

Adaptation to Conflict Frequency Without Contingency and Temporal Learning: Evidence From the Picture–Word Interference Task

Giacomo Spinelli, Jason R. Perry, and Stephen J. Lupker
University of Western Ontario

In interference tasks (e.g., Stroop, 1935), the difference between congruent and incongruent latencies (i.e., the “congruency” effect) is larger in trial blocks containing mostly congruent trials than in trial blocks containing mostly incongruent trials (the *proportion-congruent* [PC] effect). Although the PC effect has typically been interpreted as reflecting adjustments in attention toward/away from the task-irrelevant dimension (i.e., a conflict-adaptation strategy), recent research has suggested alternative accounts based on the learning of either contingencies (i.e., distractor-response associations) or of temporal expectancies (i.e., the typical response speed on previous trials), accounts in which conflict adaptation plays no role. Using the picture–word interference paradigm, we report data from two PC manipulations in which contingency learning was made impossible by using nonrepeated distractors (Experiment 1A) or both nonrepeated distractors and responses (Experiment 1B). The classic PC effect emerged in both experiments. In addition, learning of temporal expectancies could not explain the present PC effects either, as results from trial-level analyses of Experiments 1A and 1B and a nonconflict version of Experiment 1B (Experiment 2) were inconsistent with the predictions of the temporal learning account of PC effects. These results suggest that conflict adaptation remains a credible explanation for PC effects.

Public Significance Statement

This study shows that people can adapt their attention to deal with situations in which there is frequent distraction, and it rules out alternative interpretations for this behavior.

Keywords: conflict adaptation, contingency learning, picture–word interference, proportion-congruent effect, temporal learning

An established fact in cognitive research is that goal-oriented behavior requires some form of control for selection of appropriate responses in the face of conflict coming from task irrelevant information. What is less established, however, is whether control can be adaptively modulated in response to experience with conflict. With such a conflict adaptation mechanism, the cognitive control system would, presumably, not just resolve conflict, but also monitor conflict and adapt attention to relevant and irrelevant information accordingly (Botvinick, Braver, Barch, Carter, & Cohen, 2001).

Manipulations of conflict frequency in interference tasks such as the Stroop (1935) task typically produce a pattern of results that is consistent with a conflict-adaptation explanation. In the classic (color–word) Stroop task, participants are required to name the ink

color of a word while ignoring the word itself. A congruency effect typically arises, with faster (and often more accurate) responding to congruent items (e.g., the word RED in red color, RED_{red}) than to incongruent items (e.g., the word RED in blue color, RED_{blue}; MacLeod, 1991). Of interest here is the fact that the magnitude of this effect varies as a function of experience with conflict. Specifically, situations in which the proportion of congruent items is high (i.e., infrequent conflict) elicit larger congruency effects than do situations in which the proportion of congruent items is low (i.e., frequent conflict; e.g., Crump, Gong, & Milliken, 2006; Jacoby, Lindsay, & Hessels, 2003; Logan & Zbrodoff, 1979; for a review, see Bugg & Crump, 2012). These Proportion-Congruent (PC) effects are readily explained by a conflict-adaptation process. When conflict is frequent, there is regular demand for the control system to maintain attention focused on the relevant dimension. Interference from the irrelevant dimension will thus be minimized, resulting in a reduced congruency effect. On the other hand, when conflict is infrequent, the benefit of focusing on the color is rarely reinforced. As a result, interference from the irrelevant dimension on the few incongruent items that are present results in a large congruency effect.

Recent years, however, have witnessed a growing concern among researchers about the validity of conflict adaptation as an explanation for PC effects (Schmidt, 2013b; Schmidt, Notebaert, & van den Bussche, 2015). Such concern has its roots in the

This article was published Online First May 30, 2019.

Giacomo Spinelli, Jason R. Perry, and Stephen J. Lupker, Department of Psychology, University of Western Ontario.

This research was partially supported by Natural Sciences and Engineering Research Council of Canada Grant A6333 to Stephen J. Lupker and the Ontario Trillium Scholarship to Giacomo Spinelli.

Correspondence concerning this article should be addressed to Giacomo Spinelli, Department of Psychology, University of Western Ontario, London, Ontario N6A 5C2, Canada. E-mail: gspinell@uwo.ca

realization that, in speeded tasks, responding might be influenced by learned associations, or contingencies, between a stimulus and a motor response (Musen & Squire, 1993; Schmidt, Crump, Cheesman, & Besner, 2007), as well as by the formation of temporal expectancies for the emission of a response (Schmidt, 2013c). The reason these issues are relevant is that PC manipulations are typically confounded with contingency-learning biases as well as with temporal-learning biases (Schmidt, 2013c; Schmidt & Besner, 2008). As such, some combination of these factors when applied to the mechanisms involved in interference tasks appears to be able to explain the PC effects that are observed in those tasks without needing to posit a role for conflict (Kinoshita, Mozer, & Forster, 2011; Levin & Tzelgov, 2016). What is worth noting at this point is that, as will be described subsequently, these alternative accounts are essentially *facilitation* accounts. That is, their explanations for PC effects are based on the idea that some aspect of processing is facilitated as a result of participants gaining relevant information about the nature of the task. Hence, these accounts offer a radically different view of PC effects than that offered by the conflict-adaptation account, which is based on an interference-driven mechanism.

Contingency Learning

Contingency learning involves acquiring knowledge that two events tend to occur together (e.g., the presentation of the word RED typically requires the response “green”) and using that knowledge to facilitate responding (Beckers, De Houwer, & Matute, 2007). In color–word identification tasks in which the words used are not color names, contingency learning is presumed to explain why color identification is faster for a frequent word–color pair (= high-contingency item, e.g., the word BRAG presented in green color 75% of the time) than for an infrequent word–color pair (= low-contingency item, e.g., the word BRAG presented in yellow color 25% of the time; Schmidt et al., 2007; see also Musen & Squire, 1993). Essentially, according to contingency-learning accounts, participants implicitly learn contingencies between words and color responses, that is, that specific words predict specific color responses (e.g., BRAG predicts green; Schmidt et al., 2007; see also Forrin & MacLeod, 2017; Lin & MacLeod, 2018), allowing them to respond more rapidly when the word appears in its most frequent color.

The reason this issue is relevant for PC effects is that manipulating the proportion of congruent items in the Stroop task typically involves altering the frequency of specific word–color pairs as well. For example, PC experiments might involve a Mostly Congruent (MC) list in which the word RED appears in its congruent, red color 75% of the time and in the incongruent, blue color 25% of the time, and a Mostly Incongruent (MI) list in which the word RED appears in the incongruent, blue color 75% of the time and its congruent, red color 25% of the time. Doing so, however, means that RED_{red} is more frequent than RED_{blue} in the MC list, whereas RED_{blue} is more frequent than RED_{red} in the MI list. If frequent word–color pairs elicit faster responses, participants might thus speed up on RED_{red} in the MC list and RED_{blue} in the MI list. Crucially, fast responding to the congruent item RED_{red} in the MC list will lead to a relatively large congruency effect, whereas fast responding to the incongruent item RED_{blue} in the MI list will lead to a relatively small congruency effect. In

other words, learning of word–color contingencies, rather than adaptation to conflict frequency, might be responsible for the difference in magnitude of congruency effects that is typically found in PC manipulations in the Stroop task (Schmidt & Besner, 2008).

Importantly, the assumption that contingency learning is the only source of PC effects (Schmidt & Besner, 2008) implies that no PC effects should be observed in PC manipulations that control for contingency learning operations. In an effort to address this issue, a number of studies have been conducted that evaluate PC effects on contingency-controlled stimuli, that is, stimuli which are matched in contingency across MC and MI lists (Blais & Bunge, 2010; Bugg, 2014a; Bugg & Chanani, 2011; Bugg, Jacoby, & Toth, 2008; Gonthier, Braver, & Bugg, 2016; Hutchison, 2011). The rationale is that if PC effects are driven by a mechanism of adaptation to list-wide conflict frequency, that mechanism should produce a PC effect for all stimuli, including the contingency-controlled stimuli. Results from those studies do provide at least partial support for this prediction, with PC effects on contingency-controlled stimuli being reported in a number of circumstances (Bugg, 2014a; Bugg & Chanani, 2011; Gonthier et al., 2016; Hutchison, 2011), although not in all circumstances (Blais & Bunge, 2010; Bugg et al., 2008). Therefore, contingency learning by itself does not appear to offer a complete explanation of PC effects, allowing proponents of the conflict-adaptation account to argue that these contingency-controlled PC effects, when obtained, likely reflect the action of a mechanism of adaptation to conflict frequency (Bugg, 2014a). In contrast, Schmidt (2013c) has contended that those effects are better explained by a different, non-conflict learning process—temporal learning.

Temporal Learning

Whereas contingency learning is about using stimulus information to predict *what* to respond, temporal learning refers to the process of learning *when* to emit a response. Participants in speeded tasks are known to establish something like a time criterion for when to respond (i.e., the point in time at which they expect to respond) depending on the characteristics of the stimuli. For example, relatively easy stimuli are typically responded to faster when presented in a list where all of the stimuli are easy (i.e., a pure list) than when presented intermixed with harder items (i.e., a mixed list), suggesting that participants adapt their temporal expectations for response emission to the average difficulty experienced in the list (Lupker, Brown, & Colombo, 1997; Lupker, Kinoshita, Coltheart, & Taylor, 2003). Recently, Schmidt (2013c) extended this idea to explain PC effects in the Stroop task that have been obtained in the absence of biases created by contingency learning (Bugg, 2014a; Hutchison, 2011).

According to Schmidt’s (2013c) temporal learning account, participants will develop a relatively fast temporal expectancy in an MC list (because most of the items in the list elicit relatively fast responses) and a relatively slower temporal expectancy in an MI list (because most of the items in the list elicit relatively slow responses). Participants will then use those temporal expectancies to anticipate when a response should be emitted. Specifically, congruent items, but not incongruent items, will speed up in the MC list because they can be processed rapidly enough to meet the fast temporal expectancy established for that list. As a result,

the congruency effect will be relatively large in the MC list. Conversely, in an MI list, participants anticipate responding late, and as a result, there will be less pressure on them to elicit fast responses to congruent items. This situation would cause slower latencies to congruent items in an MI list relative to congruent items in an MC list. In contrast, according to Schmidt's account, a speed-up could potentially be observed for incongruent items in an MI list because they can be processed fast enough to meet the (slower) temporal expectancy established for that list. The result would be a relatively small congruency effect. In practice, however, hard-to-process stimuli appear to be relatively insensitive to temporal expectancies, at least in some situations (Kinoshita & Mozer, 2006; Kinoshita et al., 2011), meaning that the slow temporal expectancy developed for the MI list may have little impact on incongruent items. In any case, the core claim here is that learning of temporal expectancies can inflate the congruency effect in the MC list in comparison to the congruency effect in the MI list. Thus, similar to contingency learning, temporal learning can explain differences in the magnitude of congruency effects across MC and MI lists without invoking any type of conflict-adaptation mechanism.

A critical piece of evidence in support of the temporal learning account of PC effects comes from statistical analyses of PC manipulations that take into account the role of temporal expectancies that individuals develop on a trial-by-trial basis. The idea for these analyses was first proposed by Kinoshita et al. (2011) within the framework of their Adaptation to the Statistics in the Environment (ASE) model of optimal response initiation and was then extended by Schmidt (2013a) in his Parallel Episodic Processing (PEP) model of color identification (see also Schmidt, De Houwer, & Rothermund, 2016). Although the two models were developed to explain different phenomena (relatedness proportion effects in masked priming in the case of the ASE model, PC effects in regular Stroop paradigms in the case of the PEP model) and place emphasis on different aspects of response emission (adaptation to perceived difficulty in the case of the ASE model, rhythmic responding in the case of the PEP model), the models make similar assumptions. First, performance on the current trial is influenced by the participant's knowledge of response times on the previous trials (i.e., the trial history), specifically the latencies on the most recent trials. Those latencies, especially the latency on the most recent trial (response time [RT] on trial $n - 1$), would function as an index of perceived task difficulty (in the ASE model) or as an index of the rhythm of responding (in the PEP model) that could be used to form an expectancy for response initiation latency on trial n . Thus, RT on trial $n - 1$ can function as an index of temporal expectancy for trial n , with a slower RT on trial $n - 1$ leading to a slower RT on trial n (Kiger & Glass, 1981; Taylor & Lupker, 2001). Second, as noted, easier stimuli are more prone to influences from trial history than harder stimuli (although this pattern is not inevitable: Kinoshita & Mozer, 2006; Kinoshita et al., 2011). The critical implication of these assumptions is that with easy stimuli strongly affected by RT on trial $n - 1$ (i.e., they will show a large slow-down following a slower RT on trial $n - 1$) and hard stimuli only weakly affected by RT on trial $n - 1$ (i.e., they will not show a large slow-down following a slower RT on trial $n - 1$), difficulty effects (i.e., the time difference between hard and easy stimuli) will decrease as RT on trial $n - 1$ increases.

Evidence for this pattern has been obtained from experimental data that were analyzed using linear mixed-effects models. This class of models, unlike traditional means-based ANOVAs, allows one to evaluate the impact of RT on trial $n - 1$, a trial-level continuous predictor, on performance on trial n . In several investigations, use of those analyses revealed that difficulty effects caused by visible or even subliminal distractors were modulated by trial history, with there being smaller effects when the RTs were slower on trial $n - 1$ (Huber-Huber & Ansorge, 2017, 2018; Kinoshita et al., 2011; Schmidt, 2013c; Schmidt & Weissman, 2016). Most importantly for the present discussion, the fact that congruency effects (and difficulty effects in general) are modulated by temporal expectancies is relevant for PC manipulations because fast RTs inevitably occur more frequently in MC lists than in MI lists. As faster RTs on trial $n - 1$ result in larger congruency effects, MC lists will tend to produce larger congruency effects than MI lists independent of contingency-learning biases or a presumed conflict-adaptation mechanism.

