*Original Article*

QJEP

Quarterly Journal of Experimental Psychology
2020, Vol. 73(5) 739–761

DOI: 10.1177/1747021820906397
qjep.sagepub.com
SAGE

# Contingency learning as binding? Testing an exemplar view of the colour-word contingency learning effect

**James R Schmidt[1], Carina G Giesen[2] and Klaus Rothermund[2]**

## Abstract

The learning of contingent regularities between events is fundamental for interacting with our world. We are also heavily influenced by recent experiences, as frequently studied in the stimulus-response binding literature. According to one view ("unitary view"), the learning of regularities across many events and the influence of recent events on current performance can coherently be explained with one high-learning rate memory mechanism. That is, contingency learning effects and binding effects are essentially the same thing, only studied at different timescales. On the other hand, there may be more to a contingency effect than just the summation of the influence of past events (e.g., an additional impact of learned regularities). To test these possibilities, the current report reanalyses a number of datasets from the colour-word contingency learning paradigm. It is shown that the weighted sum of binding effects accumulated across many previous trials (with especially strong influence of very recent events) does explain a large chunk of the contingency effect, but not all of it. In particular, the asymptote towards which the contingency effect decreases by accounting for an increasing number of previous-trial binding effects is robustly above zero. On the other hand, we also observe evidence for higher-order interactions between binding effects at differing lags, suggesting that a mere linear accumulation of binding episodes might underestimate their influence on contingency learning. Accordingly, focusing only on episodic stimulus-response binding effects that are due to the last occurrence of a stimulus rendered contingency learning effects non-significant. Implications for memory models are discussed.

## Keywords

Contingency learning; binding; exemplars; episodic memory; mixed effect models; multicollinearity

Received: 24 June 2019; revised: 13 November 2019; accepted: 15 November 2019

## Introduction

One of the basic requirements of the human cognitive system, if not *the* most basic, is our ability to learn regularities between events in our environment (Allan, 2005; Beckers et al., 2007; Shanks, 2010). Contingency learning is the basic building block for causal learning, knowledge acquisition, and the formation of the expectancies that make our world feel ordered rather than chaotic. Also fundamental, but perhaps no different, is our ability to bind our experiences into memory traces for later retrieval (Hommel, 1998, 2004; Hommel et al., 2001; Logan, 1988). Especially recent experiences can have a particularly potent influence on our behaviour (Grant & Logan, 1993). The present report will put to the test a unitary view of contingency learning and transient stimulus-response (S-R) binding, which argues that learning and binding effects are ultimately the result of the same high-learning rate memory storage and retrieval processes.

Consider two paradigms, each developed for the purpose of studying (seemingly) two different things. The first is the colour-word contingency learning paradigm (Schmidt et al., 2007; for related procedures, see Carlson & Flowers, 1996; J. Miller, 1987; Mordkoff, 1996; Mordkoff & Halterman, 2008). In this paradigm, participants respond to the print colour of a word on each trial as quickly and accurately as

[1]LEAD-CNRS UMR5022, Université Bourgogne Franche-Comté (UBFC), Dijon, France
[2]Institute of Psychology, Friedrich-Schiller-Universität Jena, Jena, Germany

**Corresponding author:**
James R Schmidt, LEAD-CNRS UMR 5022, Université de Bourgogne, 21000 Dijon, France.
Email: james.schmidt@ubfc.fr

**Table 1.** Example contingency manipulation.

| Colour | Word | | |
|---|---|---|---|
| | Find (%) | Walk (%) | Make (%) |
| Blue | 80 | 10 | 10 |
| Red | 10 | 80 | 10 |
| Green | 10 | 10 | 80 |

Example mapping only. Which word is presented most often with which colour is different for each participant.

possible. In meaning, the words are unrelated to the colours (e.g., unlike a Stroop task; Stroop, 1935). However, each word is presented most often in one colour, as illustrated in Table 1. For instance, "find" might be presented most often in blue, "walk" most often in red, and "make" most often in green. These contingencies between the distracting words and colour responses are learned by participants, as indicated by robustly faster and more accurate responses to *high-contingency* trials (e.g., "find" in blue) relative to *low-contingency* trials (e.g., "find" in red). Critically, this paradigm is typically used as a method to study the learning of contingent regularities *across* many events (i.e., learning of which word *tends to* go with which colour response).

Next, consider a different sort of paradigm. Although there are several variants of S-R binding (or feature integration) procedures (Hommel, 1998), distracter-response binding paradigms are of particular interest for the present purposes (Frings et al., 2007; Rothermund et al., 2005). Again, participants respond to a target (e.g., print colour) while ignoring a distracter (e.g., word). Many designs even closely mirror contingency learning designs (i.e., similar stimuli, trial structure) with the notable exception that distracters are *not* correlated with targets/responses (e.g., each distracting word is presented equally often in all colours/with all responses). Instead, researchers assess performance on the second of two trials as a function of whether (a) the distracter repeats (e.g., "find" followed by "find") or changes (e.g., "find" followed by "walk"), and (b) the target colour (and therefore response) repeats (e.g., blue followed by blue) or changes (e.g., blue followed by red). The standard finding is illustrated in Figure 1. In particular, participants are faster to respond when the distracter repeats and the (target) response repeats (*DR-RR*; sometimes termed a *complete repetition*), relative to when the distracter changes and the response repeats (*DC-RR*; or *partial response repetition*). However, participants are (a little) slower to respond when the distracter repeats and the response changes (*DR-RC*; or *partial word repetition*), relative to when the distracter and response both change (*DC-RC*; or *complete alternation*). Globally, response repetitions are faster than response changes, but the Stimulus Relation × Response Relation interaction (Figure 1) is most crucial.

These binding effects are typically interpreted in terms of short-term *event files* (Hommel et al., 2001) that are

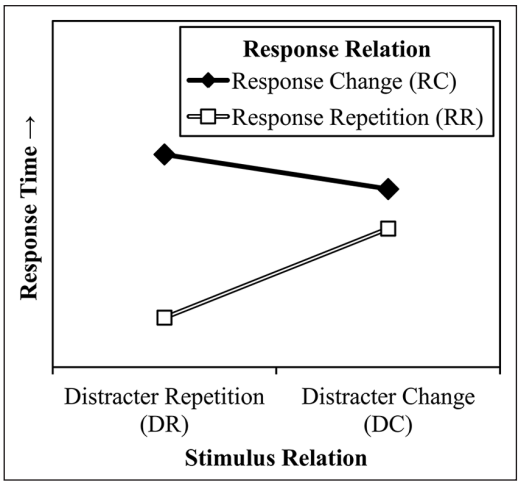


**Figure 1.** Typical results pattern indicating retrieval of distracter-response bindings.

created in short-term/working memory which are often assumed to "disintegrate" shortly after being created (Stoet & Hommel, 1999). The basic idea of stimulus-response binding is that the repetition of a stimulus (e.g., distracting word) leads to retrieval of the recently associated response that was executed on the previous trial. This can lead to facilitation if a response repetition is needed (DR-RR). However, this same similarity-based retrieval can lead to retrieval interference if a new response is required (DR-RC). For instance, after seeing "find" in blue, a second presentation of "find" will lead to a retrieval bias in favour of a blue response. However, if the current stimulus is "find" in red, then the retrieval bias towards a blue response leads the participant astray (i.e., the retrieved and thus expected blue response competes with selection of the appropriate red response).

Interestingly, the binding and contingency learning literatures have not had much of a history in communicating with each other (but see Colzato et al., 2006; Hommel & Colzato, 2009), at least until very recently (e.g., Giesen & Rothermund, 2015; Moeller & Frings, 2017a). Although not uncommon for researchers to assume that binding is related to learning (e.g., Giesen et al., 2012; Horner & Henson, 2011; Logan, 1988; Schnyer et al., 2007), viewing binding either as the first step of long-term learning or simply as single trial learning, this notion is rarely explored directly and there are many other researchers that (implicitly or explicitly) view binding and learning as two dissociable processes (Frings & Rothermund, 2011; Giesen & Rothermund, 2014; Herwig & Waszak, 2012; Moeller & Frings, 2017b).

## *Distinct mechanisms view*

Of the limited number of considerations of the potential relationship between binding and contingency learning,

**Table 2.** The frequency of different stimulus-response (S-R) binding trial types for high- and low-contingency trials with an 80% contingency.

| High-contingency trials | | | Low-contingency trials | | | Type of S-R relation Trial $n-1 \rightarrow$ Trial $n$ |
|---|---|---|---|---|---|---|
| Previous trial ($n-1$) | | Current trial ($n$) | Previous trial ($n-1$) | | Current trial ($n$) | |
| Stimulus | Frequency (%) | Stimulus | Stimulus | Frequency (%) | Stimulus | |
| ⬆ Find | 26.6 | ⬆ Find | Find | 3.3 | Find | DR-RR |
| Walk<br>Make | 6.6 | ⬆ Find | ⬆ Walk<br>Make | 30 | Find | DC-RR |
| Find<br>Find | 6.6 | ⬆ Find | ⬆ Find<br>Find | 30 | Find | DR-RC |
| ⬆ Walk<br>make<br>Walk<br>⬆ Make | 60 | ⬆ Find | Walk<br>Make<br>Walk<br>⬆ Make | 36.6 | Find | DC-RC |

DR-RR: distracter repetition, response repetition; DC-RR: distracter change, response repetition; DR-RC: distracter repetition, response change; DC-RC: distracter change, response change.
Notably, the easier DR-RR and DC-RC trials are more frequent for high-contingency trials, and the harder DC-RR and DR-RC trials are more frequent on low-contingency trials. High-contingency pairings are marked with an arrow to help conceptualise why these binding frequencies are unbalanced. See electronic version of this article for a colour table.

some researchers have suggested that the two are completely different, which can be termed a *distinct mechanisms* view. For instance, Colzato and colleagues (2006; see also, Hommel & Colzato, 2009), who provided the first detailed exploration of the potential relationships between learning and binding, suggested that learning effects are due to slow updating of association weights in long-term memory, whereas binding effects are due to temporary event files created (and swiftly destroyed) in short-term memory. They do not argue that learning and binding are completely unrelated to each other, but do argue that the processes that bring about learning and binding effects are quite distinct. The focus of their work, however, differs in important respects from the main research question we are pursuing here (e.g., they investigate S-S bindings, whereas we are mostly interested in studying the relation between associations and binding with regard to S-R connections). Methodologically, they investigated moderating effects of established associations on binding, whereas we are interested in explaining and decomposing contingency learning effects on the basis of episodic binding and retrieval. For interested readers, we provide a detailed description and discussion of their work in Supplementary Material A.

The primary argument for distinct mechanisms of Colzato and colleagues (2006; see also, Hommel & Colzato, 2009) was that contingency and binding effects did not interact with each other. In other words, the binding interaction was just as large for high contingency as for low-contingency stimuli. One might have anticipated that if learning and binding were due to the same mechanism, then some interaction might be observed between learning and binding. However, this is not necessarily the case. Indeed, as will be described in the "Teasing apart learning and binding" section (see Table 2 and Figure 2), a single mechanism can produce pure additivity. Relatedly, effects due to distinct mechanisms need not be additive, either. Supplementary Material A describes their studies in further detail and expands further on why an additive relationship between learning and binding does not necessarily entail distinct mechanisms. In that vein, the present report will take an alternative approach of measuring the influence of bindings on contingency effects directly.

### *Unitary mechanism view*

Contrary to the distinct mechanisms view discussed above, Schmidt et al. (2016) demonstrated how binding and learning effects could be, in principle, two consequences of the same learning process(es). This will be referred to as the *unitary mechanism* view. It may be important to stress up front what we do and do not mean by a "unitary mechanism." We do not mean to make the strong claim that there is only one process responsible for learning/binding or, even further, that this one process is supported by a single brain area. Rather, the notion is that the effects we study in learning and binding procedures may result from the same acquisition process(es). The conceptual point that learning and binding might be long- and short-term consequences of the same acquisition processes was made with the Parallel Episodic Processing (PEP) model (Schmidt, 2016a, 2016b, 2018), which is similar in concept to other episodic/exemplar/instance accounts of memory (e.g.,

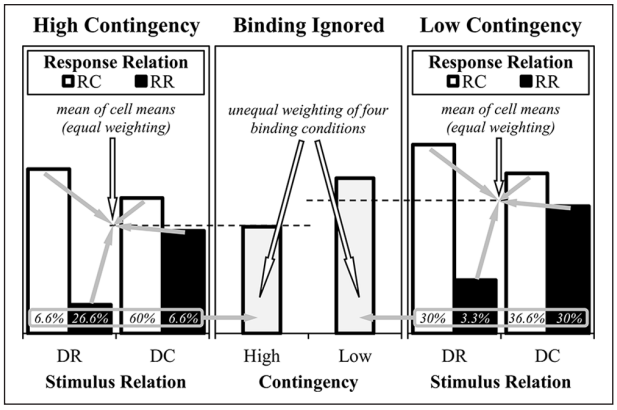


**Figure 2.** Illustration of how bindings can produce a contingency effect. Because the binding conditions are not uncorrelated with contingency, coding for binding effects will produce a reduced remaining contingency effect (difference between dashed lines) relative to when the distinction between binding conditions is ignored (difference between bars in central plot). Note that the dashed lines correspond to the mean of cell means (converging arrows), whereas the bars in the centre plot correspond to the weighted means of the same (imagined) data.

