Robust evidence for proactive conflict adaptation in the proportion-congruent paradigm

Giacomo Spinelli[1] & Stephen J. Lupker[2]

[1] Università degli Studi di Milano-Bicocca

[2] University of Western Ontario

Running Header: Robust proactive conflict adaptation

Corresponding Author:

Giacomo Spinelli

Dipartimento di Psicologia, Università degli Studi di Milano-Bicocca

Piazza dell'Ateneo Nuovo 1, 20126 Milano (MI), Italy

E-mail: giacomospinelli@hotmail.it

Word count: 22698

Author note

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. Some of the data reported in this manuscript were presented at the Psychonomic Society's 60th Annual Meeting held in Montreal, Canada, November 14-17, 2019. The raw data and the JASP files used for the analyses are publicly available at https://osf.io/rnqh6/. The study materials are available upon request. The study was not preregistered. Correspondence concerning this article may be addressed to: Giacomo Spinelli, Dipartimento di Psicologia, Università degli Studi di Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milano (MI), Italy (e-mail: giacomospinelli@hotmail.it) or Stephen J. Lupker, Department of Psychology, University of Western Ontario, London, Ontario, N6A 5C2, Canada (e-mail: lupker@uwo.ca).

Abstract

In the standard Proportion-Congruent (PC) paradigm, performance is compared between a list containing mostly congruent (MC) stimuli (e.g., the word RED in the color red in the Stroop (1935) task) and a list containing mostly incongruent (MI) stimuli (e.g., the word BLUE in red). The PC effect, the finding that the congruency effect (i.e., the latency difference between incongruent and congruent stimuli) is typically larger in an MC list, has been interpreted by the popular conflict-monitoring account (Botvinick et al., 2001) as reflecting a proactive process whereby attention to task-relevant information is adapted based on how frequently conflict from task-irrelevant information arises. Recently, however, alternative accounts of the PC effect have emerged that assume either that the PC effect reflects processes other than proactive conflict adaptation (e.g., stimulus-response contingency learning) or that proactive conflict adaptation is only engaged as a last resort (e.g., when contingency learning cannot be used to minimize interference). We examined these ideas in three experiments in which proactive conflict adaptation could be evaluated independently from processes that are normally confounded with it in the PC paradigm, while still allowing those processes, particularly contingency learning, to be used to minimize interference. Consistent with the conflict-monitoring account of the PC effect, but inconsistent with all the alternative accounts of the PC effect, evidence for proactive conflict adaptation emerged in all experiments. Although multiple processes may be engaged in the PC paradigm, this paradigm remains a valid tool for examining proactive conflict adaptation, its typical use.

Keywords: proactive control; conflict adaptation; conflict monitoring; proportion-congruent effect; Stroop

Robust evidence for proactive conflict adaptation in the proportion-congruent paradigm

A commonly held position in cognitive psychology is that control processes have a fundamental role in minimizing the impact of conflict created by task-irrelevant information when processing task-relevant information. An example of such a situation would be the Stroop (1935) task in which participants are slower at naming the ink color of an incongruent color word (e.g., the word BLUE in the color red) than that of a congruent color word (e.g., the word RED in the color red). Although this congruency effect suggests that there is conflict created by an incongruent word, conflict that produces a processing cost, the fact that, in most cases, participants are ultimately able to name the color accurately suggests that that conflict is eventually resolved through the use of some sort of control process (for a review, see MacLeod, 1991).

A less established issue is whether, in addition to resolving conflict, this control system may also have the ability to *adapt* to situations in which conflict is expected in order to more effectively deal with such situations. The conflict-monitoring model (Botvinick et al., 2001) is possibly the most representative conceptualization of this idea (see also Braver, 2012; Kane & Engle, 2003). According to this model, stimuli are continuously monitored for the presence of conflict during processing and attention is adapted between task-relevant and task-irrelevant dimensions of the stimuli accordingly. Specifically, when conflict is detected (e.g., on an incongruent trial), a top-down signal is emitted indicating a need for more focused attention to task-relevant information (i.e., the color). This signal will not be emitted when little or no conflict is detected (e.g., on a congruent trial), causing a relaxation of attention in that situation. Importantly, this mechanism does not just help to resolve the conflict experienced on any given trial, it also influences subsequent performance in a proactive or anticipatory fashion. Specifically, experiencing conflict on a trial will induce focused attention to task-relevant information on subsequent trials. With the system already prepared for conflict, the interference produced by incongruent task-irrelevant information on subsequent trials will be reduced. Thus, the control system as theorized by the conflict-monitoring model would not just be limited to the resolution of present conflict but would also have a conflict-adaptation function.

One of the most important pieces of evidence supporting the idea of a conflict-adaptation function for the control system comes from the Proportion-Congruent (PC) paradigm. In this paradigm, the frequency of congruent and incongruent items (i.e., color-word compound stimuli) in a list is manipulated in order to create both a Mostly-Congruent (MC) list in which congruent items are frequent and incongruent items are infrequent, and a Mostly-Incongruent (MI) list in which incongruent items are frequent and congruent items are infrequent. The typical result is that the congruency effect is larger in the MC list than in the MI list, a finding known as the PC effect (e.g., Logan & Zbrodoff, 1979; Logan et al., 1984; Lowe & Mitterer, 1982; for a review, see Bugg & Crump, 2012).