Support for the idea that temporal learning is at least partially responsible for PC effects comes from Schmidt's (2013c) reanalysis of the data from Hutchison's (2011) contingency-controlled items using the aforementioned linear mixed-effects model analyses. Those analyses not only replicated the finding originally reported by Hutchison (2011) of a significant PC effect, but also indicated that congruency effects decreased with increasing RT on trial $n - 1$. Furthermore, this decreased congruency effect was accompanied by a reduction (although not an elimination) of the value of the beta parameter for the PC effect (i.e., the interaction) in the model, suggesting that PC effects and temporal learning effects explain common variance in the data. Schmidt interpreted this finding as indicating that temporal learning has the potential of generating PC effects on its own, a point he reinforced by showing that his PEP model, in which temporal learning was an implemented mechanism but adaptation to conflict frequency was not, could simulate Hutchison's (2011) results. At the very least, Schmidt's analysis suggests that temporal learning contributes to PC effects in contingency-controlled situations and, therefore, its role needs to be considered when analyses of PC manipulations are conducted.

More recently, however, Cohen-Shikora, Suh, and Bugg (2018) challenged this conclusion. They noted that the critical interaction between congruency and RT on trial $n - 1$ reported by Schmidt (2013c) was obtained when the typical positively skewed RT distribution was normalized with an inverse transformation ($\ln vRT = -1000/RT$) to accommodate the assumption made by linear mixed-effects models that the dependent variable be normally distributed. A somewhat neglected downside of this type of analysis procedure is that nonlinear transformations of the dependent variable systematically alter the pattern and size of interaction terms, casting doubt on the reliability of analyses of interactions (Balota, Aschenbrenner, & Yap, 2013).

A solution to this problem is offered by generalized linear mixed-effects models, models which do not assume a normally distributed dependent variable and require, therefore, no RT transformation (Lo & Andrews, 2015). Using both inverse-transformed RTs in a linear mixed-effects model and untransformed (i.e., raw) RTs in a generalized linear mixed-effects effects model, Cohen-Shikora et al. (2018) reanalyzed Hutchison's (2011) dataset along with two additional data sets in which a PC manipulation had been

implemented while controlling for contingencies (i.e., Bugg, 2014; Gonthier et al., 2016). They reported that Schmidt's (2013c) finding that congruency effects decrease with increasing RT on trial $n - 1$ was only obtained with transformed data but not with untransformed data, with the latter data even providing evidence for the opposite pattern in some cases (i.e., congruency effects increased, rather than decreased, with increasing RT on trial $n - 1$). Furthermore, in all the data sets, the PC effect remained significant when temporal learning indices were included in the analyses, even when the value of the beta parameter for that effect was reduced due to the introduction of those indices. Finally, attempts to improve indices of temporal learning (e.g., by using mean RT on the three most recent trials as a predictor in the analyses) also yielded little evidence for the temporal learning account.

In sum, Cohen-Shikora et al.'s (2018) analyses suggest that previously reported evidence in support of the temporal learning explanation of the PC effect (Schmidt, 2013c) might have been biased by the nonlinear transformation applied to RT data. Therefore, it would be advisable that research aiming to control for temporal learning avoids this bias by using a more appropriate statistical technique, such as generalized linear mixed-effects modeling.

The Present Research

Although conflict-adaptation and nonconflict learning mechanisms are not necessarily mutually exclusive (Abrahamse, Braem, Notebaert, & Verguts, 2016; Egner, 2014), there has been a mounting debate in recent years concerning whether the classic empirical markers of conflict adaptation are, in fact, actually produced by nonattentional learning biases (e.g., contingency learning, temporal learning), biases that are typically found in manipulations designed to investigate what are presumed to be conflict adaptation effects. For some researchers, such debate has culminated in the idea that conflict adaptation might be an illusion (Schmidt et al., 2015). The fact that conflict adaptation tests are routinely used in clinical settings (e.g., Abrahamse et al., 2016; Bonnin, Houeto, Gil, & Bouquet, 2010) hints at the profound consequences borne by this idea. Presumed markers of conflict adaptation have been reported across the life span (e.g., Bugg, 2014a, 2014b) and across a variety of tasks (Bugg & Crump, 2012), with increasing reports coming from neuroimaging research (e.g., Braver, 2012; Sheth et al., 2012; West & Alain, 2000; Wilk, Ezekiel, & Morton, 2012), and these markers have been used in a number of diagnostic situations. An exact understanding of what these findings reflect is therefore crucial.

Motivated by these considerations, the present research aimed to reexamine the PC effect while at the same time accounting for potential nonconflict learning confounds. Specifically, we were interested in providing an answer to the following question: Would evidence for adaptation to conflict frequency emerge when nonconflict learning biases are controlled for or removed from the design altogether? Some attempts undertaken in this direction suggest that the answer might be "yes" (Bugg, 2014a; Bugg & Chanani, 2011; Bugg & Hutchison, 2013; Bugg, Jacoby, & Chanani, 2011; Gonthier et al., 2016; Hutchison, 2011). However, much of that research fails to consider nonconflict learning biases in their entirety and/or is based on experiments that deviate considerably from the original PC paradigm (Schmidt, 2013b, 2014a).

In addition, very few attempts have been made to control for temporal learning when analyzing PC effects (Cohen-Shikora et al., 2018; Schmidt, 2013c).

One primary objective of the present research was to examine adaptation to conflict frequency in a situation in which contingency learning could not contribute to the PC effect. According to Schmidt (2013a), learning of word–response contingencies is a two-step process: First, on each trial participants encode information about the word, the color, and the response made into episodic memory. Second, any word presented on a subsequent trial will lead to the retrieval of past episodes involving that word, with facilitation occurring if the currently presented word requires the same response as most of its previous occurrences. Note that repetition appears to be an important aspect of this process. Words need to be repeated at least a few times in the experiment for responses associated with them to be able to influence subsequent behavior (Lin & MacLeod, 2018). Because this process is based on learning a predictive association between a specific word and a specific response (Schmidt et al., 2007; but see Schmidt, Augustinova, & De Houwer, 2018), learning of word–response contingencies would thus seem to require repeating words in the relevant colors. In other words, contingency learning would be impossible without repeated word distractors.

Based on these considerations, the present experiments examined whether PC effects emerge in a PC manipulation in a Stroop-like task where no contingency learning would be possible because word distractors were never repeated (for a similar argument applied to a context-specific PC manipulation, see King, Korb, & Egner, 2012; see also Schneider, 2015, for a similar idea applied to cued task switching). Because only a limited set of words and colors can be used in the color–word Stroop task, a variant, the picture–word interference task, was used instead.¹ Experiment 1 involved two picture–word interference tasks in which the proportion of congruent trials was manipulated in a list-wide fashion (for a similar manipulation in the picture–word interference task, see Bugg & Chanani, 2011). Experiment 1A required participants to categorize unrepeated target pictures paired with unrepeated word distractors. Participants in Experiment 1B were presented with the same materials but were required to name the pictures instead. To preview the results, regular PC effects were obtained in both tasks.

Another objective of the present experiments was to investigate the role of temporal learning in PC manipulations. To accomplish this goal, the data from Experiments 1A and 1B were analyzed using RT on trial $n - 1$ as an index of temporal expectancy in a mixed-effects model analysis, similar to those of Schmidt (2013c) and Kinoshita et al. (2011). However, similar to Cohen-Shikora et al. (2018), generalized linear mixed-effects models rather than linear mixed-effects models were used in these analyses. The reason is that, as noted, RTs typically violate the assumption made

¹ Note that although there has been some debate in the literature as to whether interference effects in the picture–word interference task reflect the same underlying cognitive processes as interference effects in the color–word Stroop task (Dell'Acqua, Job, Peressotti, & Pascali, 2007), abundant support exists for a functional equivalence of the two paradigms (Lupker, 1979; Schnur & Martin, 2012; Starreveld & La Heij, 2017; van Maanen, van Rijn, & Borst, 2009), suggesting that the picture–word interference task can afford substantially larger target, distractor, and response sets than the color–word Stroop task without otherwise altering the cognitive processes engaged in the original paradigm.

by linear mixed-effects models of a normally distributed dependent variable, a problem many researchers, including Schmidt (2013c), addressed by normalizing RTs with an inverse transformation. However, as shown by Cohen-Shikora et al. (2018), this solution is inappropriate when the research interest lies in interaction terms, as those terms are typically altered by nonlinear transformations of the dependent variable. Generalized linear mixed-effects models, on the other hand, provide a better solution in that, making no assumption about the distribution of the dependent variable, they require no RT transformation and permit a clearer interpretation of interactions (Lo & Andrews, 2015). In addition to the analyses reported by Cohen-Shikora et al. (2018), the present analyses can thus shed further light on the questions of whether evidence for temporal learning will be maintained when a more appropriate statistical technique is used and, more crucially, how potential PC effects obtained in Experiments 1A and 1B relate to the effects of temporal learning.

To further strengthen the conclusions from those analyses, Experiment 2 was conducted to isolate potential effects of temporal learning from adaptation to conflict frequency in a conflict-free picture naming task. As will be discussed below, trial-level analyses of Experiments 1A and 1B and results from Experiment 2 provide converging evidence that temporal learning does not appear to pose a challenge for conflict-adaptation interpretations of PC effects, at least with the materials and tasks used in the present experiments.

Experiments 1A and 1B

If repetitions of word distractors are necessary for learning associations between specific words and specific responses, learning of such associations should be impossible when word distractors are never repeated. Such a situation should thus allow researchers to examine potential effects of adaptation to conflict frequency in the absence of the contingency-learning confound that is typically found in classic PC manipulations using the color-word Stroop task (Melara & Algom, 2003; Schmidt & Besner, 2008). As noted, to this end, a picture-word interference task was used. In the picture-word interference task, participants are required to identify a picture while ignoring a word superimposed on it. Similar to the color-word Stroop task, two types of items were used in the task variant employed in the present set of experiments: Congruent items, with words specifying the name of the picture itself (e.g., the picture of a dog with the word DOG superimposed on it), and incongruent items, with words unrelated to the picture (i.e., belonging to a different semantic category than the picture's), as well as not appearing as target pictures in the experiment (e.g., the picture of a dog had the unrelated word BED superimposed on it and no picture of a bed appeared in the experiment).

Using a between-subjects PC manipulation, participants were assigned to either an MI or an MC list.² Experiment 1A required participants to identify unrepeated target pictures paired with unrepeated word distractors as members of a semantic category and to respond vocally. Participants in Experiment 1B were presented with the same materials but were required to name the pictures instead. Note that despite the materials being the same, the word distractors used are more relevant to picture naming than they are to categorization. For example, the word DOG should help more

with naming the picture of a dog than it should help with categorizing a dog as an animal. Furthermore, unlike in picture naming, in picture categorization not only incongruent words but also congruent words are absent from the response set. Indeed, in a picture categorization task, Lupker and Katz (1981) obtained only a (nonsignificant) 12-ms difference between conditions that are analogous to the congruent and incongruent conditions of the present experiment. In contrast, picture naming was expected to elicit a much larger congruency effect because of the relevance of word distractors to the task (e.g., Underwood, 1976). Nonetheless, both picture categorization and picture naming were used in order to investigate whether the presence of PC effects might depend on the nature of the task and the basic magnitude of the congruency effect.

In response to the suggestions of two reviewers of the initial version of the present paper, we examined not only the PC effect but also the congruency sequence effect, that is, the finding that in interference tasks, congruency effects are larger following a congruent trial than following an incongruent trial (Gratton, Coles, & Donchin, 1983). Traditionally thought of as a marker of conflict adaptation (e.g., Botvinick et al., 2001), this finding, similar to the PC effect, has recently received several alternative interpretations (e.g., Hommel, Proctor, & Vu, 2004; Mayr, Awh, & Laurey, 2003), including a temporal learning interpretation (Schmidt & Weissman, 2016). This temporal learning interpretation partially relies on the same interaction that is thought to be responsible for the PC effect (i.e., decreasing congruency effects with higher RT on trial $n - 1$), an interaction that, crucially, Schmidt and Weissman (2016) observed when analyzing inverse RTs. As noted, such a situation makes the interpretation of interaction terms dubious. As such, our use of a generalized linear mixed model analysis, an analysis which permits usage of untransformed RTs, provided a valuable opportunity to assess whether the unreliability of the temporal learning interaction reported by Cohen-Shikora et al. (2018) in the context of the PC effect also applies in the context of the congruency sequence effect, another important marker of conflict adaptation. These additional analyses for Experiments 1A and 1B, along with a discussion of the control and the temporal learning account of the congruency sequence effect, can be found in the Appendix.

Method

Participants. An a priori power analysis was performed using G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) to calculate the sample size needed to have a power of .80 for obtaining a PC effect. Based on the effect size reported by Bugg and Chanani (2011) for a PC effect using contingency-controlled items in a picture-word interference task, we determined that a minimum sample size of 32 participants would be needed. Forty-eight participants took part in Experiment 1A (picture categorization) and another 51 took part in Experiment 1B (picture naming). In Experiment 1B, one participant was removed because of an equipment failure and two more were removed because of an excessive number of errors and null responses (above 25%), leaving 48 participants. Participants were all students at the University of

² Proportion of congruent items was manipulated between-subjects because of the limited number of items available.