Hintzman, 1984, 1986, 1988; Logan, 1988; Medin & Schaffer, 1978; Nosofsky, 1988a, 1988b; Nosofsky et al., 2011; Nosofsky & Palmeri, 1997). The model also takes heavy inspiration from the work of Hommel on "event files" (e.g., Hommel, 1998; Hommel et al., 2001), which may even be viewed as equivalent to exemplars if the influence of event files are assumed to be more lasting (for a detailed discussion, see Hommel, 1998).

In an exemplar store, each experienced event is encoded as an episodic trace in memory. On presentation of a new stimulus, memory traces are retrieved to the extent that they are similar to the current input. For instance, presentation of "find" in blue will lead to retrieval of memories of trials in which the word "find" was presented, memories of trials in which the colour blue was presented, and especially strong retrieval of trials in which "find" in blue was presented. Such a model can explain contingency learning effects. For instance, as "find" was presented most often in blue, most memories of the stimulus "find" will point to a blue response. This will speed responding if this high-contingency response (blue) does need to be made, but not if another, low-contingency response (e.g., red) is required.

There is also a retrieval-induced decay in the PEP model, such that older and older episodes have less and less influence on memory retrieval. This also means that recently encoded events have a particularly strong influence on retrieval (for a similar notion, see Grant & Logan, 1993; see also, Salasoo et al., 1985; Sloman et al., 1988). Thus, if "find" was *just* presented in red, then re-presentation of "find" will lead to some bias towards a red response, due to retrieval of recently encoded memory traces (which, in the model, are nothing else than short-term "bindings"). On the performance level, retrieval of short-term bindings will produce the following effects: (a) a benefit to response speed if the repeated word was just presented with the same colour response that is currently required (DR-RR), compared to colour repetition trials without word repetition (DC-RR). (b) A small cost if the repeated word was just presented with a different colour response than the one that is currently required (DR-RC), compared to colour change trials without word repetition (DC-RC; see Figure 1). Together, these produce the interaction that we normally describe as the binding effect. Of course, the model will be influenced by numerous traces of the word, allowing for learning of contingent *regularities*, but will also be impacted by recently occurring events. Indeed, the same model with one learning rate was able to simulate both binding and contingency effects (along with a range of other findings) using just one mechanism (Schmidt et al., 2016). Depending on how one looks at it, the notion is either (a) binding effects are a short-term consequence of learning or (b) learning effects are a long-term consequence of many accumulated bindings. "Binding" and "learning" in this conceptualisation are the same thing. Of course, these computational modelling results only demonstrate that one mechanism could *in principle* account for both learning and binding effects, but it remains to be seen whether empirical results support this position.

Whether or not contingency effects can be solely explained by many individual bindings (and whether binding effects can be explained as an exclusive consequence of learning) is an open question. Some experimental evidence suggests the contrary. Of particular interest are a recent series of studies by Moeller and Frings (2017a; see also Herwig & Waszak, 2012). In their procedure, the target was a letter presented at fixation and distracters were other letters presented either to the left and right (horizontal) or above and below (vertical) the target letter. In a contingency learning block, distracting letters were predictive of the target letter. The contingency effect was roughly the same for horizontally and vertically arranged distracters. In a binding block, there was no contingency between distracters and targets. The distracter-response binding effect was robustly larger with horizontally arranged distracters. They interpreted this as indicating that binding of distracters to responses only occurs with the word-like horizontal arrangement of letters, whereas vertically oriented distracters are filtered out. However, they posited that contingency learning processes are able to capitalise on any regularity. These results seem to argue against a unitary, one mechanism view of binding and learning. On the other hand, a unitary mechanism view might still be able to explain these results if it is additionally assumed that contingent attentional capture (see Chun & Jiang, 1998; Cosman & Vecera, 2014; Schmidt, 2014) of the (normally filtered) vertically arranged distracters occurs in the contingency block due to

the informativeness of the distracters. Future research might explore this notion directly.

In a second experiment by Moeller and Frings (2017a), contingency effects were found to be robustly larger with a slower task pacing (2,000 ms between trials), whereas binding effects were found to be larger with a faster task pacing (500 ms between trials). This was interpreted as evidence that binding effects are due to the influence of rapidly decaying event files, which do not survive a long trial interval (Frings, 2011), whereas contingency effects can benefit from advanced preparation time (i.e., the response can be anticipated to a stronger degree if the distracter appears well in advance of the target). Whether this pattern of results can be accommodated within a unitary mechanism account of learning and binding is less certain. Possibly, one might assume that the ever-diminishing impact of older and older episodes is determined not only by how many trials ago an event occurred (Grant & Logan, 1993; Schmidt et al., 2010, 2016) but also the real amount of time elapsed (Wixted & Ebbesen, 1991; or even temporal distinctiveness; see Horoufchin et al., 2011a, 2011b; Howard & Kahana, 2002). One decay function might fit both patterns of results (e.g., because the immediately preceding trial dominates retrieval if it was *just encoded*, but the just-encoded trace is decayed to an increasing degree with more separation in time from the prior trial). Alternatively, recently encoded events may indeed quickly decay by default (explaining decreased binding effects with delay), but distracters may be actively maintained after it has been learned that distracters are informative stimuli (explaining increased contingency effects with extra preparation time). Thus, reinterpretation of these data might be possible, but findings like these do seem to pose some challenges for a unitary mechanism view.

At minimum, however, a relationship between binding and contingency effects is logically implied by any stable model of memory. Whether conceptualised in terms of exemplars, localist associations, distributed representations, a reservoir, or otherwise (which are often mathematically equatable or at least nearly indistinguishable from each other; Barsalou, 1990; Kelly et al., 2017), decay in a memory system is *required* for a stable memory store (e.g., see Gerstner & Kistler, 2002; K. D. Miller & MacKay, 1994; Schmidt et al., 2016) and is built into all successful neural nets in one way or another. For instance, when updating distributed association weights via backpropagation (Rumelhart et al., 1986), it is by definition the case that the influence of the currently experienced event will have a larger impact on connection weights than any single older event: the influence of older events (a weighted sum in the connection weights) is weakened to the extent that the newly encoded event is recorded. That is, changing connection weights *towards* what is just now being encoded effectively means that we must adjust *away* from what was previously encoded in the connection weights (i.e., proportional to the learning/decay rate of the model). What this means is that recently experienced events will have an effect on performance exceeding that of the task-wide contingency. Conceptualised from an exemplar-based view, the influence of each event must follow some form of a decay function, whereby older and older memory traces have smaller and smaller impacts on the current trial. Of course, the slope and exact shape of the function might vary and the strength of the encoding and/or retrieval of particular events might have some variability (e.g., flashbulb memories; Brown & Kulik, 1977), but the typical pattern should be something comparable to a power function (Grant & Logan, 1993; Wixted & Ebbesen, 1991).

## *Teasing apart learning and binding*

Based on the abstract considerations in the previous section, it is plausible to assume that binding and contingency effects are related, at least to some degree. Going beyond this basic postulate, it is additionally possible that the influence of bindings from previous trials provides a *complete* account of contingency learning effects. This much more far-reaching proposal is based on the following considerations, regarding an interdependence between binding and contingency learning effects. Specifically, the different types of repetitions (DR-RR, DR-RC, DC-RR, and DC-RC) do not occur with equal frequency on high- and low-contingency trials. As illustrated in Table 2, DR-RR trials, which are responded to extremely fast, are substantially more likely on high-contingency trials. This is for a simple reason: if "find" is presented most often in blue and infrequently in red, then it is much more likely that the word "find" was presented in blue on the previous trial. As such, this will match the current high-contingency stimulus "find" in blue (DR-RR), but will not match the current low-contingency stimulus "find" in red (DR-RC). It will be an extremely rare event for the same low-contingency stimulus to repeat on successive trials. For the same reason, a DR-RC is much more likely on a low-contingency trial (e.g., "find" in blue followed by "find" in red) than on a high-contingency trial (e.g., "find" in red followed by "find" in blue).

Because DR-RR and DC-RC trials are more likely on high-contingency trials and DR-RC and DC-RR trials are more likely on low-contingency trials, binding effects can explain variance in the contingency effect. This is illustrated in Figure 2. That is, when binding factors are included in the regression, the participant mean high-contingency response time will be the mean of the cell means for (high contingency) DR-RR, DC-RR, DR-RC, and DC-RC (i.e., each contributing 25% to the mean), as illustrated in the left panel of Figure 2. The same is true of low-contingency trials, as illustrated in the right panel of Figure 2. When binding is ignored, however, the participant mean high-contingency response time is effectively a

weighted mean, given that much more of the trials will be DR-RR (26.6%) or DC-RC (60%) and very few will be DC-RR or DR-RC (6.6% each). For low-contingency trials, the relative proportion of DR-RR (3.3%) and DC-RC (36.6%) trials is much lower and the proportion of DC-RR or DR-RC trials is much higher (30% each). As illustrated in the centre panel of Figure 2, this produces mean high- and low-contingency response times (bars) that are different from the binding-controlled means (dashed lines). Most notably, if the binding interaction of the standard form is observed, then the binding-controlled contingency effect is *necessarily* smaller than the same effect without a binding control. It is also important to note that in this example (albeit, with imaginary data) learning and binding are completely additive with each other. That is, the binding interaction is identical for high- and low-contingency trials, only with an added main effect of contingency. This demonstrates how a single mechanism can be consistent with additivity (see the "Distinct mechanisms view" section).

This insight has the following implications: coding for the bindings of the previous trial, the remaining contingency effect *not* accounted for by prior-trial bindings might be reduced. Continuing this analysis "backward" many lags might eliminate the entirety of the contingency effect. However, the contingency effect will not be eliminated in this way if there is an effect of learning independent of individual bindings, as the distinct mechanisms account must predict. That is, if a contingency effect is not just the summation of many previous-trial bindings and there is a unique (e.g., long-term memory association formation) process producing learning effects, then there must remain a unique effect of the contingency manipulation independent of individual bindings. In other words, the regularity produces an effect (due to learning) separate from that of the influence of individual trials on their own (due to binding). Indeed, according to this view, we should anticipate little influence of bindings beyond one or two previous trials, after which the event files are normally assumed to have disintegrated from working memory.

Note that the relation between binding types and contingency holds not only for Trial $n-1$, but for all lags. For instance, for the current high-contingency stimulus "find" in blue, it is much more likely that the identical stimulus pairing ("find" in blue) was presented on Trial $n-1$, $n-2$, $n-3$, and so on, than a partial repetition like "find" in red. As such, the contingency effect can simply be the summation of many such bindings across a large number of trials. Or, more to the point: the (weighted) sum of binding effects across multiple lags simply *is* the contingency effect. A learned regularity is emergent from many discrete event encodings, which is the basic principle of exemplar-based models. This is similar to the suggestion of Logan (1990) that repetition priming and automaticity (i.e., skill acquisition) are the same thing (which is, in turn, the same thing we propose here as an explanation of binding and contingency learning).

To assess this possibility, Schmidt et al. (2010) provided a tentative analysis similar to one which we will also adopt here. In particular, one can encode for colour repetitions, word repetitions, and (more critically) their interaction at each lag for each trial (i.e., the binding interaction illustrated in Figure 1). Then, we can use these predictors in a regression on response times (see also Figure 2). If adding contingency as a factor to the regression explains further variance, then the contingency effect is not exclusively due to binding biases. This analysis did reveal that contingency continued to explain variance, even after accounting for binding effects up to Trial $n-5$. Furthermore, there were no statistically significant binding effects on Trials $n-2$ to $n-5$. The authors therefore concluded that there is more to the contingency effect than just a collection of binding effects at varying lags, which is inconsistent with the unitary mechanism view. In fact, the conclusion was that the contingency effect is largely unrelated to binding, which differs significantly from the conclusions of the present report.