The PC effect is easily interpreted within the conflict-monitoring framework as being the manifestation of a conflict-adaptation mechanism. According to this interpretation, frequent experience with the conflict created by incongruent trials in an MI list would induce increasingly focused attention to task-relevant information in that list. As a result of this advanced preparation for conflict, conflict will not produce a large performance cost, reducing the congruency effect in that situation. In contrast, infrequent experience with conflict in an MC list would induce increasingly relaxed attention in that list. As a result of this lack of advanced preparation for conflict, there will be a larger cost on the few trials when conflict does arise (i.e., on the few incongruent trials in that list), increasing the congruency effect in that situation.

The conflict-monitoring interpretation of the PC effect in Stroop-like tasks has had, and continues to have, a fundamental role in many lines of research employing the PC paradigm, including research on individual differences (e.g., Kane & Engle, 2003; Meier & Kane, 2013), development (e.g., Surrey et al., 2019; Wilk & Morton, 2012), aging (e.g., Bélanger et al., 2010; Cohen-Shikora et al., 2018), clinical disorders (e.g., Abrahamse et al., 2016; Bonnin et al., 2010), and the neural bases of cognitive control (e.g., De Pisapia & Braver, 2006; West & Alain, 2000). Despite its popularity, in recent years, this interpretation has also received increasing criticism, criticism that has led several researchers to reconsider the idea that the PC effect reflects a conflict-adaptation function as originally outlined in the conflict-monitoring model and to propose alternative accounts for this effect. Three alternative accounts have emerged: the item-specific account, the contingency-learning account, and the last-resort account.

The item-specific account

According to the conflict-monitoring model, conflict adaptation occurs in a proactive, item-nonspecific fashion. What that means is that, in the PC paradigm, experiencing conflict repeatedly (e.g., in an MI list) induces a focusing of attention that affects performance for all items indiscriminately. The reason is that that focusing of attention is set up to be applied to the next item before that item is processed. Similarly, experiencing lack of conflict repeatedly (e.g., in an MC list) induces a relaxation of attention before an item is processed. Therefore, responding to, for example, the incongruent word BLUE in red color would be equally difficult in an MC list whether that item followed the presentation of the congruent word BLUE in blue color (the same word as that displayed on the current trial) or whether it followed the presentation of the congruent word GREEN in green color (a different word than that displayed on the current trial), because both congruent words induce the same relaxation of attention, with specific word identities being irrelevant to the process.

Recent advances in cognitive control research, however, suggest that conflict adaptation can also occur in an item-specific fashion. For example, Jacoby et al. (2003) designed a new version of the PC paradigm in which half of the words were mainly presented in their congruent color (MC items, e.g., the words RED and BLUE presented in their corresponding colors) and the other half were mainly presented in an incongruent color (MI items, e.g., the words GREEN and YELLOW presented in their noncorresponding colors), and all words were intermixed in a single list. Similar to the standard (i.e., list-wide) PC effect, an item-specific PC effect emerged, with a larger congruency effect for MC items than for MI items (see also Crump et al., 2006).

In Jacoby et al.'s (2003) paradigm, congruent and incongruent items are equally probable in the list as a whole as well as being randomly presented. Thus, it would be impossible for an item-nonspecific conflict-adaptation process initiated before an item is presented to explain why the MC items would elicit a larger congruency effect than the MI items. In contrast, this result may be easily explained by an item-*specific* conflict-adaptation process initiated *after* the presented item is recognized. In this process, control is not regulated proactively based on general experience with conflict but *reactively* based on the experience with conflict that is specific to the word and/or color component of the presented item. For example, the recognition of an MI word (i.e., a frequently conflicting word, e.g., GREEN), would induce focused attention to the color, producing a reduced cost of conflict (i.e., a smaller congruency effect). On the other hand, the recognition of an MC word (i.e., an infrequently conflicting word, e.g., RED), would induce a relaxation of attention, thus producing an increased cost of conflict when the MC word does conflict with the color (i.e., a larger congruency effect).

Besides demonstrating that conflict adaptation can occur in an item-specific fashion (for corroborating evidence, see Bugg & Hutchison, 2013; Bugg et al., 2011; Spinelli & Lupker, 2020a, 2020b; Spinelli et al., 2020), the item-specific PC effect has raised the question of whether this item-specific conflict-adaptation process could also be the process underlying the standard PC effect. The reason that this question is relevant is that, in the standard PC paradigm, items in an MC list are typically MC items (i.e., all words in that list appear most often in their congruent color) and items in an MI list are typically MI items (i.e., all words in that list appear most often in incongruent colors). The implication is that the PC effect obtained in the standard PC paradigm is theoretically compatible not only with an item-nonspecific conflict-adaptation process, but also with an item-specific one.

More specifically, according to an item-specific account (Blais et al., 2007), the standard (i.e., list-wide) PC effect would derive from the same process that produces the difference between MC and MI items in the item-specific PC paradigm. The larger congruency effect observed in the MC list would depend on the relaxation of attention induced by the recognition of the many MC items in that list; similarly, the smaller congruency effect observed in the MI list would depend on the focusing of attention induced by the recognition of the many MI items in that list. Essentially, not only would the proactive conflict-adaptation process assumed by the original conflict-monitoring model (Botvinick et al., 2001) be unable to explain the item-specific PC effect for the reasons noted above, but it would also be unnecessary to invoke it in order to explain the standard PC effect. (note 1)

The contingency-learning account

Another alternative account of the PC effect is based on the observation that, because in PC paradigms words typically appear in some colors more often than in other colors in each of the lists being used, participants in those experiments may learn associations, or contingencies, between each word and its frequent color response in each list. This contingency-learning process would speed up latencies for high-contingency stimuli (i.e., words appearing in their frequent color) relative to low-contingency stimuli (i.e., words appearing in what are infrequent colors for them; Schmidt et al., 2007). Most importantly, this process would produce a PC effect without the need to assume conflict-adaptation processes of any sort (Schmidt, 2013a; see also Schmidt & Besner, 2008).