Western Ontario aged 18–23 years ($SD = 1.03$) and had normal or corrected-to-normal vision. All were native English speakers. Their participation was compensated with course credit.

Materials. One hundred twenty-five line drawings were sourced from the International Picture Naming Project (IPNP) database (Szekely et al., 2005). Nineteen pictures from the Internet matching as closely as possible the style of the IPNP pictures were added to the set, for a total of 144 target pictures, 480×480 pixels in size. Of these, 36 represented an animal, 36 represented a human being, 36 represented some type of food, and 36 represented a man-made object. IPNP norms and pilot testing ensured that there was high agreement among English-speaking individuals on the semantic category and name of the pictures. One hundred forty-four English word distractors, different from the modal names of the pictures, were also selected. As with the target pictures, 36 denoted an animal, 36 a human being, 36 some type of food, and 36 a man-made object. The word distractors were matched in length and CELEX frequency (Baayen, Piepenbrock, & van Rijn, 1993) with the pictures' modal names. Each picture was paired with the modal name of the picture (congruent item) and with an unrelated word belonging to one of the other three categories (incongruent item), with each of the incongruent categories being equally represented across items. For example, among the 36 pictures of animals, 12 were paired with an unrelated word denoting a person, 12 with an unrelated word denoting food, and 12 with an unrelated word denoting an object. Powerpoint software was used to superimpose the word in 32-point Courier New font in the center of the picture. A light white glow around words' letters was added to ensure the word was clearly visible. A sample of the stimuli used in Experiments 1A and 1B is presented in Figure 1.

Lists were constructed so that for half of the lists, 25% of the items were congruent (MI lists), and for the other half, 75% of the items were congruent (MC lists). Specifically, in the MI lists, 36 pictures were presented with their congruent word, and 108 pic-

tures with their incongruent word. Conversely, in the MC lists, 108 pictures were presented with their congruent word, and 36 pictures with their incongruent word. Each of the semantic categories was equally probable among the congruent and among the incongruent items (e.g., in the MC list, nine of the 36 incongruent pictures were animals, nine were human beings, nine were food items, and nine were objects, etc.). Similarly, each of the three semantic categories of incongruent word distractors was equally probable among incongruent items (e.g., in the MC list, three of the nine incongruent animal pictures appeared with an unrelated word denoting a person, three with an unrelated word denoting a food item, and three with an unrelated word denoting an object).

Lists were also counterbalanced so that each picture appeared with its congruent and incongruent word distractor in both MI and MC lists. To this end, the pictures were randomly divided into four sets: A, B, C, and D. In List 1 of the MC lists, pictures in sets A, B, and C would serve as congruent pictures and pictures in set D would serve as incongruent pictures. In List 2 of the MC lists, pictures in sets A, B, and D would serve as congruent and pictures in set C would serve as incongruent, and so on. Construction of the MI lists was done similarly, with List 1 having pictures in sets A, B, and C serving as incongruent and pictures in set D serving as congruent, and so forth. Pictures in each set included an equal number of pictures of animals, people, food, and objects. Overall, four MI and four MC lists were constructed.

Procedure. Participants were tested individually in a quiet room, seated approximately 60 cm away from a monitor upon which the stimuli were presented. Each trial began with a fixation symbol (“+”) displayed for 500 ms in the center of the screen, followed by a picture with a word superimposed on it, displayed for 3000 ms or until the participant's response. Responses were recorded with a microphone connected to the testing computer. Participants in Experiment 1A were instructed to categorize the picture using one of four semantic categories (ANIMAL, PER-



Figure 1. Sample stimuli used in Experiments 1A and 1B. Represented are congruent items (A) and incongruent items (B) for each of the four categories (ANIMAL, PERSON, FOOD, and OBJECT). In this sample, the pictures of the elephant and the shovel come from the International Picture Naming Project (Szekely et al., 2005) whereas the pictures of the astronaut and bacon were sourced from the Internet.

SON, FOOD, OBJECT) as responses. Care was taken to explain the differences between these categories to minimize potential ambiguities (e.g., living animals that are typically eaten by humans, such as chicken, being classified as food). Participants in Experiment 1B were instructed to name the picture instead. In both experiments, participants were told to ignore the word superimposed on the picture and respond as quickly and as accurately as possible. Participants were randomly assigned to one of the eight lists in Experiments 1A and 1B. Thus, each participant performed only one list for a total of 144 trials.

Prior to the experiment, participants performed a practice session involving 12 items, different from the items in the experiment and mirroring the proportion of congruent items in the upcoming list. They received no feedback. Trials were presented in a different random order for each participant. DMDX (Forster & Forster, 2003) software was used to present the stimuli and collect the data.

Results

In these experiments as well as in Experiment 2, response waveforms were manually inspected with CheckVocal (Protopapas, 2007) to determine the accuracy of the response and the correct placement of timing marks. RTs were defined as the time interval between stimulus onset and the beginning of the vocal response. Errors were marked using a conservative criterion: Any response that was not the expected response was considered an error, no matter how close it was to the expected response (e.g., “people” instead of “person” for Experiment 1A, or “cop” instead of “policeman” in Experiment 1B). Prior to the analyses, invalid trials attributable to technical failures and responses faster than 300 ms or slower than the time limit (accounting for 0.4% and 2% of the data points in Experiments 1A and 1B, respectively) were discarded. For the latency analyses, trials for which an error was made on the current trial were discarded, as were the trials for which an error or a too-fast response (<300 ms) or a too-slow response (>3000 ms) was made on the preceding trial.

The latencies and the error rates were analyzed using generalized linear mixed-effects modeling in R Version 3.4.3 (R Core Team, 2015), treating subjects and items (i.e., the target pictures) as random effects and treating Congruency (congruent vs. incongruent) and List Type (MI vs. MC) as within-subject and between-subject fixed effects, respectively (Baayen, 2008; Baayen, Davidson, & Bates, 2008). Prior to running the model, R-default treatment contrasts were changed to sum-to-zero contrasts (i.e., *contr.sum*) to help interpret lower-order effects in the presence of higher-order interactions (Levy, 2014; Singmann & Kellen, *in press*). The model was fit by maximum likelihood with the Laplace approximation technique. The lme4 package, Version 1.1–15 (Bates, Mächler, Bolker, & Walker, 2015), was used to run the generalized linear mixed-effects model and obtain probability values.

In the latency analyses, a generalized linear mixed-effects model was used instead of a linear mixed-effects model because generalized linear models, unlike linear models, do not assume a normally distributed dependent variable. Therefore, these models can accommodate the typically positively skewed distribution of raw RT data with there being no need to use nonlinear transformations, known to systematically alter interaction terms (Balota et al., 2013). A Gamma distribution was used to fit the raw RTs, with an

identity link between fixed effects and the dependent variable (Lo & Andrews, 2015). Note that convergence tests for generalized linear mixed-effects models in the current version of lme4 tend to generate many false positives (Bolker, 2018). In the following, we report the data from the BOBYQA optimizer, which returned estimates that were equivalent to other optimizers but never issued convergence warnings. Unlike the error analyses, latency analyses included RT on trial $n - 1$ as a fixed effect to control for temporal learning (Schmidt, 2013c; Schmidt & Weissman, 2016). Standardized (i.e., centered and scaled) RTs on trial $n - 1$ were used instead of raw RTs to avoid spurious correlations between the intercept and the slope and to help evaluating and interpreting the model (Bolker, 2018; Kinoshita et al., 2011; Schielzeth, 2010). The statistical model for the latency analysis was: $RT = \text{glmer}(RT \sim \text{congruency} * \text{list_type} * \text{SprevRT} + (1|\text{subject}) + (1|\text{item}), \text{family} = \text{Gamma}(\text{link} = \text{"identity"}))$, control = $\text{glmerControl}(\text{optimizer} = \text{"bobyqa"})$). The statistical model for the error rate analysis was: $\text{Accuracy} = \text{glmer}(\text{accuracy} \sim \text{congruency} * \text{list_type} + (1|\text{subject}) + (1|\text{item}), \text{family} = \text{binomial}, \text{control} = \text{glmerControl}(\text{optimizer} = \text{"bobyqa"}))$. The mean RTs and error rates based on by-subject data for Experiments 1A and 1B are shown in Table 1. Scatterplots visualizing the relation between RT on trial $n - 1$ and the congruency effect on trial n are shown in Figure 2 for Experiment 1A and in Figure 3 for Experiment 1B. The data and R scripts used for the analyses are publicly available at <https://osf.io/jnzgb/>.

Experiment 1A (picture categorization).

Response time. There were significant main effects of Congruency (congruent faster than incongruent), $\beta = -10.68$, $SE = 2.36$, $z = -4.53$, $p < .001$, and RT on trial $n - 1$ (faster responses with lower RT on trial $n - 1$), $\beta = 23.85$, $SE = 2.82$, $z = 8.44$, $p < .001$. The interaction between Congruency and List Type was significant as well, $\beta = -14.14$, $SE = 2.54$, $z = -5.56$, $p < .001$, indicating that a classic PC effect was obtained, with a larger effect of Congruency in the MC (54 ms) than in the MI condition (-2 ms). Interestingly, the interaction between Congruency and RT on trial $n - 1$, indexing temporal learning, was not significant, $\beta = -1.62$, $SE = 2.61$, $z = -0.62$, $p = .54$, but the three-way interaction between Congruency, List Type, and RT on trial $n - 1$ was, $\beta = 6.01$, $SE = 2.60$, $z = 2.31$, $p = .021$.

To explore the three-way interaction, MC and MI lists were analyzed separately. MC lists showed both main effects of Congruency, $\beta = -24.89$, $SE = 3.52$, $z = -7.07$, $p < .001$, and RT

Table 1
Mean RTs and Percentage Error Rates (and Corresponding Standard Errors) for Experiments 1A and 1B

Congruency	RTs		Error rates	
	MC list	MI list	MC list	MI list
Experiment 1A				
Congruent	893 (24)	912 (31)	1.4 (.3)	2.1 (.6)
Incongruent	947 (24)	910 (37)	2.7 (.6)	2 (.4)
Congruency effect	54	-2	1.3	-.1
Experiment 1B				
Congruent	764 (26)	797 (21)	1.1 (.3)	2 (.9)
Incongruent	1106 (38)	1041 (22)	14 (1.7)	11.2 (.9)
Congruency effect	342	244	12.9	9.2

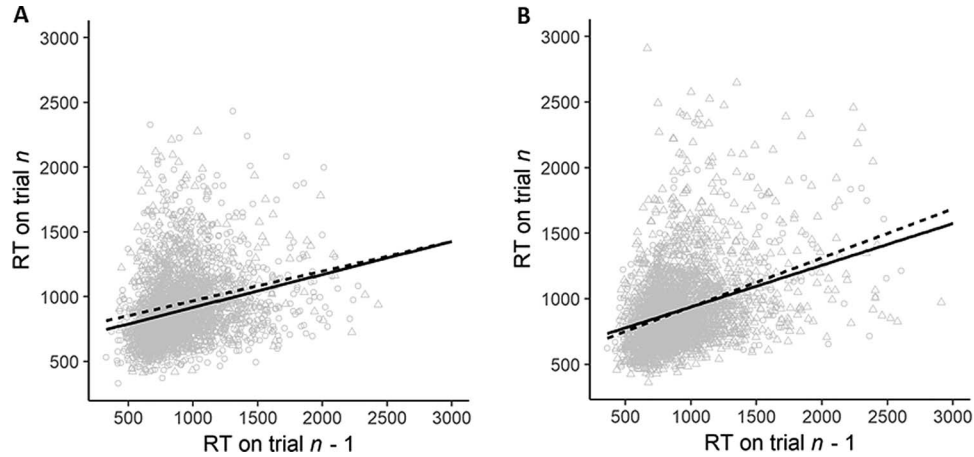


Figure 2. The impact of RT on Trial $n - 1$ on Congruency Effects on Trial n in Experiment 1A. The scatterplots represent the relation between RT on trial $n - 1$ and congruency effects on trial n in the MC list (A) and in the MI list (B). Individual observations for congruent and incongruent trials are marked with circles and triangles, respectively. Regression slopes for the congruent condition and for the incongruent condition are marked with solid and dashed lines, respectively.

on trial $n - 1$, $\beta = 25.60$, $SE = 3.99$, $z = 6.42$, $p < .001$, but no interaction between the two, $\beta = 4.51$, $SE = 3.75$, $z = 1.20$, $p = .23$. In MI lists, on the other hand, RT on trial $n - 1$ was significant, $\beta = 22.29$, $SE = 4.58$, $z = 4.86$, $p < .001$, but Congruency was not, $\beta = 3.26$, $SE = 3.66$, $z = .89$, $p = .37$. Here, the interaction between Congruency and RT on trial $n - 1$ was significant, $\beta = -8.08$, $SE = 3.91$, $z = -2.07$, $p = .039$. Note, however, that the pattern of this interaction is the opposite of that predicted by temporal learning: As illustrated in **Figure 2B**, the congruency effect on trial n increased, rather than decreased, with higher latencies on trial $n - 1$.

Error rates. Neither Congruency nor List Type was significant. The interaction between the two was marginal, $\beta = .17$, $SE = .10$, $z = 1.78$, $p = .075$, indicating a tendency for the

Congruency effect to be larger in the MC (1.3%) than in the MI condition (-.1%).