There are considerable problems with the analysis of Schmidt and colleagues (2010), however. Most notably, binding influences were only considered for Trials $n-2$ to $n-5$. Binding biases may carry over from many more trials than this. Perhaps even more important, Trial $n-1$ binding biases were not considered. The particular data used were from a task where target colour repetitions were not possible by design, though word repetitions could occur. With this design, word repetitions (DR-RC) should still be more frequent on low-contingency trials than on high-contingency trials. Both of these problems probably led to an underestimation of how large of a contribution binding effects make to the contingency effect. On the reverse side, analysis of the residual contingency effect (i.e., the contingency effect after controlling for binding biases) may allow the binding factors to "steal" variance from a true effect of contingency. As briefly explained in Supplementary Material B,[1] this is due to multicollinearity between the binding interactions and the contingency factor, which produces "omitted variable bias." By *including* the contingency factor in the regression and assessing the size of the contingency beta the more lags of binding effects that are coded in the regression, this omitted variable bias is eliminated. This article will take a more sophisticated approach to assessing the relationship between binding and contingency effects. Importantly, we note that we were not strongly motivated towards any particular view when beginning this work. We viewed the unitary mechanism notion as interesting in concept, but also saw reasons why it might not prove to be the whole story.

We did, however, view it as likely that binding effects last longer than typically assumed. This was motivated by the exemplar-based unitary view mentioned before: we

should expect prior events to matter most when they were recent, but ever-diminishing effects also from older events. A priori, such biases from older events should be small (tapering off the older the event), but non-zero. Tapering effects like this should be easy to miss if trying to study the effect specific to one lag (e.g., $n-3$), but our modelling approach should do a better job of revealing these tapering effects.

## Method

### Datasets

For the present analyses, we use Experiments 1a ($n=36$, *trials*$=300$) and 1b ($n=34$, *trials*$=300$) of Schmidt and De Houwer (2016), Experiment 1 ($n=62$, *trials*$=300$) and the control condition of Experiment 3 ($n=25$, *trials*$=180$) of Schmidt and De Houwer (2012b), and the single experiment ($n=46$, *trials*$=300$) in Schmidt and De Houwer (2012a). This represents 203 individual participant datasets, with 57,900 observations. Trials were randomised without replacement in blocks of 30 in the first two studies, but randomised with replacement in the remaining three. All experiments were three-choice colour identification tasks with three words as the predictive stimuli, similar to the example given in the "Introduction" section (see Table 1). Immediate repetitions of the target colour were possible in these studies. Further details of the individual studies can be obtained in the original reports. These studies were initially conducted to investigate questions wholly unrelated to binding.

### Analysis procedure

The initial analysis (Analysis 1) is similar in concept to that in Schmidt and colleagues (2010) but improves on the analysis in the following ways. First, the present analysis draws on a much larger collection of datasets for greater statistical power (see the "Datasets" section). Second, the present analysis considers a wider window of lags, from Trial $n$ (i.e., no lags considered) to $n-12$. In the original analysis, only $n-2$ to $n-5$ were considered. Third, and perhaps most importantly, in the new analysis we assess the size of the contingency effect that remains as a function of the number of lags of binding effects considered. That is, we consider the size of the raw contingency effect along with the size of the contingency effect remaining after accounting for the $n-1$ binding effect, the $n-1$ and $n-2$ binding effect, and so on. Thus, we can visualise the function via which the contingency decreases with increasing lags of binding effects considered. With this, we can further model whether the asymptote of such a function is zero or robustly larger than zero. That is, does the contingency effect decrease towards zero as we account for more and more previous trial bindings (zero asymptote), or is there a component of the contingency effect that cannot be explained by the individual events alone (positive asymptote)? The unitary mechanism view predicts the former. That is, if the contingency effect is simply an accumulation of binding effects from prior trials, then there should not be a main effect of contingency independent of the bindings on their own. In contrast, the distinct mechanisms view predicts a robust overall effect of the task-wide contingency, which will only be conflated with short-lived binding effects from one or two prior trials. That is, the contingency effect should reduce after accounting for binding effects from one or two trials, but should not decrease further by accounting for more and more prior trials. We will also eliminate the omitted variable bias described earlier.

Note also that while adding more and more factors to a regression (e.g., as is the case when we consider more and more prior trial binding effects) will, by definition, result in *overall* model fit improving. However, addition of extra factors *does not* necessarily cause the estimated importance of other factors in the model to decrease. In particular, adding a meaningless factor to a regression will increase the overall fit due to maximisation on random error (overfitting), but the other (e.g., meaningful) factors in the model will only be influenced randomly. Thus, the contingency effect will not simply decrease due to addition of extra binding factors unless those extra binding factors actually have explained true variance in the contingency effect. Note also that comparisons between more and less complex models with Akaike information criterion (AIC) and Bayesian information criterion (BIC; two measures of model fit explained in further detail just before the "Results" section) corrects for overfitting. Further analyses will explore the influence of intervening events on older bindings (Analysis 2) and the influence of the *last response* that was linked to the current distracting word (Analysis 3), both of which will be explained in their corresponding analysis sections.

### Data preparation and factor coding

All data analyses were performed in R 3.5.1, using the lmerTest 3.0-1 package for modelling (other packages: car 3.0-2, carData 3.0-1, retimes_0.1-2, data.table_1.11.4, lme4 1.1-17, and Matrix 1.2-14) and we use Type III sums of squares with Satterthwaite-corrected degrees of freedom. The datasets and R scripts are freely available on the Open Science Framework (see Note 1) or via the lead author, along with an Excel sheet that we used for coding the binding factors.[2] First, response times less than 150 ms were removed and the remaining response times were then normalised with an inverse transform (−1,000/RT). Transformation (especially inverse) is a standard practice when using linear mixed effects (LME) model analyses (e.g., Andrews & Lo, 2012; Kinoshita et al., 2011; Kliegl et al., 2009; Masson & Kliegl,

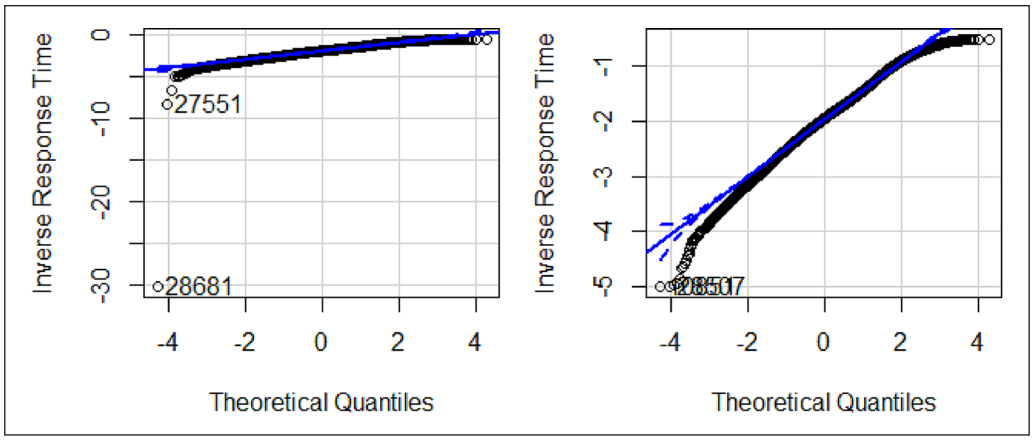


**Figure 3.** Q–Q plots of inverse-transformed current response times before (left) and after (right) a 150 ms trim.

2013; Schmidt & Weissman, 2016) and is required to meet the distributional assumptions of LME (i.e., because of the typical positive skew in response times). Q-Q plots confirmed that the inverse transform and 150 ms trim sufficiently normalised the data, as illustrated in Figure 3. We note that some have illustrated problems with analyses on inverse-transformed data (Balota et al., 2013; Cohen-Shikora et al., 2018; Lo & Andrews, 2015), but these concerns are not applicable to the present case.[3] However, we did also repeat our analyses from Analysis 1 with non-transformed data using generalised linear mixed models (GLMM)[4] and results supported the same conclusions (R scripts for these analyses are also included).

On each trial, contingency (high vs. low) was coded along with colour repetition (response repetition vs. change), word repetition (distracter repetition vs. change), and the interaction (which were coded as separate factors, with DR-RR and DC-RC trials as one level of the factor and DC-RR and DR-RC trials as the other) for each of the 12 lags (i.e., 36 binding factors in total). All trials in which participants failed to respond or produced an error were eliminated from analyses, along with the first 12 trials for each participant (i.e., because these trials could not be coded backward to Lag $n-12$). Models could then be constructed with an increasing number of lags coded as factors. For instance, Model 0 includes only the subject and item (colour) intercepts and contingency, with no binding predictors included. Model 1 includes these same factors plus the factors of colour repetition, word repetition, and the interaction at Lag $n-1$. Each larger model adds on the binding factors for an additional lag (e.g., Model 12 includes the repetition factors for all 12 lags). The dependent variable for analyses on Models 0–12 (and when comparing information criteria between them) was current trial response times.

## *Slope estimation*

Once we have an estimate of the contingency effect for each participant for each of the Models 0–12 (which, of course, requires a random intercept for contingency), we can test for the shape of the function whereby the contingency effect decreases with an increasing number of previous-trial binding effects accounted for. Thus, for slope analyses, the (single) independent variable is lag (0–12) and the dependent measure is the individual participant response time contingency betas (i.e., low-high contingency) for each lag. That is, for each participant, there is one contingency effect estimate per lag, where Lag 0 corresponds to the beta from Model 0, Lag 1 to the beta from Model 1, and so on. We will fit three models on the participant mean contingency effect as a function of lags considered (i.e., 0–12). The first is a simple *linear model*. The linear model is, of course, by definition false,[5] but is used as a reference. The second model, assumed a priori to be the correct model (Grant & Logan, 1993; Wixted & Ebbesen, 1991), is a *power model*, which fits the $RT = k + a \cdot (lag + 1)^b$ power curve (i.e., in which the contingency effect decreases rapidly after accounting for the first few lags, then continues to decrease at ever-diminishing rates). In this formula, $k$ is the asymptote of interest (i.e., the point at which the contingency effect reduces towards the more lags that are considered). A $k$ of zero would therefore indicate that binding effects fully account for the contingency effect. That is, coding for an infinite number of previous trial binding effects would fully eliminate a unique contribution of the contingency factor. This finding would support the unitary mechanism view. In contrast, a positive asymptote would suggest that there is a component of the contingency effect that is not attributable to merely the summation of individual-trial binding effects. This finding would be inconsistent with the unitary mechanism view. The third model is an *exponential model*, which fits the exponential decay function $RT = k + n \cdot e^{l \cdot lag}$, where $k$ is again the asymptote. The exponential decay function is slightly different than the power function and was included as it might plausibly provide a better fit (e.g., Heathcote et al., 2000; Myung et al., 2000), though did not. Any cross-model comparisons (for the slope or otherwise) are, by necessity, computed with

maximum likelihood estimation (MLE). Restricted maximum likelihood estimation (REML) is used otherwise.

To readers less familiar with LME or model fitting, what these analyses do is test the influence of binding and contingency learning orthogonally, in the same way that two factors in an analysis of variance (ANOVA) are measured separately. The model fitting is equivalent to treating a factor (in this case, lag) as a scale variable with an imposed form, as you would do in an ANOVA with a continuous predictor variable (e.g., training block). The linear model is like a typical regression or correlation (or continuous variable in ANOVA), assuming that each *x* increase in lag translates to a *y* change in response time for the binding factors, whereas the power and exponential functions are (two slightly different) non-linear curves assuming more potent influences of recent events that slowly taper off with increasingly older events. The AIC/BIC scores assess which model provides the best fit after controlling for degrees of freedom (i.e., to counteract "overfitting"), with lower scores indicating better fit. Importantly, binding and contingency effects are assessed orthogonally, in the same way that one might, for instance, assess the influence of a confound in an effect by manipulating the key effect and the confound separately and inserting them as two factors in an ANOVA. Thus, if there are separate mechanisms producing contingency and binding effects (which, of course, combine in the same behaviour), then measuring them separately will reveal the unique influence of each.