The reason is that, in a typical PC paradigm, any word in the MC list most frequently requires a congruent response (e.g., the word RED typically requires a “red” response as that word frequently occurs in the congruent red color). Learning these contingencies would result in a speed-up for (high-contingency) congruent colors relative to (low-contingency) incongruent colors, causing the congruency effect in that list to be sizeable. In contrast, in the MI list, a contingency-learning process would typically lead to a smaller congruency effect. This reduction in size would be most pronounced if the MI list is constructed (as it has often been; e.g., Logan et al., 1984) so that each word appears only in two colors, the infrequent congruent color and a frequent incongruent color. In that situation, contingency learning would speed up responses to the incongruent, but high-contingency, color relative to the congruent, but low-contingency, color. Alternatively, the MI list could be constructed so that no contingencies can be learned, for example, when four words and four colors are used in that list and each word appears equally often in each of the colors, one congruent and three incongruent (e.g., Cheesman & Merikle, 1986). In that case, the congruency effect would not be modified by any contingency-learning process because there are no contingencies to learn, but the effect would still be smaller than the congruency effect in an MC list in which contingency learning inflates that effect by speeding latencies to the frequent congruent items.

In sum, according to the contingency-learning account, the proactive conflict-adaptation process assumed by the original conflict-monitoring model (Botvinick et al., 2001) would be unnecessary for a PC effect to be produced in a typical PC paradigm. In fact, even an item-specific conflict-adaptation process would be unnecessary: General learning of simple stimulus-response (S-R) associations would be sufficient to provide a full account of both list-wide and item-specific PC effects.

The last-resort account

Although the item-specific account and the contingency-learning account differ as to what process underlies the PC effect, both accounts suggest that that process is not the proactive conflict-adaptation process assumed in the original conflict-monitoring model (Botvinick et al., 2001). To examine this idea further, Bugg et al. (2008) and Blais and Bunge (2010) used a modified PC paradigm which allowed a dissociation of that process from the item-specific and contingency-learning processes. In this paradigm, the items were divided into two sets, referred to as the “context” set and the “transfer” set, and congruency proportion was manipulated only for the context set. The transfer items were 50:50 congruent/incongruent and were intermixed in a list with either MC context items (creating an overall MC list) or MI context items (creating an overall MI list).

In such a paradigm, a PC effect for the context items might result from any of multiple processes (as is the case for any item in the standard PC paradigm in which no distinction is made between context and transfer items). Specifically, that effect might derive from 1) proactive conflict adaptation, based on the fact that the context items appear in an MC list vs. an MI list; 2) item-specific conflict adaptation, based on the fact that the context items are MC items in the MC list vs. MI items in the MI list; and/or 3) contingency learning, based on the fact that each of the context words has a high-contingency congruent color in the MC list but not in the MI list. More importantly, however, the only possible explanation for a PC effect for the transfer items would be proactive conflict adaptation. That is, because those items appear in different lists but are otherwise identical, neither item-specific conflict adaptation nor contingency learning can produce a PC effect for those items. Thus, a PC effect for those items would be unambiguous evidence for a proactive conflict-adaptation process, that is, the process assumed by the conflict-monitoring model. Conversely, failing to observe a PC effect for the transfer items while simultaneously observing one for the context items would suggest that, consistent with the item-specific and contingency-learning accounts, a proactive conflict-adaptation process was not being implemented by the control system.

Both Bugg et al. (2008) and Blais and Bunge (2010) reported the latter outcome, i.e., a regular PC effect for the context items but no PC effect for the transfer items. However, subsequent studies using a similar design did report a PC effect for the transfer items as well as for the context items (Bugg, 2014;

Bugg & Chanani, 2011; Gonthier et al., 2016; Hutchison, 2011; Schmidt, 2017; Spinelli & Lupker, 2020c; Spinelli et al., 2019). To explain this mixed evidence, Bugg (2014) proposed a hybrid account of the PC effect, originally termed the “associations-as-antagonists-to-control” account, which will be referred to as the “last-resort” account here for simplicity (see also Schmidt, 2019). Unlike the item-specific and contingency-learning accounts, this account does not negate the existence of a proactive conflict-adaptation process. However, based on the idea that the proactive conflict-adaptation process, unlike contingency learning, is relatively effortful (see also Braver, 2012), this account restricts the usage of that process selectively to situations in which contingency learning cannot be used on most trials in order to minimize interference from task-irrelevant information.

What are such situations? An MC list can never be one of them because contingency learning is always possible in such a list (at least for the words in the context set) and, hence, would be used instead of conflict adaptation. An MI list may not be one of those situations either if that list allows learning of contingencies between words and incongruent responses, e.g., if each of the context words appears in two colors, a high-contingency incongruent color and a low-contingency congruent color. Because contingency learning would tend to decrease the congruency effect for the context items in such a list (albeit not for transfer items), that process could certainly be used to minimize interference from task-irrelevant information on many trials. Hence, conflict adaptation would not be used in that case either. However, the same is not true for an MI list which does *not* allow learning of contingencies between words and incongruent responses, e.g., if each of the context words appears in four equally probable colors (one congruent and three incongruent). In that case, because no contingencies can be learned that would help deal with the conflict created by the frequent incongruent words in the list, a conflict-adaptation process leading to more focused attention to task-relevant information would be engaged as a last resort to achieve that goal. As a result of this proactive, item-nonspecific conflict-adaptation process, an overall reduction of the congruency effect, in comparison to that observed in an MC list, would be observed not only for the context items but also for any transfer items in that type of MI list.