Experiment 1B (picture naming).

Response time. There were significant main effects of Congruency (congruent faster than incongruent), $\beta = -143.71$, $SE = 2.62$, $z = -54.80$, $p < .001$, List Type (MI faster than MC), $\beta = 24.96$, $SE = 4.39$, $z = 5.68$, $p < .001$, and RT on trial $n - 1$ (faster responses with lower RT on trial $n - 1$), $\beta = 22.84$, $SE = 3.04$, $z = 7.51$, $p < .001$. The only significant interaction was that between Congruency and List Type, $\beta = -23.64$, $SE = 2.89$, $z = -8.19$, $p < .001$, indicating that a classic PC effect was obtained, with a larger effect of Congruency in the MC (342 ms) than in the MI condition (244 ms). Neither the interaction between Congruency and RT on trial $n - 1$, $\beta = -3.88$, $SE = 2.66$, $z = -1.46$, $p = .14$,

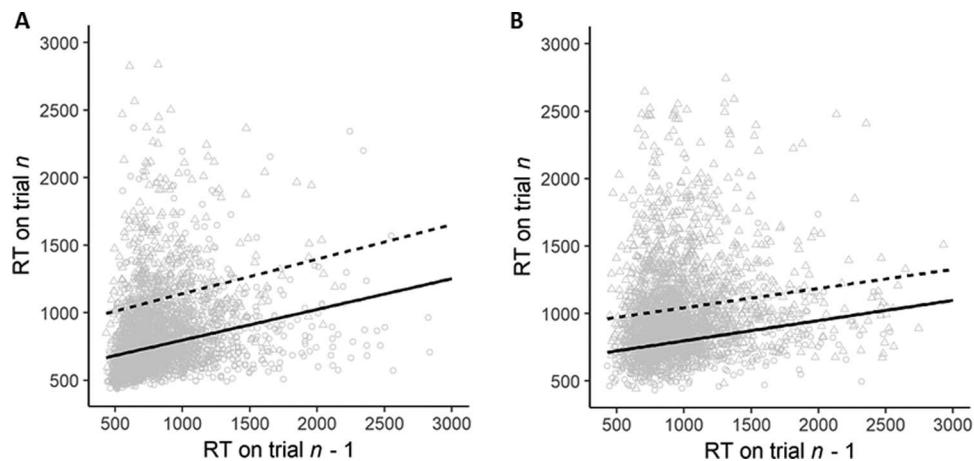


Figure 3. The Impact of RT on Trial $n - 1$ on Congruency Effects on Trial n in Experiment 1B. The scatterplots represent the relation between RT on trial $n - 1$ and congruency effects on trial n in the MC list (A) and in the MI list (B). Individual observations for congruent and incongruent trials are marked with circles and triangles, respectively. Regression slopes for the congruent condition and for the incongruent condition are marked with solid and dashed lines, respectively.

nor the three-way interaction between Congruency, List Type, and RT on trial $n - 1$, $\beta = -.18$, $SE = 3.01$, $z = -.06$, $p = .95$, was significant.

Error rates. There was a main effect of Congruency (congruent more accurate than incongruent), $\beta = 1.36$, $SE = .09$, $z = 14.71$, $p < .001$. In addition, Congruency interacted with List Type, $\beta = .27$, $SE = .09$, $z = 3.00$, $p = .003$, with the congruency effect being larger in the MC (13.0%) than in the MI condition (9.1%).

Discussion

Both Experiment 1A and Experiment 1B produced clear PC effects in a situation where learning of direct associations between words and responses was impossible. Note that, as suggested by previous findings (Lupker & Katz, 1981), the basic congruency effect was much smaller in Experiment 1A (picture categorization: 26 ms) than in Experiment 1B (picture naming: 293 ms). However, the congruency effect was similarly modulated by conflict frequency across the two tasks, with MI lists in Experiment 1B showing a congruency effect reduced by 98 ms compared with MC lists, and Experiment 1A showing the elimination of the congruency effect in MI lists. As discussed, a contingency learning account would not be able to explain these effects.

Temporal learning also does not seem to offer a reasonable explanation for the present findings. For temporal learning to account for PC effects, one would need to find that congruency effects on trial n get smaller as RT on trial $n - 1$ increases, indicating that participants use previous experience in the task to form and adjust to temporal expectancies for responding in the way suggested by Schmidt (2013c). Using generalized linear mixed-effects models to fit raw RTs, robust main effects of RT on trial $n - 1$ were found, with overall slower responses on trial n as RT on trial $n - 1$ increases. These sequence effects are routinely reported in speeded tasks (Kinoshita et al., 2011; Taylor & Lupker, 2001). More importantly, no interaction between RT on trial $n - 1$ and the congruency effect on trial n was found in Experiment 1B, whereas a complicated pattern emerged in Experiment 1A. Specifically, in Experiment 1A, MI lists (but not MC lists) showed an interaction involving the opposite pattern than was expected from the temporal learning account, i.e., the size of the congruency effect on trial n increased as the RT on trial $n - 1$ increased. Although the cause for this result is unclear, it should be noted that Cohen-Shikora et al. (2018) also reported inconsistent temporal learning patterns across the three data sets they analyzed.³ In general, it is safe to conclude from the overall pattern of results that temporal learning could not have produced, or even contributed to the production of, the PC effects reported here.

In sum, Experiments 1A and 1B showed that PC effects emerge even in the absence of temporal learning and word-response contingencies, a finding that challenges the view that mechanisms of this sort provide a sufficient account of the PC effects that are reported in the literature and that adaptation to conflict frequency may not be a mechanism humans use (Schmidt, 2013b).

To consolidate the idea that temporal learning has little to do with the PC effect obtained in Experiments 1A and 1B, Experiment 2 was conducted to disentangle conflict frequency from potential effects of temporal learning. Note that Schmidt's (2013c) temporal learning account assumes that temporal expectancies for

responding are altered as a result of any manipulation that induces appreciable differences in response rhythm. The type of manipulation which can accomplish such an alteration involves changes in the relative frequency of easy and hard stimuli, with the nature of the difficulty elicited by those stimuli playing little or no role. Because difficulty does not need to derive from conflict from an irrelevant dimension, temporal learning should not be specific to the type of task used in Experiments 1A and 1B, that is, tasks where conflict/interference from an irrelevant dimension produces the difficulty effect. That is, according to the temporal learning account of PC effects, any task in which the proportion of easy and hard items is manipulated should produce differences in the temporal expectancies being formed for responses. As a result, the magnitude of difficulty effects should parallel the pattern observed for congruency effects in the PC effect: Smaller difficulty effects in lists where most of the items are hard and larger difficulty effects in lists where most of the items are easy (Schmidt, 2013c, 2014a, 2016). Experiment 2 tested this prediction for the pictures used in Experiments 1A and 1B, which were presented without the superimposed words and modified in such a way that they were easier or harder to respond to.

Experiment 2

Following Schmidt's procedure (Schmidt, 2013c; Schmidt & Weissman, 2016), in Experiments 1A and 1B temporal learning was accounted for in the analyses by using RT on trial $n - 1$ as an index of temporal expectancy, with a lower RT on trial $n - 1$ indicating a faster temporal expectancy for trial n . However, the predicted interaction between congruency and RT on trial $n - 1$, with smaller congruency effects the higher the RT on trial $n - 1$, was not found. In fact, Experiment 1A even produced evidence for a reversed interaction in MI lists, with larger congruency effects following higher RTs on trial $n - 1$. These results are in line with recent failures to obtain regular temporal learning effects using untransformed RTs (Cohen-Shikora et al., 2018), suggesting that the nonlinear transformations reported in previously published papers (Huber-Huber & Ansorge, 2017, 2018; Kinoshita et al., 2011; Schmidt, 2013c; Schmidt & Weissman, 2016) might have systematically biased the interaction of interest in the direction predicted by temporal learning.

Statistical quirks aside, however, it must be acknowledged that supporters of temporal learning accounts have pointed out that RT

³ Following a suggestion of one of the reviewers of an earlier version of this paper, we ran an additional analysis in an attempt to determine whether part of what would seem to be noise in Experiment 1A might have resulted from response speed varying across categories. Such variability could potentially have affected the temporal learning process and, consequently, PC effects. Indeed, participants were slower with the animal (930 ms), food (927 ms), and object (949 ms) categories than with the person category (830 ms), the category that also elicited the smallest overall congruency effect (5 ms vs. 44 ms, 27 ms, and 30 ms for animal, food, and object categories, respectively). However, there was no obvious relationship between the overall category latency (and/or the overall congruency effect within a category) and the size of the PC effect (i.e., the RT difference between the congruency effect in the MC list and the congruency effect in the MI list) that the category elicited (person: 22 ms; animal: -8 ms; food: 147 ms; object: 60 ms). Do note, however, that there was, unavoidably, considerable noise in this analysis, presumably because there were very few observations (nine or fewer) in some of the cells.

on trial $n - 1$ is likely a noisy approximation of temporal expectancies (Kinoshita et al., 2011; Schmidt, 2013c), although several attempts, reported by Cohen-Shikora et al. (2018), to use a less noisy index (e.g., mean RT on the three most recent trials) also failed to produce consistent evidence for the temporal learning account of the PC effect. In fact, one could argue that internally constructed temporal expectancies might deviate considerably from the measured response time on one or more of the preceding trials, an argument that might find support in the observation that time perception is often prone to biases (e.g., Taylor & Lupker, 2006, 2007). The implication is that the analyses performed for Experiments 1A and 1B might not provide the best means of determining whether temporal learning is a potential contributor to the PC effect observed in those experiments.

A better way to deal with this issue might be found in another approach used by Schmidt (2013c, 2014a, 2016) in his attempts to demonstrate a potential role for temporal learning in the PC effect, an approach that does not require using any index of temporal expectancy in the analyses and thus avoids the potential problems associated with the noisiness of such measures. Relying on the assumption that temporal learning should operate similarly in interference tasks and in tasks in which difficulty does not derive from interference from an irrelevant dimension (e.g., perceptual tasks), this approach involves manipulating the proportion of easy items in a task of the latter type.

Indeed, the existence of a temporal learning mechanism of the sort described by Schmidt (2013c) implies that any task in which the proportion of easy and hard items is manipulated should produce differences in the magnitude of effect sizes in ways that are compatible with the changes observed for congruency effects in the PC effect. Specifically, mostly easy (ME) lists (i.e., lists in which most of the items are relatively easy to process) will favor development of a fast temporal expectancy that can be met by items that allow fast responses (i.e., “easy” items), but not by items that are relatively hard to process (i.e., “hard” items). The result is a speed-up for only the easy items and, hence, a large difficulty effect. Mostly hard (MH) lists (i.e., lists in which most of the items are relatively hard to process), on the other hand, will favor development of a slow temporal expectancy. Because participants anticipate responding relatively late, there will be no reason for them to speed up responses to easy items in this situation, causing them to produce longer latencies. In contrast, as noted, it is possible that latencies for hard items may decrease if they can be processed fast enough to meet the slower temporal expectancy, although, as Schmidt (2013c) has argued, those items tend to be insensitive to temporal expectancies (see also Kinoshita et al., 2011; Schmidt & Weissman, 2016). The end result is that learning of temporal expectancies should produce larger difficulty effects in ME lists than in MH lists.

Schmidt (2013a) did, in fact, obtain evidence of such a Proportion-Easy (PE) effect in a number of studies where no irrelevant dimension was used (Schmidt, 2013c, 2014a, 2016). For example, in a letter identification task Schmidt (2014a) found, as would be expected, shorter latencies for high-contrast letters (easy items) than for low-contrast letters (hard items). Most importantly, the size of this difficulty effect was modulated by the proportion of easy items in the list, with larger difficulty effects in ME lists than in MH lists, similar to the PC effect in the Stroop task. Although this finding is not crucial evidence that the mechanism driving PC

effects in the Stroop task and PE effects in nonconflict tasks is the same, it does suggest that temporal learning might play an important role in determining PC effects (Schmidt, 2013c). Specifically, this approach provides a proof of principle that a PC-like effect can be obtained even when little or no conflict is present in the task, suggesting that the mechanism responsible for this PC-like effect might also be operating when conflict is present, for example, in Stroop paradigms.

The goal of Experiment 2 was to examine a similar, nonconflict situation with the pictures used in Experiments 1A and 1B. Similar to Schmidt’s (2013c, 2014a, 2016) use of high-contrast and low-contrast letters, high-resolution and low-resolution pictures were used as easy and hard items, respectively, and participants were assigned to a ME list where most of the pictures had a high resolution, or a MH list where most of the pictures had a low resolution. Following Schmidt’s (2013c) temporal learning account, it was hypothesized that easy items would be responded to faster in ME than MH lists, and hard items would be responded to faster (or, at least, no more slowly) in MH than ME lists. As a result, a PE effect would be obtained, with ME lists showing a larger difficulty effect than MH lists.

It is important to note, however, that a different outcome could be expected from an alternative temporal learning account, specifically, one derived from the literature on blocking effects (Chateau & Lupker, 2003; Lupker et al., 1997, 2003; Kinoshita & Mozer, 2006; Rastle, Kinoshita, Lupker, & Coltheart, 2003; Taylor & Lupker, 2001). Blocking effects refer to the finding that when relatively easy and relatively hard items are mixed in a block (i.e., a mixed block, typically with 50% easy and 50% hard stimuli), latencies tend to be more homogeneous compared to latencies for easy versus hard items presented by themselves in pure blocks (i.e., blocks where all of the stimuli are either easy or hard). Specifically, there is a mixing cost for easy stimuli (i.e., slower latencies for easy stimuli in mixed blocks than in pure easy blocks) and a mixing benefit for hard stimuli (i.e., faster latencies for hard stimuli in mixed blocks than in pure hard blocks). Lupker and collaborators interpreted this pattern as evidence that participants in speeded tasks establish a time criterion representing the time at which they expect, and will attempt, to initiate a response. Importantly, the placement of the time criterion is dependent upon the characteristics of the stimuli in the block: The criterion will be set early in a pure easy block, late in a pure hard block, and in an intermediate position in a mixed block.

