.3
6	High	584	9	580	9	2.6	0.6	4.4	0.7
	Low	607	12	609	10	5.1	1.1	5.1	1.1
7	High	609	11	605	10	4.4	0.9	4.7	0.8
	Low	613	12	606	12	7.4	1.2	4.8	1.2
8	High	611	12	605	10	3.5	0.6	3.7	0.8
	Low	631	11	616	10	5.2	1.1	4.8	1.2
9	High	587	9	603	10	3.8	0.8	5.5	0.9
	Low	609	11	633	11	7.3	1.4	4.8	1.1

(Contd.)

Block	Contingency	Response Times				Error Percentages			
		Set A		Sets B–D		Set A		Sets B–D	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
10	High	598	10	607	9	5.4	0.9	4.4	0.9
	Low	615	12	626	9	5.9	1.1	6.4	1.3
11	High	582	10	600	10	4.2	0.7	4.5	0.7
	Low	617	13	608	11	3.7	1.0	6.5	1.5
12	High	607	9	609	11	3.8	0.7	3.3	0.7
	Low	647	12	626	13	4.3	1.0	5.7	1.1
13	High	614	11	614	10	4.9	0.9	3.6	0.9
	Low	621	11	633	11	4.1	1.1	6.5	1.3
14	High	601	9	606	11	4.6	0.8	4.3	0.9
	Low	618	12	639	12	9.0	1.7	6.4	1.4
15	High	603	9	608	8	3.2	0.7	3.6	0.8
	Low	633	14	632	11	7.9	1.6	4.5	1.1
16	High	586	9	606	10	4.4	0.8	4.7	0.8
	Low	620	11	612	10	6.6	1.5	5.6	1.2
17	High	598	10	619	9	4.3	0.9	4.5	0.8
	Low	619	12	612	9	6.4	1.5	6.1	1.5
18	High	598	11	607	10	4.4	0.7	4.3	0.9
	Low	633	11	627	13	6.5	1.4	4.6	1.1
19	High	589	9	606	9	3.7	0.6	5.9	1.0
	Low	636	10	636	13	4.9	1.1	5.9	1.2
20	High	603	9	612	11	5.4	0.9	5.7	1.0
	Low	636	12	622	11	6.2	1.3	7.1	1.3
21	High	586	9	578	10	2.5	0.9	6.7	1.4
	Low	590	8	602	9	6.7	0.9	4.5	0.8
22	High	616	11	617	12	5.1	1.2	5.4	1.2
	Low	618	10	625	10	5.4	0.9	4.9	0.8
23	High	619	12	622	12	4.8	1.1	5.4	1.2
	Low	621	10	627	9	4.9	0.9	6.0	0.9
24	High	615	11	603	11	4.5	1.3	4.2	1.0
	Low	621	9	610	9	4.8	0.7	4.6	0.8
25	High	614	12	610	12	4.5	1.1	6.3	1.3
	Low	610	10	619	11	6.2	0.9	4.0	0.6
26	Old	623	19	598	13	3.3	1.3	5.0	1.5
	Low	611	14	614	13	3.9	1.4	5.0	1.5
	New	599	10	604	11	4.5	0.9	3.7	0.7
27	Old	612	15	606	14	6.1	1.7	5.0	1.5
	Low	647	17	643	16	6.6	1.7	7.2	1.9
	New	623	10	614	9	5.2	1.0	4.4	0.8
28	Old	626	18	615	18	6.1	1.8	3.9	1.6
	Low	624	14	624	16	6.4	1.9	4.4	1.7
	New	620	11	608	12	5.8	1.0	5.3	0.9

(Contd.)

Block	Contingency	Response Times				Error Percentages			
		Set A		Sets B–D		Set A		Sets B–D	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
29	Old	618	14	630	15	2.2	1.1	5.5	1.8
	Low	615	14	624	15	9.4	2.2	5.0	1.5
	New	610	9	607	10	4.7	0.8	4.2	0.7
30	Old	609	13	595	14	4.4	1.5	3.9	1.4
	Low	643	15	605	15	5.0	1.5	3.9	1.6
	New	607	11	604	10	6.5	1.0	4.6	0.7

Table A3: Experiment 3 response time and error means and standard errors.