Consistent with this account, PC effects on transfer items have most typically been reported in PC paradigms contrasting an MC list with an MI list that does not allow contingency learning for the context items. For example, in Bugg's (2014) Experiments 1a and 2b in which context words in the MI list appeared in four equally probable colors (one congruent and three incongruent), *preventing* contingency learning, transfer items produced a PC effect, suggesting that a proactive conflict-adaptation process was engaged in those experiments (see also Bugg & Chanani, 2011; Gonthier et al., 2016; Spinelli & Lupker, 2020c; Spinelli et al., 2019). Indeed, based on this idea, Braem et al. (2019) recently recommended the use of that type of design (i.e., a design in which context words in the MI list prevent contingency learning) to elicit proactive conflict adaptation. In contrast, in Bugg's (2014) Experiments 1b and 2a in which context words in the MI list appeared in two colors (a high-contingency incongruent one and the low-contingency congruent one), *allowing* contingency learning, transfer items did not produce a PC effect, suggesting that a proactive conflict-adaptation process was not used in those experiments.

Note that the latter experiments paralleled those previously conducted by Bugg et al. (2008) and Blais and Bunge (2010), who also had presented the context words in the MI list in two colors, allowing contingency learning, and who also had failed to obtain a PC effect on the transfer items (see also Bejjani et al., 2020). Thus, this pattern of results would seem to provide reasonable evidence that the control system does have a proactive conflict-adaptation function, but that function would only be implemented as a last resort, when no advantage in dealing with conflict could be derived from contingency learning (for alternative explanations for the PC effects obtained on transfer items in those cases, see Algom & Chajut, 2019; Schmidt, 2013b, 2017; Schmidt's temporal-learning explanation, in particular, will be addressed in the General Discussion).

The present research

Overall, what the alternative accounts of the PC effect reviewed above suggest is that the PC paradigm may be far from a surefire way to evaluate the proactive conflict-adaptation process originally assumed within the conflict-monitoring model (Botvinick et al., 2001). The reason is that that process may be, in the best-case scenario (i.e., that depicted by the last-resort account), largely situational and only engaged when other processes are unavailable to guide optimal task performance. As argued by Schmidt (2019), even this best-case scenario represents a large concession to the critics of the conflict-monitoring account of the PC effect because it implies that the vast majority of the studies in the literature that have used the PC paradigm rests on faulty assumptions about the process that that paradigm measures, and would need to be reinterpreted. Further, the fact that proactive conflict adaptation may only have a subsidiary role in goal-oriented behavior significantly reduces the theoretical relevance of that process. Finally, from a practical standpoint, the fact that measuring that process requires a somewhat complicated experimental design (e.g., a design with many responses, as multiple responses are necessary to prevent contingency learning on the context items) may discourage or even prevent the use of the context/transfer design in many situations (e.g., in situations requiring a limited number of responses, such as ones involving manual responding in neuroimaging experiments). In short, the claims made by proponents of the alternative accounts of the PC effect are serious and far from inconsequential for research using the popular PC paradigm to study proactive conflict adaptation.

With the present research, we aimed to re-evaluate this situation by re-examining the critical piece of evidence that challenges the original conflict-monitoring model but is well accommodated by all three alternative accounts of the PC effect: the absence of a PC effect on transfer items in experiments that allow contingency learning on context items in the MI list (i.e., Bugg et al.'s, 2008, Experiment 1, Blais & Bunge, 2010, and Bugg's, 2014, Experiments 1b and 2a). Although that null result has been reported in several experiments, there are a few reasons to suggest caution before that null result is embraced as representing the true state of things.

First and foremost, a significant PC effect on transfer items *has* been reported a couple times in experiments that allow contingency learning on context items in the MI list, the type of situation where that effect should *not* emerge according to the item-specific, contingency-learning, and last-resort accounts (Hutchison, 2011; Schmidt, 2017). In the context of a prime-probe task requiring participants to respond to a direction word (the probe, e.g., “left”) while ignoring another direction word (the prime) which could be congruent (e.g., “left”) or incongruent (e.g., “right”) with the probe, Schmidt (2017) found (in a control condition) a PC effect on transfer items even though the primes used for the context items were paired with only two probes, allowing contingency learning for those items in both the MC list and the MI list. (That PC effect, however, disappeared in Schmidt’s experimental condition, a condition aimed to prevent use of temporal information from influencing that effect.)

In the context of the color-word Stroop task, albeit in a somewhat more complicated design than the ones reviewed thus far, a design which allowed independent examinations of list-wide PC effects, item-specific PC effects, and contingency-learning effects, Hutchison (2011) also found a (list-wide) PC effect on transfer items when, in order to create an MI list, those items were intermixed with context items using words that did not allow contingency learning. Interestingly, compared to the MC list (in which the context words, unavoidably, allowed contingency learning), that MI list in which the context words did not allow contingency learning produced the same congruency effect reduction for the transfer items as did another MI list used in that experiment in which the context words did allow contingency learning. Note, on the other hand, that in Hutchison’s design, contingency learning was possible in all lists for the transfer items, and there was evidence that that process did take place, as demonstrated by faster responses to high- than low-contingency transfer items matched on congruency and item-specific congruency proportion in all lists. In general, however, in the MI list in which the context items did not allow contingency learning, compared to the MI list in which the context items did allow contingency learning, there were substantially fewer occasions to use contingencies to minimize interference by linking a word with an incongruent color response. According to the last-resort account, it is in this type of situation that proactive conflict adaptation should be favored. The fact that, instead, a PC effect on the transfer items was also obtained when the MC list was contrasted with the other type of MI list, the list that allowed contingency learning for the context items, suggests that proactive conflict adaptation may be used more generally, inconsistent with the last-resort account as well as the item-specific and contingency-learning accounts.