This reasoning can be easily extended to comparisons among mixed lists varying in the proportion of easy items. That is, in ME lists, the criterion will be placed relatively early (although not as early as in a pure easy list), whereas in MH lists it will be placed relatively late (although not as late as in a pure hard list). As a result, both easy and hard items should be responded to faster in ME lists than in MH lists. In other words, under the assumption that adjustments of the time criterion are similar for easy and hard items, one might expect main effects of difficulty (easy faster than hard) and list type (ME faster than MH), but not necessarily their interaction, that is, difficulty effects may be equivalent in ME and MH lists. Of importance, the latter pattern (i.e., similar adjustments of the time criterion for easy and hard items) typically emerges in word naming tasks but not in (button-press) lexical decision tasks (in which only easy items appear to be affected by adjustments of a time criterion), even when using the same items in the two tasks

(Kinoshita & Mozer, 2006). Because the present experiments used naming, and blocking effects occur for pictures and words alike (Lupker et al., 2003), the expectation would be that the interaction predicted by the temporal learning account would not arise in the task investigated in Experiment 2 (i.e., picture naming).

Method

Participants. An a priori power analysis was performed using G*Power 3.1 (Faul et al., 2009) to calculate the sample size needed to have a power of .80 for obtaining a PE effect. Based on the effect size reported by Schmidt (2013c) for a PE effect in a letter identification task, we determined that a minimum of 68 participants would be needed. One hundred twelve participants took part in the experiment. Nineteen participants were removed because of an excessive number of errors and null responses (above 25%), leaving 93 participants. They were all students at the University of Western Ontario aged 18–27 years ($SD = 1.21$) and had normal or corrected-to-normal vision. All were native English speakers. They received either \$10 or course credit for their participation.

Materials. The materials were derived from those used in Experiments 1A and 1B. The congruent pictures with their superimposed word removed functioned as high-resolution, easy items. The incongruent pictures with their superimposed word removed were degraded by resizing them to a quarter of their size and then inflating them back to their original size with Bulk Resize Photos (<https://bulkresizephotos.com>), thus resulting in a lower-resolution image. Those pictures functioned as low-resolution, hard items. Other than high-resolution and low-resolution pictures replacing the congruent and incongruent pictures, respectively, lists and counterbalancing of the items were identical to those in Experiments 1A and 1B, resulting in four ME lists and four MH lists.

Procedure. The procedure was identical to that in Experiment 1B, with the exception that, of course, superimposed words were not mentioned in the instructions, and participants were simply required to name the pictures as quickly and as accurately possible.

Results

Analyses were performed in the same way as was done for Experiments 1A and 1B with the exception that the factor Congruency was replaced with the factor Difficulty (easy vs. hard) and the two levels of the factor List Type were ME and MH instead of MC and MI. In addition, in line with previous PE manipulations (Schmidt, 2013c, 2014a, 2016), RT on trial $n - 1$ was not included as a predictor in the latency analysis because there is no need to control for temporal learning in this context: Any differences between difficulty effects across the two list types should be produced by the learning of temporal expectancies induced by the Difficulty factor itself.⁴

Prior to the analyses, invalid trials attributable to equipment failures and responses faster than 300 ms or slower than the time limit, accounting for 3.3% of the data points, were discarded. Because RT on trial $n - 1$ was not used as a predictor in the latency analysis, only trials where an error was made on the current trial were discarded. The mean RTs and error rates are presented in Table 2. The data and R scripts used for the analyses are publicly available at <https://osf.io/jnzgb/>.

Reaction time. There were significant main effects of Difficulty (easy faster than hard), $\beta = -31.74$, $SE = 2.19$, $z = -14.53$,

Table 2
Mean RTs and Percentage Error Rates (and Corresponding Standard Errors) for Experiment 2

Difficulty	RTs		Error rates	
	ME list	MH list	ME list	MH list
Easy	908 (17)	948 (20)	9.1 (.5)	9.6 (.8)
Hard	972 (19)	1013 (18)	15 (1.1)	14.8 (.6)
Difficulty effect	64	65	5.8	5.1

$p < .001$, and List Type (faster responses in the ME than the MH condition), $\beta = -17.29$, $SE = 3.05$, $z = -5.67$, $p < .001$. However, Difficulty and List Type did not interact, $\beta = -.57$, $SE = 1.94$, $z = -.29$, $p = .77$, reflecting equivalent effects of Difficulty in the ME (64 ms) and MH lists (65 ms).

Error rates. The only significant effect was that of Difficulty, $\beta = .36$, $SE = .04$, $z = 9.99$, $p < .001$.

Discussion

In the present experiment, the difficulty of pictures, instead of word–picture congruency, was manipulated by using high- and low-resolution pictures, similar to the high- and low-contrast letters used by Schmidt (2013c, 2014a, 2016). Unlike Schmidt's results, however, difficulty effects were not any larger in lists where most of the trials were easy than in lists where most of the trials were hard. In fact, the magnitude of difficulty effects was identical in the two conditions, thus failing to replicate the pattern predicted by Schmidt's temporal learning account. Note, however, that the type of list participants were assigned to—ME or MH—did have an effect, with overall faster latencies in ME than MH lists. Thus, this pattern seems more consistent with the time criterion account (Lupker et al., 1997), according to which ME and MH lists should lead to relatively early and late time criteria, respectively, affecting latencies for easy and hard items in a similar way, at least in a naming situation. Most importantly, this pattern is consistent with the analyses performed for Experiments 1A and 1B in indicating that temporal learning may have little or no role in modulating difficulty effects in both interference and noninterference tasks.

General Discussion

Do humans adapt to conflict frequency? Recently, some researchers have cast doubt on this idea by pointing out that PC effects in the Stroop task might be caused by factors other than conflict adaptation, namely, word–response contingency learning and temporal learning (Schmidt, 2013b). The present research addressed this question using a picture–word interference task where contingencies were eliminated and temporal learning was

⁴ The model with RT on trial $n - 1$ as an additional predictor did not alter the pattern of results reported (i.e., there were main effects of Difficulty and List Type but no interaction between them). For the interested reader, there was a main effect of RT on trial $n - 1$ (with higher RT on trial $n - 1$ leading to longer latencies on trial n) but no interaction between Difficulty and RT on trial $n - 1$. None of the other interactions were significant either.

controlled for. Clear, contingency-free PC effects emerged in both picture categorization (Experiment 1A) and picture naming (Experiment 1B) tasks, a finding that challenges the view that contingency learning is a critical factor driving PC effects in the Stroop task. Similarly, the analysis of the impact of trial $n - 1$ latency challenges the view that temporal learning has an important role in producing PC effects. Together, these results clearly demonstrate that it is not the case that PC effects are unobservable when those factors are controlled for (Schmidt, 2013b; Schmidt & Besner, 2008).

It must be noted that although the color–word and picture–word interference tasks are thought to reflect the same underlying processes (see note 1), one important difference between the typical color–word Stroop task used in the literature and the picture–word interference task used here is that the former, but not the latter, elicits response interference. That is, in most implementations of the color–word Stroop task, incongruent words are also used as responses (e.g., the word YELLOW is presented in an experiment in which yellow color targets are also used), whereas incongruent word distractors were not responses in either Experiment 1A or Experiment 1B (e.g., BED appeared as a word distractor for the picture of a dog but not as a target picture). Using distractors that are not used as responses is known to reduce interference from the irrelevant dimension in both color–word and picture–word interference tasks (Lupker & Katz, 1981; Proctor, 1978), suggesting that response interference, among other factors, contributes to Stroop and Stroop-like effects (La Heij, 1988). As such, what the results of Experiments 1A and 1B provide is evidence for adaptation to conflict frequency even in a situation in which conflict was likely less intense than in a typical Stroop task because response interference was playing little, if any, role.

It is also important to acknowledge, however, that recent findings, published as the present research was in progress, suggest that learning of contingencies might occur at a more abstract level than previously thought. Schmidt et al. (2018) reported two color identification experiments in which words belonging to three different semantic categories were used as distractors, each category being predictive of one color. Similar to the experiments reported here, each individual word was presented only once, thus eliminating individual word–response contingencies. A category-based contingency effect was observed, with faster and more accurate responses when a category item was presented in the color in which most of the other items of that category were presented. Note that although the present experiments were designed to eliminate individual word–response contingencies, they allowed for category-based contingency learning. For example, words denoting animals were mostly associated with pictures of animals in MC lists, whereas they were equally associated with each of the four semantic categories in MI lists. Thus, participants in MC lists potentially could have used the category of the word distractor to predict the response, leading to a speed-up on high-contingency congruent items and therefore, an inflated congruency effect in MC lists.

An account of this sort, however, seems to be unlikely for a couple of reasons: First, the effects reported by Schmidt et al. (2018) (11 ms and a nonsignificant 2 ms in their Experiments 1 and 2, respectively) seem too small to offer a convincing alternative interpretation of the present findings (note that classic word–response contingency-learning effects are on the order of 40–60

ms: e.g., Lin & MacLeod, 2018; Schmidt et al., 2007). Second, although the possibility of using the category of the word distractor to predict the response might be tenable for Experiment 1A, where the response was a category name itself, applying this idea to Experiment 1B would imply that a rather complicated mechanism was in place: Participants in MC lists would have had to have used the congruent word distractor to predict the category of the picture, which would then have helped them retrieve the name of the picture (i.e., a name > category > name route). However, because congruent word distractors are the name of the picture, it is unclear why following this name > category > name route would be of any benefit for performance. Finally, it has long been established that pictures are categorized faster than words are (e.g., Lupker & Katz, 1982; Smith & Magee, 1980), and as such, using the category of word distractors to predict the category of the target pictures would be somewhat counterproductive in a speeded task. As such, adaptation to conflict frequency seems, at present, a much better explanation for the PC effects obtained here.

The Present Results From the Perspective of Bugg's (2014a) AATC Hypothesis

When considering the implications of the conclusion that a conflict–adaptation strategy is likely responsible for the results we obtained, one thing that is potentially important to note is that unlike classic PC manipulations in the Stroop task, Experiments 1A and 1B presented participants with a situation where learning of word–response contingencies was not an option at all, as the identity of word distractors could not be used to predict the response. Hence, conflict adaptation may have been essentially the only strategy available for dealing with conflict. Such is not the case, however, when engaging in routine activities in everyday life (e.g., driving to one's workplace). Those situations typically involve attending to a task in the face of stimuli that, reoccurring in time, become predictive of certain events (e.g., the fuel light on the car's dashboard signaling it is time to refuel). It is thus critical to understand how control over action is implemented in situations where a contingency-learning option is available.

In response to this concern, Bugg (2014a) proposed the associations as antagonists to top-down control (AATC) hypothesis to explain how the employment of contingency-learning and conflict-adaptation mechanisms is regulated. According to this hypothesis, the availability of reliable stimulus–response associations moderates the engagement of top-down mechanisms of conflict adaptation. Specifically, no adaptation to conflict frequency would take place if contingencies can be used to guide responding most of the time. Conflict adaptation would be, in other words, a last resort used by the control system only when learning contingencies—the default mode driving control engagement—is not feasible.

To provide some support for this hypothesis, Bugg (2014a) divided color–word Stroop stimuli into two sets, a “context” set and a “transfer” set, and manipulated conflict frequency and contingency learning for the context set only (transfer words were contingency-unbiased, appearing with congruent and incongruent colors an equal number of times). The transfer items were intermixed in the same list with context items which were either mostly congruent or mostly incongruent, so that transfer stimuli appeared in either a mostly congruent list (when mixed with MC context items) or in a mostly incongruent list (when mixed with MI context

items). Crucially, in one version of Bugg's experiments, both MC and MI context items allowed learning of contingencies, making contingency learning a very effective strategy. In contrast, in another version of Bugg's experiments, only MC context items allowed learning of contingencies, as MI context items were constructed such that there were no contingencies to learn (i.e., each of four words was presented equally often in each of four colors). Thus, contingency learning was not a useful strategy in the MI condition in this version of her experiments. Consistent with the AATC hypothesis, Bugg obtained PC effects for transfer items (i.e., evidence for a conflict-adaptation strategy being applied when responding to those items) only in the MI condition in the second version of her experiments, that is, when learning of contingencies was impossible. When learning of contingencies was possible for context items in both the MC and the MI list (i.e., when such a strategy was useful) there was no evidence of adoption of a conflict-adaptation strategy for the transfer items (for a counterargument, see Schmidt, 2014b).

Extending the AATC hypothesis to the paradigm used here, the implication for the results of Experiments 1A and 1B is straightforward: Adaptation to conflict frequency occurred not *in spite of* contingency learning being impossible, but *because* it was impossible. That is, contingency-controlled PC effects (i.e., clear markers of conflict adaptation) were obtained because there were no contingency biases at all, and adaptation to conflict frequency was the only remaining good option for maximizing performance in the task.