Block	Contingency	Response Times				Error Percentages			
		Set A		Sets B–D		Set A		Sets B–D	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
1	High	592	11	606	11	2.6	0.7	3.8	0.9
	Low	606	13	630	15	4.7	1.2	4.7	1.3
2	High	590	10	587	10	4.0	0.9	4.3	0.9
	Low	608	12	595	11	6.5	1.6	5.9	1.2
3	High	587	11	590	11	3.3	0.8	4.3	0.8
	Low	602	13	608	12	5.1	1.3	6.7	1.4
4	High	593	11	589	11	4.6	1.0	4.6	1.0
	Low	615	13	623	13	6.7	1.4	6.8	1.4
5	High	596	11	600	11	5.3	0.8	3.3	0.7
	Low	633	12	614	11	5.2	1.1	8.1	1.4
6	High	589	11	577	9	3.1	0.9	3.5	0.9
	Low	607	13	609	13	4.8	1.2	5.1	1.2
7	High	594	10	591	12	4.8	0.9	4.6	1.0
	Low	612	14	621	12	4.8	1.3	6.8	1.5
8	High	588	11	583	10	5.8	1.2	5.2	1.0
	Low	609	13	618	13	5.5	1.3	4.4	1.2
9	High	586	9	584	10	4.4	0.9	5.4	0.9
	Low	631	14	600	12	6.4	1.5	5.5	1.2
10	High	612	11	598	10	5.2	1.0	5.8	1.0
	Low	636	14	633	16	7.4	1.7	5.0	1.2
11	High	584	12	576	12	4.1	0.8	5.8	1.1
	Low	600	13	620	12	5.8	1.2	6.7	1.6
12	High	583	9	595	11	5.0	1.0	4.3	0.9
	Low	619	14	631	16	6.5	1.4	8.1	1.6
13	High	590	10	594	11	5.4	0.8	5.7	1.0
	Low	643	15	603	12	10.4	1.8	5.1	1.2
14	High	593	11	591	9	4.7	1.0	3.3	0.7
	Low	636	13	649	16	5.8	1.3	7.3	1.4
15	High	593	10	590	11	5.6	1.0	5.5	1.1
	Low	613	10	631	14	7.6	1.5	7.5	1.5

(Contd.)

Block	Contingency	Response Times				Error Percentages			
		Set A		Sets B–D		Set A		Sets B–D	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
16	High	583	9	589	11	4.5	0.9	4.9	0.9
	Low	590	11	602	13	7.2	1.2	7.5	1.5
17	High	601	9	593	11	3.3	0.6	5.0	0.9
	Low	646	16	637	10	5.8	1.4	6.5	1.4
18	High	603	11	597	11	2.9	0.7	5.2	0.8
	Low	631	13	626	13	6.1	1.3	8.5	1.3
19	High	589	12	592	9	5.3	0.9	4.7	1.0
	Low	641	13	605	11	5.8	1.5	7.2	1.3
20	High	586	9	594	9	4.9	0.9	5.8	1.0
	Low	638	13	621	12	4.7	1.2	8.4	1.7
21	High	578	12	581	9	2.6	0.7	5.5	1.1
	Low	605	12	608	13	4.8	1.1	7.1	1.5
22	High	584	11	591	11	4.4	0.8	5.3	0.9
	Low	604	16	614	14	5.9	1.3	6.8	1.4
23	High	592	10	589	11	5.0	1.1	5.1	1.0
	Low	606	11	621	14	9.1	1.8	6.1	1.3
24	High	590	11	594	9	3.9	0.9	5.1	1.1
	Low	604	12	621	15	6.2	1.3	8.5	1.6
25	High	581	11	588	10	5.4	0.9	4.8	1.0
	Low	604	10	612	11	5.8	1.2	5.1	1.3
26	High	568	9	577	11	4.5	0.9	7.5	1.0
	Low	587	12	582	11	3.8	1.0	6.5	1.4
27	High	591	11	583	9	5.1	1.0	6.3	1.3
	Low	601	10	621	15	6.4	1.6	5.1	1.4
28	High	579	11	590	11	5.0	1.0	5.7	1.0
	Low	607	12	599	13	7.5	1.5	5.7	1.5
29	High	580	9	602	11	4.9	0.9	4.8	1.0
	Low	629	13	612	13	8.7	1.7	8.0	1.9
30	High	603	11	592	10	5.2	1.1	4.1	0.9
	Low	621	15	631	15	9.1	1.6	7.7	1.7
31	High	524	8	522	8	4.9	1.0	2.6	0.7
	Low	540	8	544	11	4.8	1.1	5.8	1.2
32	High	520	8	528	8	3.2	0.8	3.6	0.8
	Low	550	10	556	10	6.7	1.5	5.8	1.4
33	High	545	9	530	8	3.9	0.7	4.6	1.0
	Low	554	9	547	9	4.1	1.0	4.7	1.4
34	High	538	8	543	11	2.0	0.6	2.8	0.7
	Low	552	13	566	11	8.1	1.6	4.1	1.0
35	High	537	8	540	9	3.0	0.8	3.1	0.7
	Low	567	10	581	12	4.1	1.1	7.4	1.4
36	High	539	10	535	8	3.7	0.9	3.4	0.8
	Low	561	11	558	11	6.1	1.3	5.8	1.3