## Results

### *Analysis 1: asymptote*

*Full model.* First, consider Model 12, which was conducted on individual trial response times with random subject and contingency intercepts, the main effect of contingency, and all binding factors across the 12 lags. The results of this model are presented in Table 3. Colour/response repetitions produced robustly faster responses (negative estimates) for many lags, as did the word repetitions. Most interesting, however, is the interaction between response and word repetitions. The estimates of these interactions are positive for most lags, indicating the standard binding effect (i.e., faster responses to DR-RR and DC-RC trials relative to DR-RC and DC-RR trials), though only significant in the first two lags. Table 4 presents the AIC and BIC differences between sequential models (e.g., Model 0–Model 1, Model 1–Model 2; raw model scores can be obtained with the R scripts). Substantial extra explanatory variance is present for the first three models, indicating a benefit for coding of these extra lags. For the remaining lags, this seems inconsistent, but Model 12 does have a better AIC fit than Model 3 (difference: 10.95). This result suggests that binding effects at ever-increasing lags do matter, but at an ever-decreasing rate (i.e., the a priori prediction of the unitary account). Together, these results therefore tell a notably different story than Schmidt and colleagues (2010): clear binding influences were observed on the contingency effect. The BIC results should probably be interpreted cautiously.[6] Note also from Table 3 that the main effect of contingency is still significant in Model 12.

*Slope.* Next, we compute the contingency effect for each of the models from Model 0–12. In particular, we compute the individual participant contingency betas for each of the models. These data are presented in Figure 4 (note that response times have been inverse transformed, so are not on a usual response time scale). Visually, we can see that the contingency effect beta drops rapidly early on after accounting for the first few lags and continues to decrease at an ever-diminishing rate afterwards. Next, we test the shape of this function to see whether the contingency effect approaches an asymptote significantly above zero. As expected, the slope fits a power function (also presented in Figure 4 as a solid line) substantially better than a linear (AIC difference: 4,840.9; BIC difference: 4,835.1) or an exponential function (both: 486.1). Importantly, the fitted power function produced a *k* (asymptote) that was significantly positive (Estimate: 0.0580548), $t(202) = 14.45$, $SE = 0.0040165$, $p < .001$, $\eta^2 = .51$. What these results suggest is that around 54% of the contingency effect is accounted for by binding effects, with the remaining 46% not. As such, these analyses are consistent with the idea that individual trial binding effects have a large impact on the magnitude of the contingency effect, but are not consistent with the idea that the contingency effect is made up of a collection of individual trial binding effects exclusively (i.e., as the unitary mechanism predicts). The distinct mechanisms view can be viewed as only partially consistent with the present results: there was a contingency effect independent of binding effects, on one hand, but previous-trial binding effects did (a) account for a substantial portion of the contingency effect and (b) continued to do so far beyond the first one or two trials. Thus, these results might be interpreted as indicating partially distinct mechanisms (but with potential caveats to follow).

### *Analysis 2: higher-order binding interactions*

The unitary mechanism view might be interpreted in two ways. According to one view, the contingency effect reflects the simple summation of individual trial binding effects. That is, the contingency effect emerges from the addition of binding effects from Lag 1 plus the binding effects at Lag 2, and so on. This is what we tested in Analysis 1 and was clearly not supported by the data. On the other hand, the influence of a memory trace on current behaviour could additionally be influenced by intervening

**Table 3.** Model 12 parameter estimates and statistics.

| Factor | Estimate | *SE* | *df* | *t* | *p* |
|---|---|---|---|---|---|
| Intercept | −1.87600 | 0.04854 | 6.099 | −38.656 | <.001*** |
| Contingency | 0.05588 | 0.01355 | 27.64 | 4.123 | <.001*** |
| Response 1 | −0.42440 | 0.00480 | 52,120 | −88.491 | <.001*** |
| Response 2 | −0.03458 | 0.00479 | 52,120 | −7.219 | <.001*** |
| Response 3 | 0.04209 | 0.00479 | 52,110 | 8.792 | <.001*** |
| Response 4 | 0.00633 | 0.00479 | 52,110 | 1.323 | .186 |
| Response 5 | 0.01437 | 0.00479 | 52,120 | 2.999 | .003** |
| Response 6 | −0.00342 | 0.00479 | 52,130 | −0.713 | .476 |
| Response 7 | −0.01330 | 0.00480 | 52,130 | −2.769 | .006** |
| Response 8 | −0.00877 | 0.00479 | 52,120 | −1.829 | .067† |
| Response 9 | −0.01182 | 0.00480 | 52,130 | −2.464 | .014* |
| Response 10 | −0.01477 | 0.00480 | 52,130 | −3.080 | .002** |
| Response 11 | −0.01234 | 0.00477 | 52,120 | −2.590 | .010** |
| Response 12 | −0.00934 | 0.00477 | 52,120 | −1.958 | .050† |
| Word 1 | −0.05400 | 0.00480 | 52,120 | −11.262 | <.001*** |
| Word 2 | −0.01679 | 0.00479 | 52,120 | −3.508 | <.001*** |
| Word 3 | −0.00165 | 0.00479 | 52,120 | −0.344 | .731 |
| Word 4 | 0.00297 | 0.00479 | 52,110 | 0.620 | .535 |
| Word 5 | −0.01023 | 0.00479 | 52,120 | −2.135 | .033* |
| Word 6 | 0.00406 | 0.00479 | 52,130 | 0.848 | .396 |
| Word 7 | 0.00739 | 0.00481 | 52,130 | 1.538 | .124 |
| Word 8 | −0.00095 | 0.00479 | 52,110 | −0.198 | .843 |
| Word 9 | 0.00943 | 0.00480 | 52,130 | 1.967 | .049* |
| Word 10 | 0.00312 | 0.00480 | 52,120 | 0.650 | .516 |
| Word 11 | 0.00178 | 0.00476 | 52,120 | 0.373 | .709 |
| Word 12 | −0.00678 | 0.00477 | 52,120 | −1.422 | .155 |
| Interaction 1 | 0.10110 | 0.00539 | 52,110 | 18.752 | <.001*** |
| Interaction 2 | 0.01438 | 0.00538 | 52,110 | 2.671 | .008** |
| Interaction 3 | 0.00508 | 0.00538 | 52,100 | 0.945 | .344 |
| Interaction 4 | −0.00035 | 0.00537 | 52,090 | −0.065 | .948 |
| Interaction 5 | 0.00308 | 0.00538 | 52,100 | 0.573 | .567 |
| Interaction 6 | −0.00213 | 0.00537 | 52,120 | −0.398 | .691 |
| Interaction 7 | 0.00626 | 0.00537 | 52,110 | 1.165 | .244 |
| Interaction 8 | 0.00632 | 0.00537 | 52,100 | 1.177 | .239 |
| Interaction 9 | 0.00264 | 0.00536 | 52,100 | 0.491 | .623 |
| Interaction 10 | 0.00394 | 0.00535 | 52,080 | 0.736 | .462 |
| Interaction 11 | 0.00060 | 0.00534 | 52,090 | 0.112 | .911 |
| Interaction 12 | −0.00069 | 0.00534 | 52,090 | −0.130 | .897 |

*SE*: standard error.
†$p < .1$, *$p < .05$, **$p < .01$, ***$p < .001$.

events. For example, a Trial $n-2$ DR-RR trial (complete repetition) may not have the same magnitude of a benefit to current trial performance if the same word, or colour, or both was also presented on Trial $n-1$. Specifically, in this case the presence of the distracter word in the more recent Trial $n-1$ might prevent retrieval of the more distant episode that was stored in Trial $n-2$ in which the distracter word was also present, which will effectively block this episode from retrieval (or, at least, reduce its impact). Furthermore, the episode that was created during Trial $n-2$ will already have been retrieved during Trial $n-1$, which will reduce its accessibility on later occasions (e.g., Schmidt et al., 2016).

Like the binding effects for individual lags, such "interactive" binding effects are also confounded with contingency. While distracter or response repetitions are equiprobable following high- and low-contingency pairings, their combinations are not. For example, "find" in blue (high contingency) is much more likely to be followed by another "find" in blue than "walk" in blue (low contingency) is to be followed by another "walk" in blue, not only on trial *n*, but also in the intervening events. That is, there is a confound in the "sequences" of binding effects. Ignoring randomisation constraints (only in some datasets), following a low-contingency stimulus there is a 9/30 chance (per lag) that only the word repeats, a 9/30

**Table 4.** Fit differences between sequential models.

| Model comparison | AIC | BIC |
|---|---|---|
| Model 0–Model 1 | 10,307.64 | 10,281.05 |
| Model 1–Model 2 | 104.05 | 77.46 |
| Model 2–Model 3 | 99.75 | 73.14 |
| Model 3–Model 4 | −1.94 | −28.53 |
| Model 4–Model 5 | 5.69 | −20.91 |
| Model 5–Model 6 | −5.13 | −31.73 |
| Model 6–Model 7 | 1.73 | −24.87 |
| Model 7–Model 8 | −1.16 | −27.76 |
| Model 8–Model 9 | 0.76 | −25.83 |
| Model 9–Model 10 | 3.83 | −22.78 |
| Model 10–Model 11 | 1.57 | −25.02 |
| Model 11–Model 12 | 5.6 | −21 |

AIC: Akaike information criterion; BIC: Bayesian information criterion.

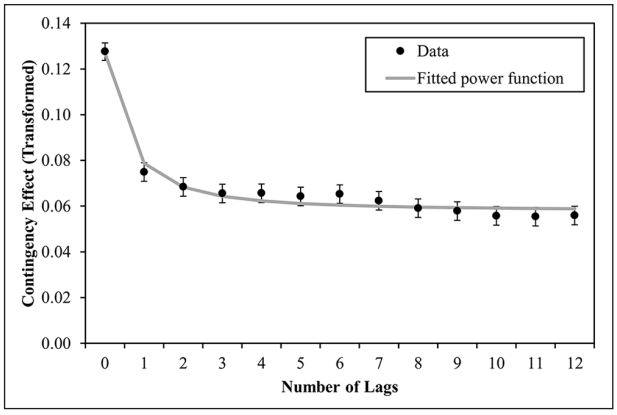


**Figure 4.** Contingency effect estimate (with standard errors) in Models 0–12 with fitted power function. Note that the estimate does not approach a zero asymptote.

chance that only the response repeats, and a small 1/30 chance of a complete repetition. By contrast, for high-contingency trials these are 2/30, 2/30, and 8/30, respectively. Thus, there is globally a higher chance that *something* repeats after a low-contingency trial and a very low chance of a complete repetition. Although rather complicated, such "interactive" binding effects should (i.e., if these interactions matter) produce more speeding of the high-contingency trials and costs for the low-contingency trials. Thus, as with the "simple" (one lag) binding effects, failing to account for "interactive" binding effects could again exaggerate the true main effect of contingency.

In sum, it is plausible to expect interactions between binding effects on different lags and of a specific a priori form: diminished binding effects for older events when following events repeated the features. It is likely impossible to test this notion in the same way as in the prior analyses. In particular, the number of hypothetical interactions between each of the lags grows exponentially at such a rapid rate the more lags considered that the number of factors would quickly exceed the number of data points. This does make it impossible to compute an intercept. However, what we can clearly do is test whether there are "higher-order" binding interactions in the data at all and, more critically, whether they influence the estimate of the contingency effect. That is, even if there are higher-order binding interactions, they may or may not account for added variance in the contingency effect (e.g., if our reasoning above is wrong). If there is no evidence for any higher-order binding interactions (or if such interactions exist but do not impact the contingency effect), then this would spell the end for any version of the unitary mechanism view that we can conceive: if binding effects are merely additive across lags (to be tested here) and additive binding effects do not account for the entire contingency effect (established in Analysis 1), then there is necessarily more to the contingency effect than just binding.

In particular, we consider to what extent the binding interaction on Trials $n-2$ to $n-4$ (short form: int2 to int4) are influenced by whether the word from Trial $n$ repeated on each of the intervening events, and also whether the colour response repeated. Thus, this includes factors for the binding interaction on Trial $n-2$ as a function of word repetition on Trial $n-1$ (short form: int2 × word1), the interaction on Trial $n-3$ as a function of word repetition on Trials $n-1$ (int3 × word1) and $n-2$ (int3 × word2), and similarly for Trial $n-4$ (following the same pattern: int4 × word1, int4 × word2, int4 × word3), in addition to the same interactions with response repetition (int2 × resp1, int3 × resp1, int3 × resp2, int4 × resp1, int4 × resp2, int4 × resp3). This is, of course, only a small subset of the possible interactions one could code for, but a theoretically interesting set. In particular, we might expect that the influence of a binding interaction (say, on Trial $n-2$) will be reduced if the word repeated in an intervening event (e.g., on Trial $n-1$). All analyses and data treatments were identical to those in Analysis 1.