Implications of the Present Research for Temporal Learning Accounts

In addition to examining adaptation to conflict frequency, the present research also sheds some light on the mechanism of temporal learning. This general form of learning assumes that participants in speeded tasks form temporal expectancies for emission of a response and, most critically, in Schmidt's (2013c) conceptualization, they adjust to those expectancies by speeding up on the trials in which they can produce a latency that matches the established expectancy. Although some evidence exists in favor of this mechanism (Kinoshita et al., 2011; Schmidt, 2013c, 2014a, 2016; Schmidt & Weissman, 2016), there is virtually no support for it in the present data. That is, the use of generalized linear mixed-effects models, a statistical technique that requires no transformation of the dependent variable (Lo & Andrews, 2015), failed to produce the predicted reduction of congruency effects with increasing RT on trial $n - 1$ in Experiments 1A and 1B. These results are consistent with those of Cohen-Shikora et al. (2018), who failed to obtain regular temporal learning effects using untransformed RTs in generalized linear mixed-effects models for a number of data sets, including Hutchison's (2011), the dataset which Schmidt first reanalyzed (with transformed RTs as the dependent variable) to make a case for temporal learning.

Because temporal learning is indexed by an interaction (i.e., that between RT on trial $n - 1$ and congruency on trial n), the present results and Cohen-Shikora et al.'s results raise the suspicion that temporal-learning interactions reported in previously published papers (Huber-Huber & Ansoorge, 2017, 2018; Kinoshita et al., 2011; Schmidt, 2013c; Schmidt & Weissman, 2016) were created by the use of nonlinear transformations of the dependent variable,

an operation that is routinely performed in linear mixed-effects modeling. It is important to again note that, although these transformations do a decent job of accommodating the assumption made by linear mixed-effects models that the dependent variable be normally distributed, they affect the size and the pattern of interactions (Balota et al., 2013). Generalized linear mixed-effects models, requiring no RT transformation, provide researchers with a safer technique to search for interactions, a technique that, moving forward, is well worth considering when interactions represent the main research interest (e.g., Yang, Chen, Spinelli, & Lupker, 2018).

Another example of the present data failing to support Schmidt's (2013c) version of a temporal learning account can be found in the results of Experiment 2. In that experiment, congruent and incongruent items were replaced with easy and hard items, items not requiring the filtering out of irrelevant information as is required by interference stimuli. The results suggested that Schmidt's version of temporal learning was not at work in this situation (i.e., when vocal responding to multiple pictures is required). That is, unlike similar investigations in a button-press letter identification task utilizing low-contrast (i.e., hard) and high-contrast (i.e., easy) letters as stimuli (Schmidt, 2013c, 2014a, 2016), the proportion of easy stimuli in the list did not influence the size of the difficulty effect. As the main point of Experiment 2 was to investigate the potential contribution of temporal learning to the PC effects found in Experiments 1A and 1B, the obvious question Experiment 2's results raise is whether it is possible to reconcile them with Schmidt's (2013c, 2014a, 2016) findings that manipulating frequency of difficulty does alter the magnitude of difficulty effects.

One important difference between Experiment 2 and Schmidt's (2013c, 2014a, 2016) experiments is the nature of the identification that is required (naming of multiple pictures vs. button-press identification of a limited set of letters). As mentioned above, button-press lexical decision and word naming tend to show different patterns of blocking effects, with naming showing equivalent benefits for both easy and hard items in a block containing mainly easy items, whereas button-press lexical decision typically produces an asymmetric pattern, with large benefits for easy items but not for hard items in a block containing mainly easy items (Kinoshita & Mozer, 2006). Extending this idea to proportion-easy manipulations, it is easy to see how a vocal-responding situation where easy and hard items are influenced by the frequency of difficulty in the same way will result in no proportion-easy effect, whereas a manual-responding situation where easy items are influenced by frequency of difficulty, but hard items much less so, will likely result in a proportion-easy effect.

Note that manual and vocal identification do differ in various ways. For example, manual responding generally constrains the number of responses available, whereas vocal responding, as in the present experiments, allows for multiple responses. Furthermore, a button press response requires participants to make a forced choice and commit to it, whereas a vocal response involves a gradual accumulation of evidence (e.g., Perea & Carreiras, 2003). As a result, participants might develop different subjective error estimates in the two situations. That is, their confidence in being able to give the correct response with sufficient time might not be the same.

Indeed, response confidence was the very factor that Kinoshita and Mozer (2006) held responsible for the different patterns of blocking effects observed in word naming and lexical decision tasks. In those tasks, high-frequency and low-frequency words were used as easy and hard items, respectively. Importantly, participants in a word-naming task can be assumed to be relatively confident about their response, even for hard items, but such may not be the case for the same hard items in lexical decision, for which a certain degree of uncertainty might remain even when a response is made (i.e., you know you will eventually name “glabrous” acceptably, but do you know for sure you will correctly classify it as a word or a nonword?).

The story changes, however, if the low-frequency words that are used, despite being harder than the high-frequency words, are familiar enough for participants to confidently classify them as words. Using these kinds of low-frequency words, Kinoshita and Mozer (2006) obtained the pattern usually found in naming, equivalent effects for low- and high-frequency words. Kinoshita and Mozer explained these findings in term of their ASE model. Simply put, the ASE model predicts that when an item is so hard that participants may never (i.e., even if they had no time pressure) be completely confident about their response, participants will not wait extra time in pure hard, compared to mixed, blocks, as doing so will not significantly improve accuracy. As a result, they will respond before they are entirely confident in pure hard blocks and no mixing benefit will be observed for those stimuli. When, however, hard items can still be responded to confidently given more time, it will be worth it to wait the extra time to confidently produce an accurate response, which will result in longer latencies in hard blocks and, hence, a mixing benefit.

One could certainly argue there might be parallels between the two situations examined by Kinoshita and Mozer (2006) and the two situations created by the present Experiment 2 versus Schmidt’s (2013c, 2014a, 2016) experiments, parallels which might explain the difference between the data patterns in the latter two situations. Although it seems unlikely that participants in Experiment 2 were completely confident about their responses to all low-resolution pictures, it is important to note that participants were presented with stimuli which often had multiple acceptable responses (in fact, several of the responses marked as errors with the conservative criterion adopted here were actually fairly acceptable responses, e.g., “tool” instead of “screwdriver,” “swimming” instead of “swimmer,” etc.). In addition, because participants were not given feedback, as is typical in naming tasks, they were never informed that they were making “errors” in many situations. In turn, this inability to know when errors were being made might have led them to assume that their responses were likely acceptable and to conclude that given enough time, they would confidently respond to both easy items (i.e., high-resolution pictures) and hard items (i.e., low-resolution pictures). Therefore, the situation in the present Experiment 2 would be much more like that in a standard naming task, implying that one would expect a speed-up for both easy and hard items in the easy block.

In contrast, participants in Schmidt’s experiments were regularly given feedback, and were presented with stimuli which had only one acceptable response among a limited set of responses. Thus, participants in Schmidt’s experiments had a better idea about how well (or badly) they were performing. Therefore, it is possible that those participants were, in some cases, constantly unsure

about the accuracy of their responses to hard items (i.e., low-contrast letters). In turn, this situation could have reduced the impact of frequency of difficulty selectively for hard items, as predicted by the ASE model, thus producing the pattern of blocking effects often found in lexical decision tasks, that is, the differences in the magnitude of the difficulty effects in ME and MH lists that he observed. An examination of the role of response modality, size of the response set, and feedback in the high/low contrast letter identification paradigm would likely help shed light on the reason why the present results and Schmidt’s differ so remarkably.⁵

Conclusion

To conclude, the reported data make a good case for the existence of a conflict-adaptation mechanism in humans. Far from being a mere illusion, such a mechanism might be an important resource in coping with tasks that require some degree of distraction suppression. Although learning about *what* to respond (contingency learning) and *when* to do it (temporal learning) might be crucial aspects in goal-oriented behavior, learning *how* to respond (i.e., learning the appropriate attentional strategy to achieve the desired goal) is another human ability that should be acknowledged.

⁵ It is important to appreciate the fact that the present discussion rests on the assumption that a temporal-learning mechanism is responsible for the pattern reported by Schmidt (2013c, 2014a, 2016) in the letter identification task. However, as recognized by Schmidt (2013c), this assumption may not be correct: If low-contrast letters are thought as stimuli creating a relatively high level of perceptual conflict, a mechanism of adaptation to the frequency of perceptual conflict could also explain his data (e.g., if participants squint their eyes more in the list containing mostly low-contrast letters, the contrast effect will be reduced). At the same time, the results from Experiment 2 constrain this putative conflict-adaptation mechanism in that they suggest that not all forms of stimulus degradation (e.g., the resolution of an image) engender a kind of perceptual conflict that people can adapt to.

References

- Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. *Psychological Bulletin, 142*, 693–728. <http://dx.doi.org/10.1037/bul0000047>
- Abrahamse, E., Ruitenberg, M., Duthoo, W., Sabbe, B., Morrens, M., & van Dijk, J. P. (2016). Conflict adaptation in schizophrenia: Reviewing past and previewing future efforts. *Cognitive Neuropsychiatry, 21*, 197–212. <http://dx.doi.org/10.1080/13546805.2016.1167679>
- Baayen, R. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511801686>
- Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412. <http://dx.doi.org/10.1016/j.jml.2007.12.005>
- Baayen, R., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX Lexical Database [CD-ROM]*. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.
- 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. <http://dx.doi.org/10.1037/a0032186>

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. <http://dx.doi.org/10.18637/jss.v067.i01>
- Beckers, T., De Houwer, J., & Matute, H. (2007). Human contingency learning. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *60*, 289–290. <http://dx.doi.org/10.1080/17470210601000532>
- Blais, C., & Bunge, S. (2010). Behavioral and neural evidence for item-specific performance monitoring. *Journal of Cognitive Neuroscience*, *22*, 2758–2767. <http://dx.doi.org/10.1162/jocn.2009.21365>
- Bolker, B. (2018). *GLMM FAQ*. Retrieved from <https://bbolker.github.io/mixedmodels-misc/glmmFAQ.html>
- Bonnin, C. A., Houeto, J. L., Gil, R., & Bouquet, C. A. (2010). Adjustments of conflict monitoring in Parkinson's disease. *Neuropsychology*, *24*, 542–546. <http://dx.doi.org/10.1037/a0018384>
- 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. <http://dx.doi.org/10.1037/0033-295X.108.3.624>
- Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*, 106–113. <http://dx.doi.org/10.1016/j.tics.2011.12.010>
- Bugg, J. M. (2014a). Conflict-triggered top-down control: Default mode, last resort, or no such thing? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 567–587. <http://dx.doi.org/10.1037/a0035032>
- Bugg, J. M. (2014b). Evidence for the sparing of reactive cognitive control with age. *Psychology and Aging*, *29*, 115–127. <http://dx.doi.org/10.1037/a0035270>
- Bugg, J. M., & Chanani, S. (2011). List-wide control is not entirely elusive: Evidence from picture-word Stroop. *Psychonomic Bulletin & Review*, *18*, 930–936. <http://dx.doi.org/10.3758/s13423-011-0112-y>
- Bugg, J. M., & Crump, M. J. (2012). In support of a distinction between voluntary and stimulus-driven control: A review of the literature on proportion congruent effects. *Frontiers in Psychology*, *3*, 367. <http://dx.doi.org/10.3389/fpsyg.2012.00367>
- Bugg, J. M., & Hutchison, K. A. (2013). Converging evidence for control of color-word Stroop interference at the item level. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 433–449. <http://dx.doi.org/10.1037/a0029145>
- Bugg, J. M., Jacoby, L. L., & Chanani, S. (2011). Why it is too early to lose control in accounts of item-specific proportion congruency effects. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 844–859. <http://dx.doi.org/10.1037/a0019957>
- Bugg, J. M., Jacoby, L. L., & Toth, J. P. (2008). Multiple levels of control in the Stroop task. *Memory & Cognition*, *36*, 1484–1494. <http://dx.doi.org/10.3758/mc.36.8.1484>
- Chateau, D., & Lupker, S. J. (2003). Strategic effects in word naming: Examining the route-emphasis versus time-criterion accounts. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 139–151. <http://dx.doi.org/10.1037/0096-1523.29.1.139>
- 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*. Advance online publication. <http://dx.doi.org/10.1037/xlm0000670>
- Crump, M. J., Gong, Z., & Milliken, B. (2006). The context-specific proportion congruent Stroop effect: Location as a contextual cue. *Psychonomic Bulletin & Review*, *13*, 316–321. <http://dx.doi.org/10.3758/BF03193850>
- Dell'Acqua, R., Job, R., Peressotti, F., & Pascali, A. (2007). The picture-word interference effect is not a Stroop effect. *Psychonomic Bulletin & Review*, *14*, 717–722. <http://dx.doi.org/10.3758/BF03196827>
- Egner, T. (2014). Creatures of habit (and control): A multi-level learning perspective on the modulation of congruency effects. *Frontiers in Psychology*, *5*, 1247. <http://dx.doi.org/10.3389/fpsyg.2014.01247>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. <http://dx.doi.org/10.3758/BRM.41.4.1149>
- Forrin, N. D., & MacLeod, C. M. (2017). Relative speed of processing determines color-word contingency learning. *Memory & Cognition*, *45*, 1206–1222. <http://dx.doi.org/10.3758/s13421-017-0721-4>
- Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116–124. <http://dx.doi.org/10.3758/BF03195503>
- Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive control in the Stroop task. *Memory & Cognition*, *44*, 778–788. <http://dx.doi.org/10.3758/s13421-016-0591-1>
- Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography & Clinical Neurophysiology*, *55*, 468–484. [http://dx.doi.org/10.1016/0013-4694\(83\)90135-9](http://dx.doi.org/10.1016/0013-4694(83)90135-9)
- 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. <http://dx.doi.org/10.1007/s00426-003-0132-y>
- Huber-Huber, C., & Ansorge, U. (2017). The role of RT carry-over for congruence sequence effects in masked priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 757–780. <http://dx.doi.org/10.1037/xlm0000331>
- Huber-Huber, C., & Ansorge, U. (2018). Unconscious conflict adaptation without feature-repetitions and response time carry-over. *Journal of Experimental Psychology: Human Perception and Performance*, *44*, 169–175. <http://dx.doi.org/10.1037/xhp0000450>
- Hutchison, K. A. (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 851–860. <http://dx.doi.org/10.1037/a0023437>
- Jacoby, L. L., Lindsay, D. S., & Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin & Review*, *10*, 638–644. <http://dx.doi.org/10.3758/BF03196526>
- Kiger, J. I., & Glass, A. L. (1981). Context effects in sentence verification. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 688–700. <http://dx.doi.org/10.1037/0096-1523.7.3.688>
- King, J. A., Korb, F. M., & Egner, T. (2012). Priming of control: Implicit contextual cuing of top-down attentional set. *The Journal of Neuroscience*, *32*, 8192–8200. <http://dx.doi.org/10.1523/JNEUROSCI.0934-12.2012>
- Kinoshita, S., & Mozer, M. C. (2006). How lexical decision is affected by recent experience: Symmetric versus asymmetric frequency-blocking effects. *Memory & Cognition*, *34*, 726–742. <http://dx.doi.org/10.3758/BF03193591>
- 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. <http://dx.doi.org/10.1037/a0024230>
- La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Memory & Cognition*, *16*, 400–410. <http://dx.doi.org/10.3758/BF03214220>
- 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. <http://dx.doi.org/10.1016/j.actpsy.2015.12.009>
- Levy, R. (2014). *Using R formulae to test for main effects in the presence of higher order interactions*. Retrieved from <https://arxiv.org/abs/1405.2094>
- Lin, O. Y. H., & MacLeod, C. M. (2018). The acquisition of simple associations as observed in color-word contingency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 99–106. <http://dx.doi.org/10.1037/xlm0000436>