Second, there may be reasons other than prioritization of contingency learning (the type of explanation proposed by the last-resort account) that could explain the failures to obtain PC effects on transfer items that have been reported in some studies. One such reason may be the common tendency in the context/transfer paradigm for the stimuli in the context set to occur more frequently than the stimuli in the transfer set. For example, in Bugg et al.'s (2008) Experiment 1 and Bugg's (2014) Experiment 2a, each individual color and each individual word in the context set was twice as frequent as each individual color and each individual word in the transfer set, and four times as frequent in Bugg's (2014) Experiment 1b. As recently argued by Braem et al. (2019), this design choice certainly has the advantage of creating a stronger congruency-proportion manipulation at the list level because, e.g., mixing 75%-congruent context items with 50%-congruent transfer items only yields a 62.5%-congruent list if context and transfer items are overall equally frequent, while yielding a 70%-congruent list if context items are four times more frequent than the transfer items (a similar point would apply when considering the MI list). However, the relatively infrequent presentation of transfer items in both lists may have had the side effect of attracting the participant's attention to those surprising items (Johnston et al., 1990), making it harder to observe any difference in attentional states between lists (relaxed attention in the MC list vs. focused attention in the MI list) on those items (for a similar point, see Gonthier et al., 2016).

Consistent with this idea, unpublished data from picture-word interference experiments conducted in our lab (Spinelli & Lupker, in preparation) show that when unrepeated transfer items are intermixed with context items that are also unrepeated, a regular PC effect emerges on the transfer items. However, when individual context items are repeated several times but transfer items remain unrepeated, thus becoming individually less frequent (i.e., more surprising) than the context items, the PC effect on those items vanishes. (note 2)

Third, another reason why a PC effect on transfer items may not have been obtained in some studies, particularly in Blais and Bunge's (2010) study, is the use of multiple arbitrary responses. Blais and Bunge used 8 to 12 ink colors overall (although only 4 colors per block were presented) and required participants to respond to those colors with manual (keypress) responses (unlike most other relevant studies, where vocal responses were required), while receiving no feedback on their performance. Because associations between colors and manual responses are completely arbitrary, participants in those experiments were tasked not only with identifying the ink colors while ignoring the words (the standard Stroop instructions) but also with learning and maintaining the instructed S-R mappings throughout the experiment. When knowledge of many such mappings is required, the latter task may create such high working-memory demands that participants are prevented from applying proactive conflict adaptation, a type of process for which intact working-memory resources may be required (Braver, 2012).

Consistent with this idea, Bejjani et al. (2020) recently failed to obtain a PC effect on transfer items in a similar situation as that examined by Blais and Bunge (2010), i.e., a manual-response picture-word interference task with no performance feedback. Interestingly, however, they did obtain that effect when performance (i.e., accuracy) feedback was displayed, possibly because the presence of feedback allowed participants to rely less heavily on their working memory to maintain the instructed S-R mappings (although for alternative explanations, see Bejjani et al., 2020).

The main point is that many of the reported failures to obtain evidence for proactive conflict adaptation in the form of a PC effect on transfer items, particularly in situations where context items in the MI list allowed contingency learning, may have to do with suboptimal methodological choices that were made in those studies. In the present research, we addressed the question of whether that evidence would emerge in those situations using what we believe to be a more appropriate methodology, specifically, a context/transfer paradigm in which: 1) context and transfer items were presented equally often (to avoid attention capture from infrequent transfer items), and 2) vocal responses to colors were required (to avoid working-memory demands created by arbitrary S-R mappings).

Importantly, these methodological changes not only represent an improvement over past studies in that they remove potential sources of noise but also they are not directly relevant to any of the alternative accounts of the PC effect that have been proposed. What is relevant to those accounts is whether the PC paradigm being used could be affected by item-specific conflict adaptation (the item-specific account), contingency learning (the contingency-learning account), or overall reliability of contingency learning in the MI list (the last-resort account). The situation that we examined in the present research, a PC paradigm where context words in the MI list allowed contingency learning, ticks all those boxes because it is a situation in which all three accounts predict that no evidence for proactive conflict adaptation (i.e., no PC effect on the transfer items) should emerge. (note 3) Thus, the present research represented a fair and strong test of those accounts against the original conflict-monitoring idea that proactive conflict adaptation should occur even in that type of situation.

Experiment 1

In this experiment, we adapted the context/transfer paradigm developed by Bugg et al. (2008), Blais and Bunge (2010), and Bugg (2014) to a vocal Stroop task involving eight colors and the corresponding color names. The frequency of the color-word combinations in one of the counterbalancings of the experiment is presented in Table 1 for the MC list and Table 2 for the MI list. In each list, two non-overlapping sets of items were created by combining four colors with the corresponding color names. One set (e.g., the colors yellow, black, blue, pink, and their corresponding color names) was used for the transfer items, which were 50:50 congruent/incongruent in both lists. The other set (e.g., the colors red, white, green, purple, and their corresponding color names) was used for the context items, which were MC items in the MC list and MI items in the MI list.