Table 5 presents the results of Model 4 with these added interactions. Notably, some of these interactions are significant. This is especially the case for the interaction between the $n-2$ binding interaction and the $n-1$ response repetition effects (i.e., the most recent potential binding interaction). The positive parameter indicates that the size of the binding interaction is *decreased* when the colour response repeats on the intervening trial. There is, of course, some degree of noisiness in many of the other parameters, which we return to in the next analysis. The contingency betas were further extracted for each participant using the normal Model 4 (as in Analysis 1) and the present version containing the added higher-order interactions. The difference in the contingency effect was robustly smaller when the higher-order interactions were *included* (estimate: 0.06026) than when they were *excluded* (estimate: 0.06562; difference: 0.00536), $t(202)=24.08$, $p<.001$. Both AIC (difference: 114.13) and BIC (7.73) were also lower in the larger

**Table 5.** Model 4 with higher-order binding interactions.

| Factor | Estimate | *SE* | *df* | *t* | *p* |
|---|---|---|---|---|---|
| Intercept | −1.90600 | 0.04819 | 5.725 | −39.544 | <.001*** |
| Contingency | 0.06026 | 0.01114 | 13 | 5.408 | <.001*** |
| Response 1 | −0.45360 | 0.00607 | 52,090 | −74.771 | <.001*** |
| Response 2 | −0.04739 | 0.00579 | 52,140 | −8.189 | <.001*** |
| Response 3 | 0.04598 | 0.00535 | 52,130 | 8.601 | <.001*** |
| Response 4 | 0.00634 | 0.00478 | 52,120 | 1.326 | .185 |
| Word 1 | −0.06689 | 0.00609 | 52,140 | −10.989 | <.001*** |
| Word 2 | −0.02005 | 0.00578 | 52,140 | −3.467 | <.001*** |
| Word 3 | −0.00407 | 0.00537 | 52,140 | −0.758 | .448 |
| Word 4 | 0.00246 | 0.00479 | 52,130 | 0.515 | .607 |
| Interaction 1 | 0.10850 | 0.00552 | 52,110 | 19.674 | <.001*** |
| Interaction 2 | 0.00536 | 0.00587 | 52,100 | 0.914 | .360 |
| Interaction 3 | −0.00199 | 0.00624 | 52,080 | −0.32 | .749 |
| Interaction 4 | −0.01029 | 0.00656 | 52,100 | −1.57 | .116 |
| int2 × resp1 | 0.04963 | 0.00547 | 52,140 | 9.076 | <.001*** |
| int3 × resp1 | 0.01179 | 0.00548 | 52,120 | 2.151 | .031* |
| int3 × resp2 | 0.00720 | 0.00539 | 52,140 | 1.336 | .182 |
| int4 × resp1 | 0.00522 | 0.00546 | 52,140 | 0.956 | .339 |
| int4 × resp2 | 0.02252 | 0.00541 | 52,140 | 4.164 | <.001*** |
| int4 × resp3 | −0.01141 | 0.00528 | 52,140 | −2.161 | .031* |
| int2 × word1 | −0.01299 | 0.00550 | 52,120 | −2.36 | .018* |
| int3 × word1 | −0.00219 | 0.00550 | 52,140 | −0.398 | .691 |
| int3 × word2 | 0.00268 | 0.00539 | 52,130 | 0.498 | .619 |
| int4 × word1 | 0.01166 | 0.00549 | 52,140 | 2.126 | .033* |
| int4 × word2 | −0.00873 | 0.00541 | 52,140 | −1.613 | .107 |
| int4 × word3 | 0.00949 | 0.00531 | 52,140 | 1.787 | .074† |

*SE*: standard error.
†$p < .1$, *$p < .05$, **$p < .01$, ***$p < .001$.

model. Combined, these results demonstrate that (a) there are interactive effects between bindings at different lags, and more critically (b) these do explain added variance in the contingency effect. This analysis had the potential to conclusively falsify the unitary mechanism view. However, although we cannot continue this sort of analysis backward for enough lags to compute a slope and intercept for the contingency effect, the observed interactive binding effects instead leave open the possibility that the contingency effect might be solely explainable by binding, which (if true) would be consistent with the unitary mechanism view.

## *Analysis 3: previous response*

In the previous analysis, it was shown that the magnitude of binding effects for a given Trial $n-x$ is influenced by intervening events. In particular, binding effects for a given lag were reduced if an intervening event repeated the same stimuli/responses. These analyses were not perfect, largely due to the fact that we cannot code for all potential "interactive binding effects." Indeed, some of the noisiness in parameters could easily be explained by omitted effects we could not code for. For instance, an interaction between a binding event on Trial $n-4$ and Trial $n-1$ ignores what happened on Trials $n-3$ and $n-2$. Still, Analysis 2 demonstrated clearly that the simple "additive" binding model in Analysis 1 missed legitimate binding influences. Following up on this finding and related to the above-mentioned limitation with Analysis 2, we assess the magnitude of the contingency effect as a function of whether the *last occurrence* of the currently presented word was linked to the same or a different colour response in Analysis 3. For instance, if the current trial is "find" in blue, then the *response relation* factor codes whether "find" was last presented with a blue response versus another colour response. Of course, the last presentation of "find" could have been on the immediately previous trial or a number of trials before. We further code this lag in the *lag* factor, which was first coded as the number of trials prior to the current one where the current word last appeared (e.g., "1" if the same word occurred on trial $n-1$, "2" if the same word last occurred on trial $n-2$). As suggested by an anonymous reviewer, this predictor was then inverse transformed (−10/lag) such that recent matches (e.g., 1) were more heavily weighted than much more distant matches (e.g., 12), then centred on the mean.

The core advantage of this analysis[7] is that it provides us with estimates of the real strength of binding effects that are not underestimated due to the inclusion of sequences containing intervening occurrences of the word stimulus. Including sequences with intermediate occurrences into the analysis inevitably reduces the estimated strength of the binding effect by averaging across sequences in which binding effects differ with regard to their strength. Focusing only on binding and retrieval effects in sequences without these intermediate occurrences might therefore give us a closer estimate of the true strength of pure stimulus-based response retrieval when this effect is strong and not mitigated by counteracting influences. Inclusion of the lag factor is important not only as a test of the a priori notion that response relation effects should decrease as lag increases, but also controls for the fact that response relation repeats tend to occur more recently for high-contingency trials than for low-contingency trials.

Similarly, it is also worth considering the relation in the reverse direction. That is, whether the last occurrence of the colour/response was with the same or different word is also important. Between the last presentation of the word and the current trial, it is also possible that there was an intermediate response repetition (i.e., with a different word). This will occur equally often with high- and low-contingency trials (e.g., on $.\overline{154}$ of Trial $n-1$ trials), which therefore does not represent a problem. Slightly more complicated is when splitting contingency by response relation condition (i.e., when including the contingency · response relation interaction in the model). In this case, intermediate response repetitions are not equally distributed across the four contingency by response relation cells (high-repeat: $.\overline{09}$; high-change: $.4\overline{09}$; low-repeat: $.4\overline{09}$; low-change: $.1\overline{26}$).

Other than the change in factors, all analyses and data treatments were performed in the identical manner as in Analyses 1 and 2 with one exception. Trials on which there was not a previous presentation of the word (i.e., the first appearance of each distracting word) or of the colour (i.e., the first appearance of each colour) were excluded (e.g., rather than the first 12 trials of the experiment), as response relation factors cannot be coded for these trials.

To begin with, we consider only the response relation and lag factors, with their interaction, in Model A to establish a baseline response relation effect and how this is modulated by lag. The results of this and the following two models are presented in Table 6. As can be observed, there was a significant response relation effect in Model A, with responses overall faster if the last response to the distracter was the same as the current response. There was also a main effect of lag, indicating overall slower responses with increasing lag. More importantly, response relation and lag interacted. The negative estimate indicates that, as predicted, the response relation effect decreases with increasing lag.

Next, we consider Model B, which is identical to Model A except that the main effect of contingency was also added. The same effects observed in Model A were also present in Model B. Notably, the newly added contingency factor did not produce a significant effect. Furthermore, AIC was numerically comparable (difference: −0.3) and BIC was lower (−9.1) in the smaller Model A, demonstrating worse fit for Model B. In other words, contingency did not robustly explain anything on top of the response relation effects. Comparing individual participant contingency betas from Model B to a model containing only contingency as a factor, the contingency effect was reduced by 92% (from 0.12545 to 0.01049), $t(202) = 156.91$, $p < .001$.

Finally, we consider Model C, in which the full factorial interactions between contingency, response relation, and lag were added. In this model, contingency did come out as significant, albeit significantly negative. Contingency further interacted with response relation and lag. Exploring these interactions further, we observed that the contingency effect was significantly negative with a response relation repeat (−0.07726), $t(35{,}210) = 4.610$, $SE = 0.01676$, $p < .001$, and significantly positive with a response relation change (0.02487), $t(18{,}120) = 3.657$, $SE = 0.00680$, $p < .001$. These effects should, however, be interpreted with caution given the intermediate response repetition confound already discussed. In particular, it is more likely that there was an intermediate response repetition (which speed responses) on low relative to high-contingency trials when there is a response relation repeat (explaining the unusual negative contingency effect) and the reverse when there is a response relation change (explaining the positive contingency effect). We therefore suspect that these interactive patterns are spurious.[8] Together these results suggest even more strongly still that binding influences have a sizable impact on the magnitude of the contingency effect.

As one final consideration, it might be suggested that binding contrasts are not completely separated from the task-wide contingency in another respect. Current-trial contingency was controlled for in our prior analyses, indicating no direct influence of the contingency *on the current trial*. However, it might be proposed that the contingency influences the strength of binding, resulting in stronger binding effects for high-frequency pairings. Such a "prior-trial" contingency effect would actually be inappropriately attributed to binding. For a current high-contingency trial, a response relation repetition ("complete repetition") necessarily implies that the last presentation of the word was also a high-contingency trial, whereas for a response relation change the last presentation of the word was necessarily a low-contingency trial. This is actually (partially) reversed for low-contingency trials: the last presentation of the word for a response relation repetition was necessarily a low-contingency trial, whereas for response relation change it could have been either high or

**Table 6.** Models A–C.

| Factor | Estimate | *SE* | *df* | *t* | *p* |
|---|---|---|---|---|---|
| Model A | | | | | |
| Intercept | −2.00500 | 0.04020 | 6.726 | −49.880 | <.001*** |
| Response relation (Rel) | 0.16070 | 0.00422 | 53,520 | 38.130 | <.001*** |
| Lag | 0.05635 | 0.00072 | 53,520 | 77.970 | <.001*** |
| Rel × Lag | −0.07646 | 0.00125 | 53,520 | −61.290 | <.001*** |
| Model B | | | | | |
| Intercept | −2.01400 | 0.04074 | 7.075 | −49.440 | <.001*** |
| Contingency (Cont) | 0.00847 | 0.00633 | 53,520 | 1.340 | .180 |
| Rel | 0.15650 | 0.00527 | 53,520 | 29.670 | <.001*** |
| Lag | 0.05635 | 0.00072 | 53,520 | 77.970 | <.001*** |
| Rel × Lag | −0.07646 | 0.00125 | 53,520 | −61.290 | <.001*** |
| Model C | | | | | |
| Intercept | −1.92400 | 0.04341 | 9.442 | −44.335 | <.001*** |
| Cont | −0.07908 | 0.01670 | 53,530 | −4.735 | <.001*** |
| Rel | 0.04360 | 0.02043 | 53,520 | 2.134 | .033* |
| Lag | 0.06807 | 0.00505 | 53,520 | 13.493 | <.001*** |
| Cont × Rel | 0.10290 | 0.01804 | 53,520 | 5.704 | <.001*** |
| Cont × Lag | −0.01143 | 0.00488 | 53,520 | −2.341 | .019* |
| Rel × Lag | −0.07567 | 0.00601 | 53,520 | −12.602 | <.001*** |
| Cont × Rel × Lag | 0.00322 | 0.00529 | 53,520 | 0.608 | .543 |

*SE*: standard error.
†$p < .1$. *$p < .05$. **$p < .01$. ***$p < .001$.

low contingency (i.e., the high-contingency response to the word or the remaining low-contingency response).