- 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, 1171. <http://dx.doi.org/10.3389/fpsyg.2015.01171>
- 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. <http://dx.doi.org/10.3758/BF03197535>
- Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition*, 7, 485–495. <http://dx.doi.org/10.3758/BF03198265>
- Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 570–590. <http://dx.doi.org/10.1037/0278-7393.23.3.570>
- Lupker, S. J., & Katz, A. N. (1981). Input, decision, and response factors in picture-word interference. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 269–282. <http://dx.doi.org/10.1037/0278-7393.7.4.269>
- Lupker, S. J., & Katz, A. N. (1982). Can automatic picture processing influence word judgments? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 418–434. <http://dx.doi.org/10.1037/0278-7393.8.5.418>
- Lupker, S. J., Kinoshita, S., Coltheart, M., & Taylor, T. E. (2003). Mixing costs and mixing benefits in naming words, pictures, and sums. *Journal of Memory and Language*, 49, 556–575. [http://dx.doi.org/10.1016/S0749-596X\(03\)00094-9](http://dx.doi.org/10.1016/S0749-596X(03)00094-9)
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163–203. <http://dx.doi.org/10.1037/0033-2909.109.2.163>
- Mayr, U., Awh, E., & Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature Neuroscience*, 6, 450–452. <http://dx.doi.org/10.1038/nn1051>
- Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, 110, 422–471. <http://dx.doi.org/10.1037/0033-295X.110.3.422>
- Musen, G., & Squire, L. R. (1993). Implicit learning of color-word associations using a Stroop paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 789–798. <http://dx.doi.org/10.1037/0278-7393.19.4.789>
- Perea, M., & Carreiras, M. (2003). Sequential effects in the lexical decision task: The role of the item frequency of the previous trial. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 56, 385–401. <http://dx.doi.org/10.1080/02724980244000387>
- Proctor, R. W. (1978). Sources of color-word interference in the Stroop color-naming task. *Perception & Psychophysics*, 23, 413–419. <http://dx.doi.org/10.3758/BF03204145>
- Protopapas, A. (2007). CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, 39, 859–862. <http://dx.doi.org/10.3758/BF03192979>
- Rastle, K., Kinoshita, S., Lupker, S. J., & Coltheart, M. (2003). Cross-task strategic effects. *Memory & Cognition*, 31, 867–876. <http://dx.doi.org/10.3758/BF03196441>
- R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>
- Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, 1, 103–113. <http://dx.doi.org/10.1111/j.2041-210X.2010.00012.x>
- Schmidt, J. R. (2013a). The Parallel Episodic Processing (PEP) model: Dissociating contingency and conflict adaptation in the item-specific proportion congruent paradigm. *Acta Psychologica*, 142, 119–126. <http://dx.doi.org/10.1016/j.actpsy.2012.11.004>
- Schmidt, J. R. (2013b). Questioning conflict adaptation: Proportion congruent and Gratton effects reconsidered. *Psychonomic Bulletin & Review*, 20, 615–630. <http://dx.doi.org/10.3758/s13423-012-0373-0>
- Schmidt, J. R. (2013c). Temporal learning and list-level proportion congruency: Conflict adaptation or learning when to respond? *PLoS ONE*, 8, e82320. <http://dx.doi.org/10.1371/journal.pone.0082320>
- Schmidt, J. R. (2014a). List-level transfer effects in temporal learning: Further complications for the list-level proportion congruent effect. *Journal of Cognitive Psychology*, 26, 373–385. <http://dx.doi.org/10.1080/20445911.2014.896367>
- Schmidt, J. R. (2014b). Contingencies and attentional capture: The importance of matching stimulus informativeness in the item-specific proportion congruent task. *Frontiers in Psychology*, 5, 540. <http://dx.doi.org/10.3389/fpsyg.2014.00540>
- Schmidt, J. R. (2016). Temporal learning and rhythmic responding: No reduction in the proportion easy effect with variable response-stimulus intervals. *Frontiers in Psychology*, 7, 634. <http://dx.doi.org/10.3389/fpsyg.2016.00634>
- Schmidt, J. R., Augustinova, M., & De Houwer, J. (2018). Category learning in the color-word contingency learning paradigm. *Psychonomic Bulletin & Review*, 25, 658–666. <http://dx.doi.org/10.3758/s13423-018-1430-0>
- 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. <http://dx.doi.org/10.1037/0278-7393.34.3.514>
- Schmidt, J. R., Crump, M. J., Cheesman, J., & Besner, D. (2007). Contingency learning without awareness: Evidence for implicit control. *Consciousness and Cognition*, 16, 421–435. <http://dx.doi.org/10.1016/j.concog.2006.06.010>
- 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. <http://dx.doi.org/10.1016/j.cogpsych.2016.10.004>
- Schmidt, J. R., Notebaert, W., & Bussche, E. V. (2015). Is conflict adaptation an illusion? *Frontiers in Psychology*, 6, 172. <http://dx.doi.org/10.3389/fpsyg.2015.00172>
- Schmidt, J. R., & Weissman, D. H. (2014). Congruency sequence effects without feature integration or contingency learning confounds. *PLoS ONE*, 9, e102337. <http://dx.doi.org/10.1371/journal.pone.0102337>
- Schmidt, J. R., & Weissman, D. H. (2016). Congruency sequence effects and previous response times: Conflict adaptation or temporal learning? *Psychological Research*, 80, 590–607. <http://dx.doi.org/10.1007/s00426-015-0681-x>
- Schneider, D. W. (2015). Attentional control of response selection in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 41, 1315–1324. <http://dx.doi.org/10.1037/xhp0000091>
- Schnur, T. T., & Martin, R. (2012). Semantic picture-word interference is a postperceptual effect. *Psychonomic Bulletin & Review*, 19, 301–308. <http://dx.doi.org/10.3758/s13423-011-0190-x>
- Sheth, S. A., Mian, M. K., Patel, S. R., Asaad, W. F., Williams, Z. M., Dougherty, D. D., . . . Eskandar, E. N. (2012). Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature*, 488, 218–221. <http://dx.doi.org/10.1038/nature11239>
- Singmann, H., & Kellen, D. (in press). An introduction to mixed models for experimental psychology. In D. Spieler & E. Schumacher (Eds.), *New methods in cognitive psychology*. New York, NY: Routledge.
- Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture—Word processing. *Journal of Experimental Psychology: General*, 109, 373–392. <http://dx.doi.org/10.1037/0096-3445.109.4.373>
- Starreveld, P. A., & La Heij, W. (2017). Picture-word interference is a Stroop effect: A theoretical analysis and new empirical findings. *Psy-*

- chonomic Bulletin & Review*, 24, 721–733. <http://dx.doi.org/10.3758/s13423-016-1167-6>
- Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662. <http://dx.doi.org/10.1037/h0054651>
- Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., . . . Bates, E. (2005). Timed action and object naming. *Cortex*, 41, 7–25. [http://dx.doi.org/10.1016/S0010-9452\(08\)70174-6](http://dx.doi.org/10.1016/S0010-9452(08)70174-6)
- Taylor, T. E., & Lupker, S. J. (2001). Sequential effects in naming: A time-criterion account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 117–138. <http://dx.doi.org/10.1037/0278-7393.27.1.117>
- Taylor, T. E., & Lupker, S. J. (2006). Time perception and word recognition: An elaboration of the time-criterion account. *Perception & Psychophysics*, 68, 933–945. <http://dx.doi.org/10.3758/BF03193356>
- Taylor, T. E., & Lupker, S. J. (2007). Sequential effects in time perception. *Psychonomic Bulletin & Review*, 14, 70–74. <http://dx.doi.org/10.3758/BF03194030>
- Underwood, G. (1976). Semantic interference from unattended printed words. *British Journal of Psychology*, 67, 327–338. <http://dx.doi.org/10.1111/j.2044-8295.1976.tb01519.x>
- van Maanen, L., van Rijn, H., & Borst, J. P. (2009). Stroop and picture-word interference are two sides of the same coin. *Psychonomic Bulletin & Review*, 16, 987–999. <http://dx.doi.org/10.3758/PBR.16.6.987>
- Weissman, D. H., Jiang, J., & Egner, T. (2014). Determinants of congruency sequence effects without learning and memory confounds. *Journal of Experimental Psychology: Human Perception and Performance*, 40, 2022–2037. <http://dx.doi.org/10.1037/a0037454>
- West, R., & Alain, C. (2000). Age-related decline in inhibitory control contributes to the increased Stroop effect observed in older adults. *Psychophysiology*, 37, 179–189. <http://dx.doi.org/10.1111/1469-8986.3720179>
- Wilk, H. A., Ezekiel, F., & Morton, J. B. (2012). Brain regions associated with moment-to-moment adjustments in control and stable task-set maintenance. *NeuroImage*, 59, 1960–1967. <http://dx.doi.org/10.1016/j.neuroimage.2011.09.011>
- Yang, H., Chen, J., Spinelli, G., & Lupker, S. J. (2018). The impact of text orientation on form priming effects in four-character Chinese words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. <http://dx.doi.org/10.1037/xlm0000655>

Appendix

Examining the Impact of Temporal Learning on the Congruency Sequence Effect

The congruency sequence effect refers to the finding that, in interference tasks, congruency effects are larger following a congruent trial than following an incongruent trial (Gratton et al., 1983). The traditional, control-based account of this effect (Botvinick et al., 2001) holds that experiencing conflict during an incongruent trial would lead participants to focus attention to the target dimension, thus reducing interference on subsequent trials; conversely, experiencing little or no conflict during a congruent trial would lead to relaxed attention, thus increasing interference on subsequent trials. Like the control account of the PC effect, this explanation has also faced some challenges: For example, in most paradigms, repetitions of stimulus features from one trial to the next seem to contribute to the congruency sequence effect (e.g., Hommel et al., 2004; Mayr et al., 2003), although a congruency sequence effect is still observed when this confound and others are removed (e.g., Schmidt & Weissman, 2014; Weissman, Jiang, & Egner, 2014).

Recently, however, Schmidt and Weissman (2016) proposed that the congruency sequence effect observed when potential confounds are accounted for is best interpreted as being the result of a temporal learning mechanism rather than the result of a conflict-adaptation mechanism. This temporal learning explanation is similar to the one proposed for PC effects. Following a trial in which a fast response was emitted (typically, a congruent trial), participants will develop a relatively fast temporal expectancy which will speed up responding to a subsequent item that could be processed

rapidly enough to meet that fast temporal expectancy (a situation typically occurring on a congruent trial). Because this speed-up will typically benefit congruent items but not incongruent items, the result will be an inflated congruency effect following a congruent trial, consistent with the pattern of the congruency sequence effect. Conversely, following a trial in which a slow response was emitted (typically, an incongruent trial), participants will develop a relatively slow temporal expectancy. This temporal expectancy could potentially speed up responding to a subsequent slow item that could be processed fast enough to meet that slower temporal expectancy (a situation typically occurring on an incongruent trial), although this result may not be observed in practice because, as noted, temporal expectancies often have little impact on hard-to-process stimuli (Kinoshita & Mozer, 2006; Kinoshita et al., 2011). In any case, the point is that following an incongruent stimulus, there is a potential speed-up for incongruent items. However, in comparison with what happens when the preceding response was fast, there would be no pressure to produce faster responses for fast (i.e., congruent) items. The result would be a congruency effect which should be, if anything, relatively small—again, consistent with the pattern of the congruency sequence effect. In sum, a temporal learning mechanism of this sort would seem capable of creating a pattern of results that mimics the congruency sequence effect with no need to assume a conflict adaptation mechanism.