Table 1

Template for the Frequency of Color-Word Combinations in the MC List in Experiment 1

		Word							
		Context				Transfer			
Color		RED	WHITE	GREEN	PURPLE	YELLOW	BLACK	BLUE	PINK
Context	Red	21	1	1	1				
	White	1	21	1	1				
	Green	1	1	21	1				
	Purple	1	1	1	21				
Transfer	Yellow					12	4	4	4
	Black					4	12	4	4
	Blue					4	4	12	4
	Pink					4	4	4	12

Table 2

Template for the Frequency of Color-Word Combinations in the MI List in Experiment 1

		Word							
		Context				Transfer			
Color		RED	WHITE	GREEN	PURPLE	YELLOW	BLACK	BLUE	PINK
Context	Red	1	21	1	1				
	White	1	1	21	1				
	Green	1	1	1	21				
	Purple	21	1	1	1				
Transfer	Yellow					12	4	4	4
	Black					4	12	4	4
	Blue					4	4	12	4
	Pink					4	4	4	12

Note. The (incongruent) items in the cells shaded in grey are those that were used to examine contingency learning independently from other processes. The items shaded in light grey were the high-contingency incongruent items and those shaded in dark grey were the low-contingency incongruent items.

Inevitably, the context words in the MC list allowed contingency learning, having the congruent color as the high-contingency color and three incongruent colors as low-contingency colors. For example, the context word RED appeared many more times in red (the high-contingency color) than in white, green, and purple (the other colors in the context set, which for the word RED were the low-contingency colors; see Table 1). More importantly, the context words in the MI list also allowed contingency learning because for each word in this set, one incongruent color was used as the high-contingency color and the other three colors (the other two incongruent colors and the congruent color) were used as low-contingency colors. For example, the context word RED appeared many more times in purple (the high-contingency color) than in red, white, and green (the low-contingency colors; see Table 2).

This design allowed us to undertake a number of relevant analyses. First, for context items, it allowed the classic Proportion-Congruent analysis contrasting congruent vs. incongruent items in the MC vs. MI list (with high-contingency and low-contingency incongruent items in the context set of the MI list being collapsed). Although this analysis virtually always produces a PC effect (i.e., a larger congruency effect for context items in the MC list than in the MI list), it is not possible to establish its source. The reason, as noted, is that a PC effect on the context items may be caused by proactive conflict adaptation (because the context items appear in an MC list vs. an MI list), item-specific conflict adaptation (because the context items are MC items in the MC list vs. MI items in the MI list), and/or contingency learning (because each of the context words has a high-contingency congruent color in the MC list but not in the MI list, where the congruent color is a low-contingency color).

However, unlike most experiments using the context/transfer paradigm (except for Hutchison, 2011), this design did allow a contingency-learning analysis. This analysis was based on the contrast between high-contingency incongruent context items in the MI list (shaded in light grey in Table 2, e.g., RED in purple) and low-contingency incongruent context items in the same list (shaded in dark grey in Table 2, e.g., RED in white). Because those items belonged to the same list (i.e., the MI list) and the same item set (i.e., the MI context set), whatever difference emerges in that contrast could not be the result of proactive or item-specific conflict adaptation. Instead, it must be the result of a contingency-learning process, which should favor responding to a word when it appears in its high-contingency (incongruent) color relative to its low-contingency (incongruent) colors (for a similar design, see Spinelli & Lupker, 2020b). By dissociating contingency learning from item-specific and proactive conflict-adaptation processes, this analysis allowed an examination of whether participants do engage in contingency learning in PC paradigms, as both the contingency-learning account and the last-resort account assume.

On the other hand, note that because this contingency-learning contrast was located only within the MI context set, a set of stimuli that only appeared in the MI list, it was impossible to examine the impact of the PC manipulation on the contingency-learning process itself. Similarly, it was impossible to examine the impact of the PC manipulation on the item-specific conflict-adaptation process that the context items allowed participants to engage. For such examinations to be possible, the contingency-learning and item-specific contrasts would have to be located in the transfer set, a situation that Hutchison (2011) created in his experiment. For the present purposes, however, it was sufficient to show that contingency learning took place in the MI list, the type of list in which, according to the last-resort account, the possibility of learning contingencies will determine the dominant process in that list. Our contingency-learning contrast in the MI context set was appropriate for that purpose.

Finally and most importantly, the present design allowed the decisive analysis for examining the presence of a proactive conflict-adaptation process, a Proportion-Congruent analysis on transfer items contrasting congruent vs. incongruent items in the MC vs. MI lists. As noted, neither item-specific conflict adaptation nor contingency learning could affect the results of this analysis because the analysis is based on items – the transfer items – that are identical in the two lists. Therefore, a PC effect on those items (i.e., a larger congruency effect in the MC list than in the MI list) must be the result of a proactive conflict-adaptation process driven by the nature of the context items. (note 4)

As also noted, although this result is the one predicted by the original conflict-monitoring model (Botvinick et al., 2001), it is not the result predicted by the alternative accounts of the PC effect. According to those accounts, no PC effect should emerge on the transfer items (i.e., congruency effects for those items should be equivalent in the MC and MI lists), either because the existence of a proactive conflict-adaptation process capable of generating that effect is denied (the item-specific and contingency-learning accounts) or because the possibility that that process would be engaged in the type of situation examined in this experiment is negated, due to the fact that contingencies can be learned for context words in the MI list, therefore making contingency learning the preferred process (the last-resort account).