For current (i.e., Trial *n*) high-contingency trials, we cannot test this assumption, because response relation (repeat vs. change) is perfectly confounded with the contingency during the previous occurrence (high vs. low contingency). However, for current low-contingency trials, the "confound" is not complete. Instead, we can distinguish between three different types of "last occurrence" trials: (a) response relation repetitions where the last occurrence was low contingency (repeatLC), (b) response relation changes where the last occurrence was low contingency (changeLC), and (c) response relation changes where the last occurrence was high contingency (changeHC). Based on simple binding, of course, we should expect that repeatLC trials should be faster than changeLC and changeHC trials. No differences should be expected between the two types of response relation changes, however. On the other hand, if the contingency does influence binding, as discussed above, then we should expect more interference for the changeHC trials. To test this notion, we compared each pair of conditions in a trial type (e.g., changeLC vs. changeHC) by lag LME. As shown in Table 7, we did observe faster responses to repeatLC trials relative to both changeLC and changeHC. This is exactly as predicted by a simple binding account. However, we did not observe a difference between changeLC and changeHC trials, inconsistent with the "contingency influences binding" notion.

## Discussion

The binding and contingency learning literatures are both large enterprises. Excluding a recent enthusiasm for consideration of the potential relation between learning and binding (Giesen & Rothermund, 2015; Moeller & Frings, 2017a; Schmidt et al., 2016), these two literatures have historically been non-communicative with each other (i.e., with only sporadic exceptions). The goal of the present investigation was to further explore the notion that contingency and binding effects might be the result of one unitary memory mechanism. This notion contrasts with previous suggestions (e.g., Colzato et al., 2006; Hommel & Colzato, 2009) that the two types of effects are due to entirely different processes (e.g., unbinding costs vs. association adjustment) operating on different types of memory codes (e.g., event files vs. episodic or associative memory traces) in different memory stores (e.g., short- vs. long-term memory).

The present results suggest two things. First, there was a clear influence of recent bindings on the magnitude of the contingency effect. That is, at least to some degree the contingency effect is a (compound) binding effect in disguise. Although the most notable effects were for the two immediately preceding trials in Analysis 1, larger models coding up to 12 previous trials continued to explain more variance in the contingency effect. This diverges from the conclusions of Schmidt and colleagues (2010). The influence of bindings at longer lags will not always be apparent if attempting to assess the unique binding effect for a single

**Table 7.** Analysis of previous contingency types.

| Factor | Estimate | *SE* | *df* | *t* | *p* |
|---|---|---|---|---|---|
| repeatLC vs. changeLC | | | | | |
| Intercept | −1.95300 | 0.04332 | 5.748 | −45.075 | <.001*** |
| Trial Type | −0.11820 | 0.00958 | 9,107 | −12.349 | <.001*** |
| Lag | 0.01085 | 0.00276 | 9,097 | 3.937 | <.001*** |
| Trial Type × Lag | 0.03511 | 0.00276 | 9,099 | 12.727 | <.001*** |
| repeatLC vs. changeHC | | | | | |
| Intercept | −1.94500 | 0.04611 | 5.502 | −42.181 | <.001*** |
| Trial type | −0.12030 | 0.01315 | 1,702 | −9.149 | <.001*** |
| Lag | 0.01286 | 0.00363 | 1,807 | 3.546 | <.001*** |
| Trial Type × Lag | 0.03443 | 0.00362 | 1,807 | 9.501 | <.001*** |
| changeLC vs. changeHC | | | | | |
| Intercept | −1.82900 | 0.04308 | 5.593 | −42.441 | <.001*** |
| Trial Type | −0.00540 | 0.00795 | 9,410 | −0.679 | .497 |
| Lag | −0.02386 | 0.00241 | 9,417 | −9.909 | <.001*** |
| Trial Type × Lag | −0.00040 | 0.00241 | 9,415 | −0.166 | .868 |

*SE*: standard error.
†$p < .1$, *$p < .05$, **$p < .01$, ***$p < .001$.

lag, because the unique contribution at a given lag grows vanishingly smaller the further away it occurred from the current trial. However, if one pools across many such trials, it becomes clear that variance is being explained past one or two trials. This therefore also suggests that while bindings are certainly more potent influencers of behaviour when occurring very recently (e.g., Frings, 2011), "event files" do not completely disintegrate after a few seconds (Stoet & Hommel, 1999). Future research might aim to assess whether similar residual effects of bindings can be observed in more typical, non-learning binding procedures (e.g., with regression approaches similar to the current report). If the present story is correct, then one should expect that coding backwards for more and more prior bindings should lead to more and more explained variance in response times, likely following a comparable power curve as we observed in the present studies.

Relatedly, it is worth pointing out that bindings are conflated with the contingency effect as long as the contingency manipulation is still in place (e.g., as in the datasets used in this article). However, this is no longer the case once the contingency is removed, for instance, in a test block where the same words are presented equally often in all colours. In this "unlearning" or test scenario, the distribution of the four binding conditions is the same for high- and low-contingency trials (i.e., aside from carryover from the prior phase). It would therefore be interesting to carry out analyses similar to the present ones across learning and unlearning phases (and perhaps also relearning and/or counterconditioning). Consistent with the story of this work, adaptations to changes in the contingencies occur very quickly (e.g., Lin & MacLeod, 2018; Schmidt & De Houwer, 2016; Schmidt et al., 2010), but it would be interesting to assess whether one learning function could explain changes in performance across phases.

Second, the results clearly show that the unitary mechanism account of learning and binding is, at minimum, mostly correct, though may or may not be the whole story. That is to say, learning and binding effects clearly are not orthogonal to one another. Our Analysis 1 demonstrated that a substantial portion of the contingency effect is accounted for by binding. However, there is a unique effect of contingency that exceeds that attributable to the *additive* effects of individual exemplars. As such, a learning-as-binding account that assumes no interactivity between bindings from multiple lags (e.g., that what happens on Trial $n-1$ does not influence binding effects from Trial $n-2$) is clearly wrong (or rather: not the complete story). However, if we do assume such interactive effects, the pure learning-as-binding view might still be salvageable. Indeed, we found robust evidence for interactive binding effects in our Analysis 2, where binding effects for a given lag were diminished if an intervening event repeated the distracting stimulus or target/response. We further investigated the binding influence of the *last occurrence* of a stimulus in Analysis 3. These analyses paint a somewhat different picture than Analysis 1, with binding now accounting for nearly all (or at least the bulk) of the contingency learning effect, rendering the unique contribution of the contingency learning factor quite small and non-significant. The combined results suggest that the unitary mechanism view is correct in the assumption that what we measure in learning and binding paradigms is, at minimum, mostly assessing the same acquisition processes. That is, the results suggest that there may be a partial separation between the mechanisms that contribute to contingency and binding effects, but there is also a considerable degree of overlap.

Similarly, the distinct mechanisms view either is consistent or inconsistent with the present results depending

on how extremely one interprets it. The notion that the binding effect and the contingency effect (perhaps: after controlling for one or two prior bindings) tap into two completely different mechanisms in a process-pure way is clearly inconsistent with the present data. Instead, most of what is being studied with a contingency effect measure is a combination of prior-trial bindings. Phrased the reverse direction, much of what is studied in a binding effect is a short-term consequence of learning. However, this does not rule out some degree of unique process contributions to each effect. In other words, the truth might lie in less extreme percentages: the processes that bring about contingency and binding effects overlap heavily, but not completely. As an aside, the present report also demonstrated why we need to be careful about how we explore unique variance explained by correlated predictors, as discussed briefly in Supplementary Material B.

### *The unitary view*

Overall, the results of our Analysis 1 were not consistent with a pure learning-as-binding model, at least as conceptualised in terms of additive binding effects. That is, binding effects from multiple lags did not "add up" to explain the entirety of the contingency effect. The contingency factor continued to explain variance when competing in the multiple regression with the binding factors. By one view, this might seem surprising. Effectively, a contingency is defined by the relative history of co-occurrences of pairs of stimulus/response features. Whether conceptualised in terms of discrete memory traces or association weights, the current state of the system is primarily determined by the updates that occur on each experienced event. For this reason, an explanation of the contingency effect in terms of the influences of individual events might seem to be almost inherently true (at least partially, it should be). Thus, the results of our Analysis 1 might seem puzzling from this perspective.

However, a purely additive interpretation of the unitary, contingency-as-binding view might not be correct. Indeed, we think that it is reasonable to assume that binding effects at different lags *should* interact. Although we did begin with the notion that additive binding contrasts might possibly be able to explain all variance in the contingency effect, there are good reasons to suppose that this might not be true. In particular, the influence of a given event on current-trial performance presumably should be influenced by intervening events. Indeed, Analysis 2 revealed that the binding interaction for a given lag *is* moderated by repetitions of the target response (or distracter) in the intermediate events before the current trial. For instance, an $n-2$ DR-RR of "find" in blue does not have the same magnitude of influence on current-trial performance if the colour response (or distracter) also repeated on Trial $n-1$. This makes sense, as the memory updating that occurs on Trial $n-1$ should presumably influence the memory trace from Trial $n-2$. For instance, the exemplar encoded on Trial $n-2$ should be retrieved if stimuli repeat on Trial $n-1$ (but not if stimuli do not repeat) and this should influence subsequent retrievability of the $n-2$ exemplar. In the PEP model, for instance, retrieval-induced decay should make the exemplar less retrievable.

The same interactive processes might prove true of associative, distributed, or other models of memory (which may be mathematically equatable with exemplar-based accounts; Kelly et al., 2017). That is, the constant updating of association weights should presumably produce a similar sort of decay function, whereby the association weight is simply a recency-weighted average of previous event encodings. However, *which* weights are adjusted on a given trial will be influenced by the stimuli presented and response made on a given trial. In particular, adjustments of association weights on Trial $n-1$ should influence associations for similar stimuli (activated associations) more than for non-activated stimuli. As such, if Trial $n-1$ repeats a stimulus from Trial $n-2$, the influence of the Trial $n-2$ event on association weights will be weakened.

Most critically, however, we further found that these "interactive binding effects" explained further variance in the contingency effect in Analysis 2. Thus, there are binding influences that were not coded in the initial regression analysis (i.e., Analysis 1). An analysis that does not take into account these complex interdependencies thus leads to an underestimation of binding effects by averaging across different types of sequences for which binding effects differ systematically: the weaker binding effects of previous episodes during sequences that contain an intervening occurrence of either the current stimulus or the response dampens the predictive power of those trials in which no such intervening repetition occurred. These interactions, as previously discussed, are systematically in favour of high-contingency trials and to the detriment of low-contingency trials. As such, to the extent that the contingency factor is able to capitalise on the unmodelled variance for these higher-order binding interactions, the "true" effect of contingency will be overestimated (and the true effect of binding underestimated).

Simply coding for the "full gamut" of the potential interactive patterns in a slope analysis akin to those presented in Analysis 1 will prove impossible, as the number of interactive contrasts will grow at an exponential rate the more lags one considers (i.e., such that the number of factors will rapidly exceed the number of data points). To overcome these complexities, we introduced a new way of testing the effects of binding that focuses just on the last occurrence of the distracting stimulus for the current trial in Analysis 3, contrasting trials in which the last episode contains either the response that matches or does not match the current response requirements. Such an analysis is not fraught with problems relating to

complex interactions of competing and interacting retrieval processes, and it yielded a fairly clear picture: apparently, binding effects regarding the last occurrence of a word stimulus explained away the majority of the contingency effect (92%). The contingency effect that remained was non-significant, albeit trending in the correct numerical direction. The extra influence of older, weaker exemplars (i.e., not-last occurrences of the stimulus) might explain the remaining variance. Thus, Analysis 3 had its limitations, too. Unfortunately, combining all the desirable features of each of the three analyses into one large analysis does not seem possible with the present datasets. Future experimental research might aim to develop procedures optimised to capture specific interactive binding effects (e.g., by strategically manipulating the series of items), rather than testing for such effects after the fact as we did in the present report.

### *Implications for memory models*

The exemplar-based, unitary account explained above provides a relatively coherent and parsimonious account of a range of phenomena. Binding (or encoding) of events into memory and the influence of the subsequent retrieval of these events on performance suggests strong ties between areas of research seemingly investigating unrelated phenomena. Related conceptual points have been made several times in a number of focal literatures. For instance, simple binding "biases" (or "confounds") have repeatedly been noted in a variety of domains, such as in the congruency sequence effect (CSE; Hommel et al., 2004; Mayr et al., 2003), proportion congruent effect (PCE; Risko et al., 2008), switch cost (Goschke, 2000; Schmidt & Liefooghe, 2016), and negative priming (Rothermund et al., 2005). It has similarly been noted that improvements with practice (skill acquisition) and repetition priming can been explained by the same power function of memory: (a) an ever-increasing store of exemplars allows for faster and faster responses with practice, and (b) the decay (or decreasing retrievability of memory traces) with time explains repetition priming effects (Logan, 1990). Similarly, we suggest that (a) decreasing influences of older and older traces and (b) accumulation of more memory traces of frequent over infrequent events explains binding and learning effects, respectively. A simplified illustration in Figure 5 illustrates this point.