(Appendix continues)

To support their temporal learning interpretation of the congruency sequence effect, Schmidt and Weissman (2016) reanalyzed Schmidt and Weissman's (2014) data, a confound-minimized study of the congruency sequence effect in the prime-probe task, using RT on trial $n - 1$ as an index of temporal expectancy for trial n in a linear mixed-effects model analysis. They reasoned that the finding of an interaction between RT on trial $n - 1$ and congruency on trial n whereby congruency effects diminish with higher RT on trial $n - 1$ would be evidence that a temporal learning mechanism is being used. Indeed, they obtained not only such an interaction but also a reduction (although not an elimination) of the value of the beta parameter for the congruency sequence effect (i.e., the interaction between congruency and congruency on trial $n - 1$) in the model. Using similar reasoning as that used by Schmidt (2013c) for the PC effect, Schmidt and Weissman (2016) interpreted these results as indicating that temporal learning can generate a congruency sequence effect on its own, an idea they reinforced by successfully simulating the experimental data with an upgraded version of Schmidt's (2013c) PEP model in which temporal learning was an implemented mechanism but trial-to-trial conflict adaptation was not.

However, a fundamental problem with Schmidt and Weissman's (2016) results is that, similar to what done by Schmidt (2013c) in the context of the PC effect, RTs were inverse-transformed to accommodate the assumption of a normally distributed dependent variable made by linear mixed-effects models. As noted, such a transformation can substantially alter the pattern of interactions and thus casts serious doubts on the interpretation of interactions, including the critical interaction between RT on trial $n - 1$ and congruency, the interaction that indexes temporal learning. In the following, we present additional analyses of Experiments 1A and 1B to examine whether the problems that emerged for the temporal learning explanation of the PC effect when a more appropriate analysis is used (i.e., generalized linear mixed-effects models with untransformed RTs; Cohen-Shikora et al., 2018) also emerge when considering the congruency sequence effect.

Results

The analyses were based on the same data as those used for the analyses reported in the main text of the article, with the exception that trials for which an error was made on the preceding trial were removed from both the latency and the error analyses, as is standard for analyses of congruency sequence effects. Furthermore, in order to minimize the impact of feature and response

repetitions (Hommel et al., 2004), for Experiment 1A (picture categorization) we removed the trials in which the category of the picture (and hence, the correct response) on trial n matched the category of the picture (and correct response) on trial $n - 1$ (e.g., the picture of dog preceded by the picture of a cat, with both pictures requiring the response ANIMAL). The statistical models were also the same as those used for the analyses reported in the main text of the article, with the exception that Congruency on trial $n - 1$ was included as an additional fixed effect. The mean RTs and error rates for by-subject data for these analyses of Experiments 1A and 1B are presented in Tables A1 and A2, respectively.

Experiment 1A (picture categorization).

Response time. There was a main effect of RT on trial $n - 1$ (faster responses with lower RT on trial $n - 1$), $\beta = 24.01$, $SE = 3.42$, $z = 7.01$, $p < .001$, but not Congruency, $\beta = -3.86$, $SE = 3.31$, $z = -1.16$, $p = .25$. The interaction between Congruency and List Type, that is, the PC effect, was significant, $\beta = -8.73$, $SE = 3.49$, $z = -2.51$, $p = .012$. Congruency on trial $n - 1$ marginally interacted with List Type, $\beta = 6.75$, $SE = 3.52$, $z = 1.92$, $p = .055$, indicating that in MC lists responses tended to be overall faster when trial $n - 1$ was incongruent, a pattern that was reversed in MI lists. Most importantly, Congruency on trial $n - 1$ and Congruency did not interact, $\beta = -2.63$, $SE = 3.44$, $z = -.76$, $p = .47$, although there was a marginal three-way interaction between Congruency on trial $n - 1$, Congruency, and List Type, $\beta = -5.52$, $SE = 3.30$, $z = -1.68$, $p = .094$. As in the analysis presented in the main text of the article, there was also a three-way interaction between Congruency, List Type, and RT on trial $n - 1$, $\beta = 7.29$, $SE = 3.29$, $z = 2.21$, $p = .027$.

Table A1
Mean RTs and Percentage Error Rates (and Corresponding Standard Errors) for the Congruency-Sequence-Effect Analysis of Experiment 1A

Congruency	RTs		Error rates	
	MC list	MI list	MC list	MI list
Previous congruent				
Congruent	920 (.25)	931 (.36)	1.3 (.3)	.8 (.8)
Incongruent	970 (.25)	913 (.37)	2.7 (.7)	2 (.8)
Congruency effect	50	-18	1.4	1.2
Previous incongruent				
Congruent	923 (.24)	935 (.32)	3.9 (.8)	2.7 (.7)
Incongruent	928 (.32)	938 (.39)	4.3 (1.7)	2.2 (.5)
Congruency effect	5	3	.4	-.5

(Appendix continues)

Table A2
Mean RTs and Percentage Error Rates (and Corresponding Standard Errors) for the Congruency-Sequence-Effect Analysis of Experiment 1B

Congruency	RTs		Error rates	
	MC list	MI list	MC list	MI list
Previous congruent				
Congruent	762 (27)	791 (26)	1 (.3)	2.6 (2.1)
Incongruent	1111 (40)	1054 (26)	15.3 (2.1)	11.4 (1.5)
Congruency effect	349	263	15.2	8.8
Previous incongruent				
Congruent	770 (24)	802 (22)	1.3 (.5)	2 (.8)
Incongruent	1083 (42)	1037 (23)	8.5 (2)	11 (1.1)
Congruency effect	313	235	7.2	9

The three-way interactions were explored by analyzing MC and MI lists separately. In MC lists, the main effects of Congruency on trial $n - 1$ (faster responses when trial $n - 1$ was incongruent), $\beta = 10.46$, $SE = 5.08$, $z = 2.06$, $p = .040$, Congruency (congruent faster than incongruent), $\beta = -11.96$, $SE = 5.03$, $z = -2.38$, $p = .017$, and RT on trial $n - 1$, $\beta = 24.86$, $SE = 5.54$, $z = 4.49$, $p < .001$, were all significant. In addition, there was a marginal interaction between Congruency on trial $n - 1$ and Congruency, $\beta = -9.36$, $SE = 5.19$, $z = -1.80$, $p = .071$. This interaction indicates a regular congruency sequence effect, with a tendency for the congruency effect to be reduced following an incongruent trial (5 ms) compared with the congruency effect following a congruent trial (50 ms). Note that this reduction occurred because responses to incongruent trials were faster when following another incongruent trial (928 ms) than when following a congruent trial (970 ms), $\beta = 39.70$, $SE = 17.71$, $z = 2.24$, $p = .025$, whereas Congruency on trial $n - 1$ had no impact on congruent trials, $\beta = 2.15$, $SE = 10.38$, $z = .21$, $p = .84$. Note that, as in the analysis reported in the main text of the article, Congruency and RT on trial $n - 1$ did not interact, $\beta = 5.90$, $SE = 5.49$, $z = 1.07$, $p = .28$, suggesting that there was no temporal learning mechanism being used.

In MI lists, the only significant effect was RT on trial $n - 1$, $\beta = 22.29$, $SE = 4.58$, $z = 4.86$, $p < .001$. In particular, there was neither an interaction between Congruency on trial $n - 1$ and Congruency, $\beta = 2.43$, $SE = 5.48$, $z = .51$, $p = .61$, nor a numerical tendency for a congruency sequence effect. Similar to the analysis reported in the main text of the article, there was a tendency for congruency effects to increase with higher RT on trial $n - 1$, which is the reverse of the pattern predicted by the temporal learning account (i.e., decreasing congruency effects with higher RT on trial $n - 1$). However, the interaction between Congruency and RT on trial $n - 1$ was not significant in this analysis, $\beta = -7.17$, $SE = 5.48$, $z = -1.31$, $p = .19$.

Error rates. Both Congruency, $\beta = .26$, $SE = .16$, $z = 1.67$, $p = .096$, and List Type, $\beta = -.31$, $SE = .18$, $z = -1.69$, $p = .091$, were marginally significant, with congruent items showing a tendency to elicit less errors than incongruent items and MC lists showing a tendency to elicit more errors than MI lists. Congruency on trial $n - 1$ was significant, $\beta = .42$, $SE = .16$, $z = 2.64$, $p = .008$, indicating that participants were more accurate following a congruent trial (1.6%) than an incongruent trial (3.2%). No other effect reached significance.

Experiment 1B (picture naming).

Response time. The main effects of Congruency (congruent faster than incongruent), $\beta = -143.42$, $SE = 3.44$, $z = -41.74$, $p < .001$, Congruency on trial $n - 1$ (faster responses following an incongruent trial), $\beta = 15.27$, $SE = 3.56$, $z = 4.29$, $p < .001$, List Type (MI faster than MC), $\beta = 22.68$, $SE = 7.63$, $z = 2.97$, $p = .003$, and RT on trial $n - 1$ (faster responses with lower RT on trial $n - 1$), $\beta = 30.80$, $SE = 4.32$, $z = 7.13$, $p < .001$, were all significant. Congruency and List Type interacted, $\beta = -20.67$, $SE = 4.12$, $z = -5.02$, $p < .001$, indicating a regular PC effect. Congruency also interacted with Congruency on trial $n - 1$, $\beta = -11.79$, $SE = 3.62$, $z = -3.26$, $p = .001$. This interaction indicates a regular congruency sequence effect, with a reduced congruency effect following an incongruent trial (274 ms) than following a congruent trial (307 ms). Again, the main reason for this reduction was that responses to incongruent trials were faster when following another incongruent trial (1060 ms) than when following a congruent trial (1083 ms), $\beta = 54.11$, $SE = 11.71$, $z = 4.62$, $p < .001$. In contrast, Congruency on trial $n - 1$ had no impact on congruent trials, $\beta = 6.96$, $SE = 8.32$, $z = .84$, $p = .40$.

There was also an interaction between Congruency on trial $n - 1$ and RT on trial $n - 1$, $\beta = 11.89$, $SE = 4.64$, $z = -2.56$, $p = .010$, with lower RT on trial $n - 1$ producing a larger speed-up for responses on trial n if trial $n - 1$ was congruent than if it was incongruent, and a marginal interaction between Congruency and RT on trial $n - 1$, $\beta = -7.44$, $SE = 3.97$, $z = -1.87$, $p = .061$, with a tendency for congruency effects to increase with higher RT on trial $n - 1$. The former interaction seems consistent with the idea that fast temporal expectancies produced by easy-to-process stimuli (i.e., congruent) have a larger impact on performance than do slower temporal expectancies produced by hard-to-process stimuli (i.e., incongruent; Schmidt & Weissman, 2016). On the other hand, the finding that congruency effects increased with higher RT on trial $n - 1$ reflects, once again, the reverse of the pattern predicted by the temporal learning account, according to which higher RT on trial $n - 1$ should reduce congruency effects.

Error rates. Congruency (congruent more accurate than incongruent) was the only significant effect, $\beta = 1.28$, $SE = .11$, $z = 11.28$, $p < .001$.

(Appendix continues)

Conclusion

Similar to what was found for the PC effect, temporal learning does not seem to provide a convincing explanation for the congruency sequence effect in the present dataset. According to the temporal learning account, a congruency sequence effect should emerge as a consequence of a mechanism whereby congruency effects decrease with higher RT on trial $n - 1$. However, in an analysis in which untransformed RTs were used (thus avoiding potential problems associated with nonlinear transformations of the dependent variable), we found, if anything, marginal evidence for the opposite pattern (i.e., congruency effects increasing with higher RT on trial $n - 1$) in Experiment 1B and a numerical tendency in the same direction in the MI list of Experiment 1A. Although this situation suggests that no temporal learning mechanism was being used, a regular congruency sequence effect emerged nonetheless.

It is worth noting that in the present analyses, not only temporal learning was controlled for but also feature and response repetitions were either removed (Experiment 1A) or minimal to begin with (i.e., there were no response repetitions in Experiment 1B because each trial required a different response). Therefore, we are inclined to interpret the congruency sequence effect that was obtained as resulting from a trial-to-trial conflict-adaptation mechanism (Botvinick et al., 2001), with recent experience with conflict

leading to focused attention, thereby decreasing interference, and recent experience with little or no conflict leading to relaxed attention, thereby increasing interference. This explanation would be consistent with the finding that the congruency sequence effect was mainly caused by facilitation for incongruent trials following another incongruent trial, a pattern that would reflect reduced interference when interference has recently been dealt with. This explanation would also seem to accommodate the fact that in the MI list in Experiment 1A, no congruency sequence effect was obtained. The high number of incongruent trials produced a complete elimination of the congruency effect in that list, suggesting that there was little conflict to adapt to. Indeed, it seems reasonable to assume that some amount of conflict is necessary in order for a trial-to-trial conflict adaptation mechanism to be operable. The core claim, in any case, is that not only the PC effect but also the congruency sequence effect, another important marker of conflict adaptation, emerges when potential confounds are accounted for, and most importantly, temporal learning does not seem to offer a convincing alternative to a control-based interpretation of this effect.

Received October 23, 2018

Revision received February 12, 2019

Accepted February 27, 2019 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <https://my.apa.org/portal/alerts/> and you will be notified by e-mail when issues of interest to you become available!