(Contd.)

Block	Contingency	Response Times				Error Percentages			
		Set A		Sets B–D		Set A		Sets B–D	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
37	High	542	9	545	9	4.6	1.0	4.4	1.2
	Low	582	11	563	10	5.4	1.3	3.8	1.4
38	High	543	9	539	9	2.8	0.8	3.6	0.8
	Low	576	12	566	10	7.0	1.7	7.4	1.7
39	High	539	9	542	8	3.7	0.8	4.1	1.0
	Low	572	11	570	10	5.7	1.8	5.4	1.3
40	High	543	10	546	10	3.0	0.7	4.8	0.9
	Low	576	12	577	12	4.7	1.3	8.8	1.8
41	High	535	9	520	9	4.1	0.8	6.0	1.2
	Low	566	12	582	12	7.1	1.3	5.4	1.3
42	High	554	9	557	8	2.9	0.8	4.1	1.0
	Low	582	12	582	12	7.4	1.5	6.4	1.4
43	High	557	9	555	8	4.1	0.7	3.9	0.9
	Low	589	10	569	13	3.7	1.0	7.2	1.5
44	High	553	9	545	9	2.8	0.7	4.3	0.9
	Low	590	11	585	11	5.4	1.2	6.1	1.3
45	High	549	10	555	11	3.6	0.8	3.8	0.9
	Low	594	13	596	12	9.1	1.9	7.4	1.6
46	High	563	11	574	10	2.6	0.7	5.9	1.0
	Low	590	11	585	11	7.8	1.5	8.4	1.5
47	High	543	9	577	9	5.3	1.2	5.3	1.1
	Low	588	12	588	13	9.1	1.5	2.8	1.0
48	High	564	11	578	10	3.6	0.8	6.1	1.1
	Low	601	13	586	12	7.0	1.6	5.7	1.3
49	High	562	10	565	10	4.5	0.8	4.6	0.9
	Low	601	15	578	12	8.6	1.6	6.4	1.3
50	High	569	10	577	10	5.1	1.0	4.5	0.9
	Low	588	11	598	12	8.8	1.5	7.8	1.5
51	High	562	12	568	14	4.1	1.2	5.5	1.4
	Low	584	9	585	10	7.5	1.4	6.6	1.2
52	High	566	12	577	13	5.1	1.3	4.4	1.2
	Low	585	8	596	10	5.4	0.9	6.0	1.3
53	High	566	10	572	9	2.7	0.9	7.2	1.9
	Low	595	11	593	10	7.5	1.2	5.2	1.0
54	High	564	12	579	14	4.4	1.5	5.1	1.5
	Low	586	11	586	10	7.9	1.1	7.5	1.3
55	High	570	12	581	12	4.4	1.1	5.4	1.2
	Low	585	10	593	9	6.5	1.1	4.9	1.1
56	Old	541	13	562	14	3.6	1.6	8.0	2.4
	Low	594	14	562	13	6.6	2.1	5.3	2.0
	New	573	9	573	10	6.5	1.0	5.5	1.1

(Contd.)