In addition to allowing a dissociation of contingency learning from other processes, the other novel aspect of this experiment was that it avoided two potentially concerning features of previous studies in the context/transfer paradigm: the infrequent presentation of transfer items relative to context items (Bugg, 2014; Bugg et al., 2008) and the use of multiple arbitrary responses (Blais & Bunge, 2010). As discussed, those features may induce the engagement of processes (i.e., attention capture from infrequent items and active maintenance of arbitrary S-R mappings in working memory, respectively) which are irrelevant to current accounts of the PC effect but may make it difficult to observe that effect, particularly on the transfer items for which only one process (proactive conflict adaptation) would be capable of producing such an effect. Those irrelevant processes were negated in this and the following experiments because 1) each of the words and colors in the context and transfer sets were presented equally often in each list (e.g., because the transfer word YELLOW appeared as often as any context word, that word was not an unusual (attention-capturing) stimulus) and 2) vocal responses to colors were used (i.e., responses which, being non-arbitrary, involved no particular working-memory demands). The predictions of the extant accounts of the PC effect, particularly regarding the emergence of this effect on the transfer items in a situation favoring contingency learning, could thus be examined without those processes potentially affecting the results.

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 PC effect. This analysis was based on the sizes of the PC effects on transfer items reported by Bugg (2014) in Experiments 1a and 2b, the only color-word Stroop experiments in the literature besides Hutchison's (2011) where transfer items did produce a PC effect. Based on the smallest of those effect sizes (η_p^2 = .190, reported for Bugg's (2014) Experiment 1a), we determined that a minimum sample size of 38 participants would be needed.

Forty-eight participants took part in the experiment. All participants in this experiment were students at the University of Western Ontario (age 17–31 years) who participated for course credit. They were all native English speakers and had normal or corrected-to-normal vision.

Materials

Eight color names (RED, WHITE, GREEN, PURPLE, YELLOW, BLACK, BLUE, and PINK) were used as distractors, and the corresponding colors (red [R: 255; G: 0; B: 0], white [R: 255; G: 255; B: 255], green [R: 0; G: 192; B: 0], purple [R: 128; G: 0; B: 192], yellow [R: 255; G: 255; B: 0], black [R: 0; G: 0; B: 0], blue [R: 0; G: 0; B: 255], and pink [R: 255; G: 128; B: 192]) were used as targets. The stimuli were divided into two sets, with RED, WHITE, GREEN, PURPLE and their corresponding colors forming one set, and YELLOW, BLACK, BLUE, PINK, and their corresponding colors forming the other set (see Tables 1 and 2). One set (e.g., RED, WHITE, GREEN, PURPLE and their corresponding colors) served as the context set and the other set (e.g., YELLOW, BLACK, BLUE, PINK, and their corresponding colors) served as the transfer set for each participant.

In the MC list, each word in the context set (e.g., RED) appeared 21 times with the congruent color (e.g., red, the high-contingency color) and once with each of the three incongruent colors in the set (e.g., white, green, and purple, the low-contingency colors). Overall, there were 84 congruent items and 12 incongruent items in the context set in the MC list, an item-specific congruency proportion of 87.5%. Similarly, in the MI list, each word in the context set (e.g., RED) appeared 21 times with one incongruent color (e.g., purple, the high-contingency color) and once with each of the other three colors in the set, including the congruent color (e.g., red, white, and green, the low-contingency colors). Overall, there were 4 congruent items and 92 incongruent items in the context set in the MI list, an item-specific congruency proportion of 4.17%.

Each word in the transfer set (e.g., YELLOW) appeared 12 times with the congruent color (e.g., yellow) and 4 times with each of the incongruent colors in the set (e.g., black, blue, and pink) in both lists. Overall, there were 48 congruent items and 48 incongruent items in the transfer set in both lists, an item-specific congruency proportion of 50%. However, considering both context and transfer items, there were overall 132 congruent items and 60 incongruent items in the MC list (a list-wide congruency proportion of 68.75%) and 52 congruent items and 140 incongruent items in the MI list (a list-wide congruency proportion of 27.08%). Note, further, that the contingency manipulations were parallel for words in the MC and MI lists because in each list, participants could learn a 21-to-1 contingency for context words (i.e., the (congruent or incongruent) high-contingency color was 21 times more likely than any of the (congruent or incongruent) low-contingency colors) and a 3-to-1 contingency for transfer words (i.e., the (congruent) high-contingency color was 3 times more likely than any of the (incongruent) low-contingency colors). For each list, we also calculated C, a chi-square based contingency coefficient measuring the strength of the correlation between word and color values, using Melara and Algom's

(2003) formula. The absolute value of C was .88 for both lists, indicating that the absolute strength of color-word correlations was the same in the two lists (for a discussion of the potential role of color-word correlations in Stroop tasks, see footnote 3).

In both lists, the context set and the transfer set were randomly intermixed. The assignment of the two sets of words/colors as context versus transfer items was counterbalanced across participants, as was the order with which the MC and MI lists were presented. (note 5) The specific incongruent color serving as the high-contingency color for context words in the MI list was also counterbalanced across participants.

Procedure

Each trial began with a fixation symbol ("+") displayed for 250 ms in the center of the screen followed by a colored word displayed for 2000 ms or until the participant's response, which was recorded with a microphone connected to the testing computer. Participants were instructed to name the color of the word as quickly and as accurately as possible while ignoring the word itself. They were told about the colors that would appear in the experiment. Stimuli were presented in uppercase Courier New font, pt. 14, against a medium grey background (R: 128; G: 128; B: 128). No feedback was provided. There was a self-paced pause between the two lists. The order of trials within each list was randomized. Initially, participants performed a practice session including 8 trials in which a string of Xs ("XXXX") was presented in each of the eight colors used in the experiment. The experiment was run using DMDX (Forster & Forster, 2003) software. This research was approved by the Research Ethics Board of the University of Western Ontario (protocol # 108956).