### *Learning/binding as confounds*

As mentioned earlier, binding and learning effects, in some instances, are not the key effects of interest, but rather confounds to something else that authors are aiming to study. For instance, the CSE is the observation that congruency effects (e.g., in the Stroop task) are smaller following an incongruent trial than following a congruent trial (Gratton et al., 1992). The CSE is typically used as a means to study conflict-driven attentional adjustments (Botvinick et al., 1999), though binding biases systematically confound this effect (Hommel et al., 2004; Mayr et al., 2003) without proper controls (e.g., because complete repetitions are only possible on a sequence of two congruent or two incongruent trials, but not on trials where congruency changes from congruent to incongruent or vice versa). Similarly, the PCE is the observation that congruency effects are smaller when most trials are incongruent relative to when most trials are congruent (Logan & Zbrodoff, 1979). Like the CSE, the PCE is typically used to study conflict-driven attentional adaptation (Botvinick et al., 2001), but this effect is confounded with contingencies (Schmidt & Besner, 2008) between distracters and targets (without proper controls). As suggested by Schmidt (2019) and reinforced by the present analyses, these "two" biases might actually be regarded as one. Approaches to dealing with such confounds already exist (Braem et al., 2019), but the present results might also suggest that we should not only be attentive to overall regularities between stimuli, but also the recent pairings when aiming to rule out confounds. For instance, if a PCE is confounded by a contingency, then we should expect especially large biases for recent stimulus pairings and not just the overall contingency, which has received some (Risko et al., 2008) but limited attention.

### *Distinct mechanisms possibilities*

In the previous sections, we considered the unitary view of learning and binding. However, it is worth considering some of the reasons why there might be more to the contingency effect than just the individual events in isolation from one another. For example, the cognitive system might not only encode events, but also perform an averaging of many events to detect typical patterns, which may then be consolidated to a long-term store (or just extra traces in the same store). Relatedly, a participant might happen to notice (consciously or unconsciously) that there are *regularities* between events (e.g., "find" *tends to* be presented in blue), producing (explicit or implicit) knowledge of these regularities. Thus, the participant might encode additional traces of the sort *the word "find" tends to be presented in blue*. Alternatively, noticed regularities might remain primed in working memory. Relatedly, participants may only notice part of a regularity (Perruchet & Amorim, 1992; Perruchet et al., 1997). For instance, a consciously noticed distracter and response repetition (DR-RR) of a stimulus (though not necessarily the task-wide contingency) might be actively maintained in short-term memory for some time (e.g., "I just saw 'find' in blue twice in a row"). All of these factors might produce an impact of the contingency on behaviour that exceeds that of the retrieval of individual traces. Indeed, one possibility is that the

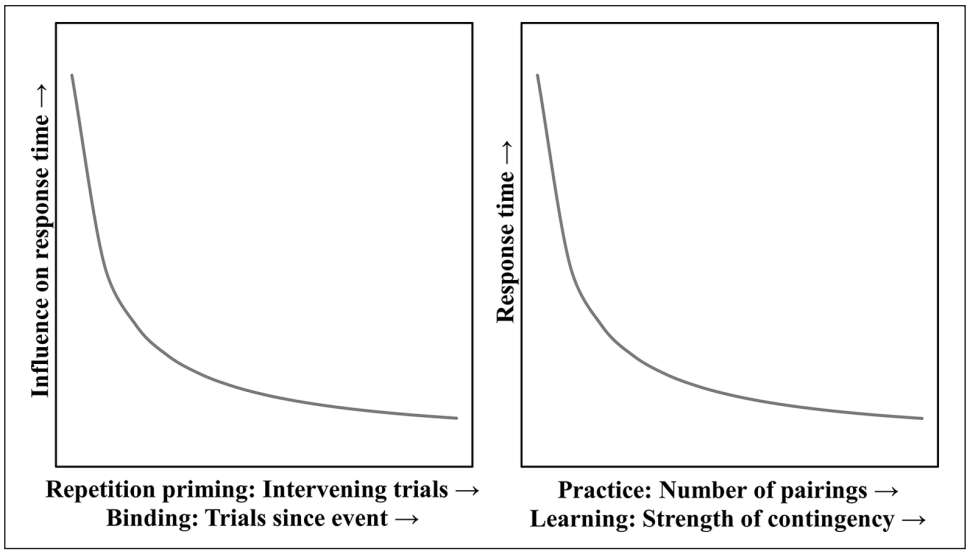


**Figure 5.** The influence of events on current behaviour as a function of how long ago the event was encoded (left) and the speed of responding after summing across multiple similar events that have been encoded (right). Both should presumably follow some function comparable to this power function.

contingency effect is exclusively due to individual-trial effects in participants that are unaware of the contingencies, but those who do become aware produce an effect of the regularity on top of individual-trial biases. Future work might therefore explore whether the "true" effect of the regularity (i.e., beyond individual bindings) is awareness dependent.

Another consideration when relating learning and binding is the timescale over which we are discussing learning. Most "learning effects" that we study in the lab are, for instance, the result of just-acquired regularities experienced during the course of an experiment, what we might call early learning. If we consider learning consolidated over considerably longer periods of time (e.g., the meaning of words in a language, associations between objects and their typical colouration), then the relation between learning and binding may become less direct. For example, Colzato and colleagues (2006; see also Hommel & Colzato, 2009) observed some differences in the magnitude and persistence of binding effects for heavily related relative to poorly related features. For instance, fruits (e.g., strawberry) are highly associated with colour (e.g., red) whereas shapes (e.g., triangle) are not. This might make it quite difficult to ignore the colour in which a fruit is presented, but less so for a simple triangle. Indeed, binding effects vanished rather quickly for the latter stimuli, but not for the former. This learning-derived attentional influence (which also occurs on shorter timescales in the case of contingent attentional capture; Cosman & Vecera, 2014; Jiang & Chun, 2001) might be regarded as an effect of learning that is distinct from binding (related also to the discussion of Moeller & Frings, 2017a, in the "Introduction" section).

### *Binding-as-learning*

This work focused primarily on binding as a possible account of contingency learning effects, but the reverse is also implied by the unitary account tested here. That is, binding effects could, at least in part, be a consequence of learning. Here, too, there could in principle be more to a binding effect than simple stronger retrieval of recently encoded events from long-term memory. For instance, binding effects could, in part, be additionally influenced by experiences kept active in working/short-term memory from the previous trial (event files), as typically proposed in the binding literature. Further research exploring potential unique influences on binding not attributable to learning might be interesting in this respect (e.g., Moeller & Frings, 2017a). Furthermore, to the extent that contingency and binding effects *are* due to different mechanisms, further investigation of the possible interrelations between learning and binding might be explored (e.g., Giesen & Rothermund, 2015).

In discussing this and related research with colleagues, one sentiment we have heard is that if binding effects are "just" a short-term consequence of learning, then this makes binding less interesting. We disagree. In other domains in which learning or binding "biases" represent confounds in attempting to study something else (e.g., higher-order control), the type of perspective presented here does potentially undermine the goal of a research domain. We do not think that the same is true for the relation between learning and binding. Instead, learning and binding can be viewed as two sides of the same coin and studying the influence of recent experiences on behaviour is no less interesting if closely related to the influence of

frequent experiences on behaviour. If anything, a unitary mechanism view is exciting and might indicate that there is much to gain by further cross-pollinating across the two research domains.

### Limitations

One limitation of the present research is that we focused exclusively on one particular learning task, the colour-word contingency learning procedure. This was, in part, due to availability of data and in part due to the structural similarity between the colour-word contingency learning procedure and distracter-response binding preparations. However, future research should aim to perform similar analyses on other types of learning procedures, such as the flanker contingency paradigm (Carlson & Flowers, 1996; J. Miller, 1987), shape-colour learning paradigm (Levin & Tzelgov, 2016; Schmidt & De Houwer, 2019), or, with modification, sequence learning (Nissen & Bullemer, 1987). Such work could clarify whether the observations of the present report are general to incidental learning more globally or whether different types of learning environments produce, for instance, stronger influences of the global regularity above that attributed to individual bindings.

Indeed, it could be the case that certain paradigms produce larger effects of regularities than for the colour-word contingency learning paradigm. For example, in the studies of Moeller and Frings (2017a) described earlier, distracting stimuli (flanking letters) and target stimuli (target letters) were both spatially and temporally separated. Especially the temporal separation between the pre-exposed flankers and the following target might be relevant. The extra time may allow participants to use knowledge of the regularities (maybe even explicitly) to anticipate the upcoming target. If so, we might imagine different results with the current analyses applied to such a temporal flanker paradigm.

Another limitation of this work is that we did not consider the distinction between stimulus-stimulus and stimulus-response learning and binding. In both colour-word contingency and distracter-response binding procedures, participants respond to the target on the basis of the target identity, typically with fixed target-to-response mappings (but not always; for example, Giesen & Rothermund, 2014). As such, a repetition of the target entails a repetition of the response, and a target alternation entails a response change as well. Both learning and binding effects might, however, be influenced by both stimulus-stimulus regularities and by stimulus-response regularities. For instance, participants could be learning that "find" is presented most often in blue (stimulus-stimulus pairings) and/or that "find" is presented most often with the J-Key response (stimulus-response pairings). In the present report, this does not reflect a confound per se, as the four conditions of the binding interaction are identical both in terms of stimulus-stimulus and stimulus-response bindings. However, separate analyses of stimulus-stimulus and stimulus-response regularities would be informative.

We might also wonder about whether the contribution of the regularity (vs. the individual-trial effects) changes over the course of extended training. For instance, it could be that earlier on in learning performance is primarily determined by the influences of individual bindings, but after more extensive practice the regularity begins to have a unique additional effect. The experiments reanalysed in this article were relatively short (longest: 300 trials). An interesting extension of this work would be to perform the same types of analyses on different phases of learning in more lengthy training procedures.

## Conclusion

The present report aimed to test a unitary view of contingency and binding effects. Coding for binding effects at multiple lags did not seem to fully account for the contingency effect (Analysis 1). That is, the present results show that there is some component of the contingency effect that is not accounted for by the additive influence of individual memory traces alone. On the other hand, this is only true if it is assumed that each event has a separate, additive influence on current trial performance, which is probably unreasonable. In particular, it is reasonable to expect that the influence of a given event on current-trial performance should be influenced by intervening events. Indeed, we found additional evidence for interactions between binding effects at different lags, with binding effects being diminished if an intervening event repeated stimuli/responses from the initial binding (Analysis 2). Fully accounting for *all* potential interactive binding effects is impossible, but we were able to show that these "interactive" binding effects explain yet more variance in the contingency effect. This implies that even more of the contingency effect is accounted for by binding than what Analysis 1 suggests. Our results were therefore also not sufficient to rule out a pure learning-as-binding view entirely, in contrast to a prior (but problematic) attempt to assess the same question (Schmidt et al., 2010). Relatedly, focusing on binding influences on the basis of the last occurrence of the current distracting stimulus explained away a substantial portion of the contingency effect (Analysis 3). At minimum, a substantial portion of the contingency effect does seem to be due to individual bindings, which indicates that the effects studied in learning and binding procedures are strongly related. We hope that future experimental and modelling work will help to further tease apart the unitary and distinct mechanism views of learning and binding effects.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the French "Investissements d'Avenir" programme, project ISITE-BFC (contract ANR15-IDEX-0003) to J.R.S.