Block	Contingency	Response Times				Error Percentages			
		Set A		Sets B–D		Set A		Sets B–D	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
57	Old	574	14	578	15	2.6	1.3	3.3	1.4
	Low	593	15	597	18	7.6	2.3	6.6	2.1
	New	581	9	599	10	6.0	1.2	5.8	1.2
58	Old	580	15	566	15	2.7	1.6	7.6	2.1
	Low	574	12	617	18	5.3	2.2	4.0	1.5
	New	584	11	573	10	6.5	1.0	4.8	0.9
59	Old	592	19	590	14	6.6	1.9	9.3	2.3
	Low	590	18	585	15	9.6	2.4	8.0	2.4
	New	569	9	572	9	5.5	1.1	6.1	1.0
60	Old	588	17	574	15	6.6	1.9	4.6	1.6
	Low	608	16	584	13	5.9	1.8	6.0	2.1
	New	571	11	578	10	5.3	1.1	7.5	1.2

Data Accessibility Statement
The experiment files, along with the raw data, participant averaged data, and R scripts are available on the Open Science Framework (https://osf.io/7fwae/).

Funding Information
This research was supported by the French "Investissements d'Avenir" program, project ISITE-BFC (contract ANR15-IDEX-0003) to James R. Schmidt, Grant BOF09/01M00209 of Ghent University to Jan De Houwer, and Research Program G073317N of the Research Foundation – Flanders (FWO), and the Research Fund of KU Leuven (C14/17/047) to Agnes Moors.

Competing Interests
The authors have no competing interests to declare.

Author Contributions
All authors contributed to conception and design, revised the article, and approved the submitted version for publication. JRS contributed to data acquisition, analyzed and interpreted the data, and drafted the article.

References

Allan, L. G. (2005). Learning of contingent relationships. *Learning & Behavior*, *33*, 127–129. DOI: https://doi.org/10.3758/BF03196057

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. DOI: https://doi.org/10.1037/a0028458

Beckers, T., De Houwer, J., & **Matute, H.** (2007). Editorial: Human contingency learning. *Quarterly Journal of Experimental Psychology*, *60*, 289–290. DOI: https://doi.org/10.1080/17470210601000532

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory*, *11*, 485–494. DOI: https://doi.org/10.1101/lm.78804

Briggs, G. E. (1954). Acquisition, extinction, and recovery functions in retroactive inhibition. *Journal of Experimental Psychology*, *47*, 285–293. DOI: https://doi.org/10.1037/h0060251

Carlson, K. A., & **Flowers, J. H.** (1996). Intentional versus unintentional use of contingencies between perceptual events. *Perception & Psychophysics*, *58*, 460–470. DOI: https://doi.org/10.3758/BF03206821

De Houwer, J. (2019). On how definitions of habits can complicate habit research. *Frontiers in Psychology*, *10*, Article 2642. DOI: https://doi.org/10.3389/fpsyg.2019.02642

De Houwer, J., Tanaka, A., Moors, A., & **Tibboel, H.** (2018). Kicking the habit: Why evidence for habits in humans might be overestimated. *Motivation Science*, *4*, 50–59. DOI: https://doi.org/10.1037/mot0000065

Dickinson, A. (1985). Actions and habits: The development of behavioural autonom. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *308*, 67–78. DOI: https://doi.org/10.1098/rstb.1985.0010

Ellis, N. R., & **Dulaney, C. L.** (1991). Further evidence for cognitive inertia of persons with mental retardation. *American Journal on Mental Retardation*, *95*, 613–621.

Evans, J. S. T., & **Stanovich, K. E.** (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*, 223–241. DOI: https://doi.org/10.1177/1745691612460685

Forrin, N. D., & **MacLeod, C. M.** (2017). Relative speed of processing determines color-word contingency learning. *Memory & Cognition*, *45*, 1206–1222. DOI: https://doi.org/10.3758/s13421-017-0721-4

Frings, C., Rothermund, K., & **Wentura, D.** (2007). Distractor repetitions retrieve previous responses to targets. *Quarterly Journal of Experimental*

Psychology, *60*, 1367–1377. DOI: https://doi.org/10.1080/17470210600955645

Geukes, S., Gaskell, M. G., & **Zwitserlood, P.** (2015). Stroop effects from newly learned color words: Effects of memory consolidation and episodic context. *Frontiers in Psychology*, *6*, Article 278. DOI: https://doi.org/10.3389/fpsyg.2015.00278

Giesen, C., Schmidt, J. R., & **Rothermund, K.** (2020). The law of recency: An episodic stimulus-response retrieval account of habit acquisition. *Frontiers in Psychology*, *10*, Article 2927. DOI: https://doi.org/10.3389/fpsyg.2019.02927

Grant, S. C., & **Logan, G. D.** (1993). The loss of repetition priming and automaticity over time as a function of degree of initial learning. *Memory & Cognition*, *21*, 611–618. DOI: https://doi.org/10.3758/BF03197193

Gul, A., & **Humphreys, G. W.** (2015). Practice and colour-word integration in Stroop interference. *Psicológica*, *36*, 37–67.

Heathcote, A., Brown, S., & **Mewhort, D. J. K.** (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review*, *7*, 185–207. DOI: https://doi.org/10.3758/BF03212979

Heyes, C. M., & **Dickinson, A.** (1990). The intentionality of animal action. *Mind and Language*, *5*, 87–104. DOI: https://doi.org/10.1111/j.1468-0017.1990.tb00154.x

Lemercier, C. (2009). Incidental learning of incongruent items in a manual version of the Stroop task. *Perceptual and Motor Skills*, *108*, 705–716. DOI: https://doi.org/10.2466/pms.108.3.705-716

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. DOI: https://doi.org/10.1016/j.actpsy.2015.12.009

Lewicki, P. (1985). Nonconscious biasing effects of single instances on subsequent judgments. *Journal of Personality and Social Psychology*, *48*, 563–574. DOI: https://doi.org/10.1037/0022-3514.48.3.563

Lewicki, P. (1986). Processing information about covariations that cannot be articulated. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 135–146. DOI: https://doi.org/10.1037/0278-7393.12.1.135

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. DOI: https://doi.org/10.1037/xlm0000436

Lin, Y.-H. (2015). *Characterizing colour-word contingency learning* (Doctoral dissertation). Retrieved from http://hdl.handle.net/10012/9279. DOI: https://doi.org/10.1037/e528942014-039

Lindsay, S., & **Gaskell, M. G.** (2013). Lexical integration of novel words without sleep. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 608–622. DOI: https://doi.org/10.1037/a0029243

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527. DOI: https://doi.org/10.1037/0033-295X.95.4.492

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, *22*, 1–35. DOI: https://doi.org/10.1016/0010-0285(90)90002-L

MacLeod, C. M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. *Memory & Cognition*, *26*, 201–211. DOI: https://doi.org/10.3758/BF03201133

MacLeod, C. M. (2019). Learning simple associations. *Canadian Psychology*, *60*, 3–13. DOI: https://doi.org/10.1037/cap0000170

MacLeod, C. M., & **Dunbar, K.** (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 126–135. DOI: https://doi.org/10.1037/0278-7393.14.1.126

Miller, J. (1987). Priming is not necessary for selective-attention failures: Semantic effects of unattended, unprimed letters. *Perception & Psychophysics*, *41*, 419–434. DOI: https://doi.org/10.3758/BF03203035