## ORCID iD

James R Schmidt https://orcid.org/0000-0002-0412-396X

## Open practices

The data and materials from the present experiment are publicly available at the Open Science Framework website: https://osf.io/h2vep/

## Supplementary material

The supplementary material is available at: qjep.sagepub.com

## Notes

1. Data and R scripts are publicly available on the Open Science Framework: https://osf.io/h2vep/. More information is available from the lead author on request.
2. Because the original studies were not concerned with binding, the datasets did not include variables for word and colour repetitions.
3. Data transformations can change not only the magnitude but direction of an interaction of two main effects (i.e., whether it is additive, overadditive, or underadditive), with the typical transforms (e.g., inverse) biasing towards underadditivity. This is primarily only a concern for assessing additivity and is not applicable to the crossover type interactions (see Loftus, 1978, for general concerns about interpreting non-crossover interactions) that we investigate in the present report (or simple main effects). Globally, whether transformed or non-transformed data should be used depends on the situation (e.g., non-transformed data can also be inappropriate in many scenarios).
4. These analyses were conducted with a Gamma distribution and identity link function, as previously recommended by Lo and Andrews (2015). These generalised linear mixed models (GLMMs) did have trouble converging, however.
5. A linear model with any non-zero slope, positive or negative, implies asymptotic performance of, respectively, $\infty$ or $-\infty$.
6. The Bayesian information criterion (BIC) results are decidedly different (difference between Models 3 and 12: −228.43), which suggests that the more complex model is less likely. This discrepancy is due to the much larger penalty for added factors in BIC ($\ln(n)\cdot k$) than Akaike information criterion (AIC) ($2k$). Given that the binding interaction is the only factor per lag that we are interested in (word and response repetitions being orthogonal with contingency) and that we assume a priori that binding effects at longer lags exist but are vanishingly small, the BIC penalty is unreasonable (e.g., because it heavily "punishes" larger models for adding largely uninfluential factors, like the main effects of word repetition, and weaker effects at longer lags; related to the issue of tapering effects discussed by Burnham & Anderson, 2002).
7. Note that we initially performed some analyses with a simple analysis of variance (ANOVA) that supported the same general conclusions that we will report here, but two anonymous reviewers suggested that sticking with a similar analysis approach as Analyses 1 and 2 was preferable. Code for the original ANOVA tests are also included in the R scripts.
8. An anonymous reviewer (the same who suggested the addition of our lag factor) suggested that we also include a "word relation" factor (i.e., coding for whether the last presentation of the colour/response was with the same vs. different word) and a corresponding response-linked lag factor along with contingency, response relation, and lag in one large LME. Although an interesting idea, this proves impossible due to nesting of some of the factors (e.g., it is impossible to have a response relation repeat or change preceding a word relation repeat). This leads to rank deficiency. Incidentally, contingency does not come out as significant in such an analysis (see R scripts), though it is difficult to interpret such a model.

## References

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

Andrews, S., & Lo, S. (2012). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 152–163. https://doi.org/10.1037/a0024953

Balota, D. A., Aschenbrenner, A. J., & Yap, M. J. (2013). Additive effects of word frequency and stimulus quality: The influence of trial history and data transformations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1563–1571. https://doi.org/10.1037/a0032186

Barsalou, L. W. (1990). On the indistinguishability of exemplar memory and abstraction in category representation. In T. K. Srull & R. S. Wyer Jr. (Eds.), *Advances in social cognition: Content and process specificity in the effects of prior experiences* (Vol. 3, pp. 61–88). Lawrence Erlbaum.

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

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.

Botvinick, M. M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, *402*, 179–181.

Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J. C., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring adaptive control in conflict tasks. *Trends in Cognitive Sciences*, *23*(9), 769–783.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, *5*, 73–99.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). Springer.

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

Chun, M. M., & Jiang, Y. H. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71. https://doi.org/10.1006/cogp.1998.0681

Cohen-Shikora, E. R., Suh, J., & Bugg, J. M. (2018). Assessing the temporal learning account of the list-wide proportion congruence effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(9), 1703–1723. https://doi.org/10.1037/xlm0000670

Colzato, L. S., Raffone, A., & Hommel, B. (2006). What do we learn from binding features? Evidence for multilevel feature integration. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 705–716. https://doi.org/10.1037/0096-1523.32.3.705

Cosman, J. D., & Vecera, S. P. (2014). Establishment of an attentional set via statistical learning. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1–6. https://doi.org/10.1037/a0034489

Frings, C. (2011). On the decay of distractor-response episodes. *Experimental Psychology*, *58*, 125–131. https://doi.org/10.1027/1618-3169/a000077

Frings, C., & Rothermund, K. (2011). To be or not to be . . . included in an event file: Integration and retrieval of distractors in stimulus-response episodes is influenced by perceptual grouping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1209–1227.

Frings, C., Rothermund, K., & Wentura, D. (2007). Distractor repetitions retrieve previous responses to targets. *Quarterly Journal of Experimental Psychology*, *60*, 1367–1377.

Gerstner, W., & Kistler, W. M. (2002). Mathematical formulations of Hebbian learning. *Biological Cybernetics*, *87*, 404–415. https://doi.org/10.1007/s00422-002-0353-y

Giesen, C., Frings, C., & Rothermund, K. (2012). Differences in the strength of distractor inhibition do not affect distractor-response bindings. *Memory & Cognition*, *40*, 373–387. https://doi.org/10.3758/s13421-011-0157-1

Giesen, C., & Rothermund, K. (2014). Distractor repetitions retrieve previous responses and previous targets: Experimental dissociations of distractor-response and distractor-target bindings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 645–659. https://doi.org/10.1037/a0035278

Giesen, C., & Rothermund, K. (2015). Adapting to stimulus-response contingencies without noticing them. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 1475–1481. https://doi.org/10.1037/xhp0000122

Goschke, T. (2000). Intentional reconfiguration and involuntary persistence in task set switching. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance* (Vol. 18, pp. 331–355). MIT Press.

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. https://doi.org/10.3758/BF03197193

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*, 480–506.

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.

Herwig, A., & Waszak, F. (2012). Action-effect bindings and ideomotor learning in intention- and stimulus-based actions. *Frontiers in Psychology*, *3*, Article 444. https://doi.org/10.3389/fpsyg.2012.00444

Hintzman, D. L. (1984). Minerva 2: A simulation model of human memory. *Behavior Research Methods Instruments & Computers*, *16*, 96–101.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411–428.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528–551.

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition*, *5*, 183–216.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, *8*, 494–500.

Hommel, B., & Colzato, L. S. (2009). When an object is more than a binding of its features: Evidence for two mechanisms of visual feature integration. *Visual Cognition*, *17*, 120–140. https://doi.org/10.1080/13506280802349787

Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849–878. https://doi.org/10.1017/S0140525x01000103

Hommel, B., Proctor, R. W., & Vu, K. P. L. (2004). A feature-integration account of sequential effects in the Simon task. *Psychological Research*, *68*, 1–17.

Horner, A. J., & Henson, R. N. (2011). Stimulus–response bindings code both abstract and specific representations of stimuli: Evidence from a classification priming design that reverses multiple levels of response representation. *Memory & Cognition*, *39*, 1457–1471. https://doi.org/10.3758/s13421-011-0118-8

Horoufchin, H., Philipp, A. M., & Koch, I. (2011a). The dissipating task-repetition benefit in cued task switching: Task-set decay or temporal distinctiveness? *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 455–472. https://doi.org/10.1037/a0020557

Horoufchin, H., Philipp, A. M., & Koch, I. (2011b). Temporal distinctiveness and repetition benefits in task switching: Disentangling stimulus-related and response-related contributions. *Quarterly Journal of Experimental Psychology*, *64*, 434–446. https://doi.org/10.1080/17470218.2010.496857

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*, 269–299. https://doi.org/10.1006/jmps.2001.1388

Jiang, Y. H., & Chun, M. M. (2001). Selective attention modulates implicit learning. *Quarterly Journal of Experimental Psychology*, *54*, 1105–1124. https://doi.org/10.1080/02724980042000516

Kelly, M. A., Mewhort, D. J. K., & West, R. L. (2017). The memory tesseract: Mathematical equivalence between composite and separate storage memory models. *Journal of Mathematical Psychology*, *77*, 142–155. https://doi.org/10.1016/j.jmp.2016.10.006

Kinoshita, S., Mozer, M. C., & Forster, K. I. (2011). Dynamic adaptation to history of trial difficulty explains the effect of congruency proportion on masked priming. *Journal of Experimental Psychology: General*, *140*, 622–636.

Kliegl, R., Masson, M. E. J., & Richter, E. M. (2009). A linear mixed model analysis of masked repetition priming. *Visual Cognition*, *18*, 655–681.

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

Lin, O.-Y., & 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.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, Article 1171. https://doi.org/10.3389/fpsyg.2015.01171

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition*, *6*, 312–319.

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

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, *22*, 1–35.

Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition*, *7*, 166–174. https://doi.org/10.3758/BF03197535

Masson, M. E. J., & Kliegl, R. (2013). Modulation of additive and interactive effects in lexical decision by trial history. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 898–914. https://doi.org/10.1037/a0029180

Mayr, U., Awh, E., & Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature Neuroscience*, *6*, 450–452.

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

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

Miller, K. D., & MacKay, J. D. C. (1994). The role of constraints in Hebbian learning. *Neural Computation*, *6*, 100–126.

Moeller, B., & Frings, C. (2017a). Dissociation of binding and learning processes. *Attention, Perception, & Psychophysics*, *79*, 2590–2605.

Moeller, B., & Frings, C. (2017b). Overlearned responses hinder S-R binding. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 1–5.

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

Mordkoff, J. T., & Halterman, R. (2008). Feature integration without visual attention: Evidence from the correlated flankers task. *Psychonomic Bulletin & Review*, *15*, 385–389. https://doi.org/10.3758/Pbr.15.2.385

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.

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

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

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

Nosofsky, R. M., Little, D. R., Donkin, C., & Fific, M. (2011). Short-term memory scanning viewed as exemplar-based categorization. *Psychological Review*, *118*, 280–315. https://doi.org/10.1037/a0022494

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266–300. https://doi.org/10.1037/0033-295X.104.2.266

Perruchet, P., & Amorim, M. A. (1992). Conscious knowledge and changes in performance in sequence learning: Evidence against dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 785–800. https://doi.org/10.1037/0278-7393.18.4.785

Perruchet, P., Bigand, E., & Benoit-Gouin, F. (1997). The emergence of explicit knowledge during the early phase of learning in sequential reaction time tasks. *Psychological Research*, *60*, 4–13. https://doi.org/10.1007/BF00419676

Risko, E. F., Blais, C., Stolz, J. A., & Besner, D. (2008). Nonstrategic contributions to putatively strategic effects in selective attention tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1044–1052.

Rothermund, K., Wentura, D., & De Houwer, J. (2005). Retrieval of incidental stimulus–response associations as a source of negative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 482–495.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533–536.

Salasoo, A., Shiffrin, R. M., & Feustel, T. C. (1985). Building permanent memory codes: Codification and repetition effects in word identification. *Journal of Experimental Psychology: General*, *114*, 50–77.

Schmidt, J. R. (2014). Contingencies and attentional capture: The importance of matching stimulus informativeness in the item-specific proportion congruent task. *Frontiers in Psychology, 5*, Article 540.

Schmidt, J. R. (2016a). Context-specific proportion congruent effects: An episodic learning account and computational model. *Frontiers in Psychology*, *7*, Article 1806.

Schmidt, J. R. (2016b). Proportion congruency and practice: A contingency learning account of asymmetric list shifting effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 1496–1505.

Schmidt, J. R. (2018). Best not to bet on the horserace: A comment on Forrin and MacLeod (2017) and a relevant stimulus-response compatibility view of colour-word contingency learning asymmetries. *Memory & Cognition*, *46*, 326–335.

Schmidt, J. R. (2019). Evidence against conflict monitoring and adaptation: An updated review. *Psychonomic Bulletin & Review*, *26*, 753–771.

Schmidt, J. R., & Besner, D. (2008). The Stroop effect: Why proportion congruent has nothing to do with congruency and everything to do with contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 514–523.

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

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

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

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

Schmidt, J. R., & De Houwer, J. (2019). Cue competition and incidental learning: No blocking or overshadowing in the colour-word contingency learning procedure without instructions to learn. *Collabra: Psychology*, *5*, Article 15.

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

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

Schmidt, J. R., & Liefooghe, B. (2016). Feature integration and task switching: Diminished switch costs after controlling for stimulus, response, and cue repetitions. *PLOS ONE*, *11*, Article e0151188.

Schmidt, J. R., & Weissman, D. H. (2016). Congruency sequence effects and previous response times: Conflict adaptation or temporal learning? *Psychological Research*, *80*, 590–607.

Schnyer, D. M., Dobbins, I. G., Nicholls, L., Davis, S., Verfaellie, M., & Schacter, D. L. (2007). Item to decision mapping in rapid response learning. *Memory & Cognition*, *35*, 1472–1482.

Shanks, D. R. (2010). Learning: From association to cognition. *Annual Review of Psychology*, *61*, 273–301.

Sloman, S. A., Hayman, C. A. G., Ohta, N., Law, J., & Tulving, E. (1988). Forgetting in primed fragment completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 223–239.

Stoet, G., & Hommel, B. (1999). Action planning and the temporal binding of response codes. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1625–1640.

Stroop, J. R. (1935). Studies on interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–661.

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science*, *2*, 409–415. https://doi.org/10.1111/j.1467-9280.1991.tb00175.x