Moors, A. (2016). Automaticity: Componential, Causal, and Mechanistic Explanations. *Annual Review of Psychology*, *67*, 263–287. DOI: https://doi.org/10.1146/annurev-psych-122414-033550

Moors, A., Boddez, Y., & **De Houwer, J.** (2017). The power of goal-directed processes in the causation of emotional and other actions. *Emotion Review*, *9*, 310–318. DOI: https://doi.org/10.1177/1754073916669595

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. DOI: https://doi.org/10.1037/0278-7393.19.4.789

Myung, I. J., Kim, C., & **Pitt, M. A.** (2000). Toward an explanation of the power law artifact: Insights from response surface analysis. *Memory & Cognition*, *28*, 832–840. DOI: https://doi.org/10.3758/BF03198418

Newell, A., & **Rosenbloom, P. S.** (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.

Nissen, M. J., & **Bullemer, P.** (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1–32. DOI: https://doi.org/10.1016/0010-0285(87)90002-8

Pavlov, I. P. (1927). *Conditioned reflexes*. London: Oxford University Press.

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. DOI: https://doi.org/10.1016/j.concog.2006.06.010

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

Schmidt, J. R., & **De Houwer, J.** (2012b). Does temporal contiguity moderate contingency learning in a speeded performance task? *Quarterly Journal of Experimental Psychology*, *65*, 408–425. DOI: https://doi.org/10.1080/17470218.2011.632486

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

Schmidt, J. R., & **De Houwer, J.** (2016a). Contingency learning tracks with stimulus-response proportion: No evidence of misprediction costs. *Experimental Psychology*, *63*, 79–88. DOI: https://doi.org/10.1027/1618-3169/a000313

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. DOI: https://doi.org/10.1016/j.lmot.2016.09.002

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. DOI: https://doi.org/10.1016/j.concog.2009.12.016

Schmidt, J. R., De Houwer, J., & **Rothermund, K.** (2016). The Parallel Episodic Processing (PEP) Model 2.0: A single computational model of stimulus-response binding, contingency learning, power curves, and mixing costs. *Cognitive Psychology*, *91*, 82–108. DOI: https://doi.org/10.1016/j.cogpsych.2016.10.004

Schmidt, J. R., Giesen, C., & **Rothermund, K.** (in press). Contingency learning as binding? Testing an exemplar view of the colour-word contingency learning effect. *Quarterly Journal of Experimental Psychology*.

Shanks, D. R. (2010). Learning: From association to cognition. *Annual Review of Psychology*, *61*, 273–301. DOI: https://doi.org/10.1146/annurev.psych.093008.100519

Stevens, A., Schwarz, J., Schwarz, B., Ruf, I., Kolter, T., & **Czekalla, J.** (2002). Implicit and explicit learning in schizophrenics treated with olanzapine and with classic neuroleptics. *Psychopharmacology*, *160*, 299–306. DOI: https://doi.org/10.1007/s00213-001-0974-1

Thorndike, E. L. (1911). *Animal intelligence*. New York: MacMillan.

Urry, K., Burns, N. R., & **Baetu, I.** (2015). Accuracy-based measures provide a better measure of sequence learning than reaction time-based measures. *Frontiers in Psychology*, *6*, Article 1158. DOI: https://doi.org/10.3389/fpsyg.2015.01158

Wood, W., & **Rünger, D.** (2016). Psychology of habit. *Annual Review of Psychology*, *67*, 289–314. DOI: https://doi.org/10.1146/annurev-psych-122414-033417

Peer review comments
The author(s) of this paper chose the Streamlined Review option, and the Editor's decision letter can be downloaded at:
http://doi.org/10.1525/collabra.320.pr

How to cite this article: Schmidt, J. R., De Houwer, J., & Moors, A. (2020). Learning Habits: Does Overtraining Lead to Resistance to New Learning? *Collabra: Psychology*, 6(1): 21. DOI: https://doi.org/10.1525/collabra.320

Senior Editor: Simine Vazire

Editor: Simine Vazire

Submitted: 08 January 2020 **Accepted:** 24 February 2020 **Published:** 07 April 2020