Results

For this and the following experiments, the waveforms of responses were manually inspected with CheckVocal (Protopapas, 2007) in order to determine the accuracy of the response and the correct placement of timing marks. Prior to the analyses, invalid trials due to technical failures and responses faster than 300 ms or slower than 2000 ms, the time limit (accounting for .6% of the data points), were discarded. Separate analyses were conducted for context and transfer items, paralleling previous research using the context/transfer paradigm (e.g., Bugg, 2014; Bugg et al., 2008).

For both context and transfer items, a Proportion-Congruent analysis was conducted on both latencies and errors with Congruency (Congruent vs. Incongruent) and List Type (Mostly congruent vs. Mostly incongruent) as within-subject factors. For the context items, no distinction was made between the low-contingency and the high-contingency incongruent items in the MI list (the data from all incongruent context items in that list were collapsed). With this analysis, we aimed to examine the presence of a PC effect on the context items and, more importantly, on the transfer items.

For context items, a contingency-learning analysis was also conducted on both latencies and errors contrasting the low-contingency incongruent items in the MI list (the items shaded in dark grey in Table 2) and the high-contingency incongruent items also in the MI list (the items shaded in light grey in Table 2), a within-subject contrast. With this analysis, we aimed to examine the presence of a contingency-learning process on the context items. To this end, a paired-samples *t*-test was conducted contrasting the null hypothesis of equivalent RTs and error rates for high- and low-contingency items to the one-sided alternative hypothesis of smaller RTs and error rates for high- than low-contingency items.

In addition to traditional null-hypothesis significance testing analyses, in order to quantify the evidence supporting the presence vs. the absence of theoretically relevant effects (the interaction between Congruency and List Type, i.e., the PC effect, for context and transfer items, and the contingency-learning effect for context items), we also performed Bayes Factor analyses for those effects. These analyses were performed in JASP version 0.14.1 (JASP Team, 2020) by comparing the model without the effect of interest (interpreted as the null hypothesis H_0) and the model with that effect (interpreted as the alternative hypothesis H_1) using the default settings. The result of this comparison was BF_{10}, with $BF_{10} > 1$ suggesting evidence in support of H_1 (i.e., the presence of the effect), and $BF_{10} < 1$ suggesting evidence in support of H_0 (i.e., the absence of the effect) ($BF_{10} = 1$ would suggest equal evidence for the two hypotheses). Note that for the contingency-learning analysis, the notation BF_{-0} (note the minus in the subscript) will be used to denote the directional nature of the alternative hypothesis (i.e., high-contingency smaller than low-contingency). Jeffreys's (1961) classification scheme (as reported in adjusted form by Lee and Wagenmakers, 2013) was used to help interpret the size of the Bayes Factor.

The mean RTs and error rates are presented in Table 3. For this and the following experiments, the raw data and JASP files used for the analyses are publicly available at https://osf.io/rnqh6/. The study materials are available upon request. The study was not preregistered.

Table 3

Mean RTs and Percentage Error Rates (and Corresponding 95% Confidence Intervals) for Context and Transfer Items in Experiment 1

	RTs		Error rates	
Item type	MC list	MI list	MC list	MI list
Context items				
Congruent	686 [660, 712]	742 [697, 788]	.15 [-.01, .31]	.52 [-.53, 1.57]
Incongruent	857 [818, 897]	800 [764, 837]	2.45 [1.22, 3.67]	1.76 [1.14, 2.39]
Congruency Effect	171	58	2.30	1.24
High-contingency		799 [762, 836]		1.73 [1.09, 2.36]
Low-contingency		814 [777, 851]		2.19 [.75, 3.64]
Contingency-learning Effect		15		.46
Transfer items				
Congruent	705 [677, 733]	730 [695, 765]	.52 [.16, .89]	.57 [.18, .96]
Incongruent	836 [799, 873]	824 [784, 864]	3.79 [2.58, 5.01]	2.41 [1.59, 3.23]
Congruency Effect	131	84	3.27	1.84

Context items

Proportion-Congruent analysis

RTs. There was a main effect of Congruency, $F(1, 46) = 208.58$, $MSE = 3019$, $p < .001$, $\eta_p^2 = .816$, indicating faster responses to congruent than incongruent items, but no main effect of List Type, $F(1, 46) < .01$, $MSE = 6068$, $p = .969$, $\eta_p^2 < .001$. However, List Type interacted with Congruency, $F(1, 46) = 38.95$, $MSE = 3969$, $p < .001$, $\eta_p^2 = .453$. The interaction reflected the typical PC effect, with a larger congruency effect in the MC list (171 ms) than in the MI list (58 ms). The Bayes Factor for the comparison between the model with the interaction and the model without it was $BF_{10} = 483067.74 \pm 4.47\%$, meaning that the data were 483067.74 times more likely to occur under the hypothesis of an interaction than under the hypothesis of no interaction. In Jeffreys's (1961) classification scheme, this value would suggest "extreme" evidence for the presence of the interaction.

Error rates. The only significant effect was the main effect of Congruency, $F(1, 46) = 20.86$, $MSE = .001$, $p < .001$, $\eta_p^2 = .307$, with congruent items eliciting fewer errors than incongruent items. Numerically, the congruency effect was larger in the MC list (2.30%) than in the MI list (1.24%), however, there was no significant interaction between Congruency and List type, $F(1, 46) = 1.53$, $MSE = .001$, $p = .222$, $\eta_p^2 = .032$. On the other hand, the Bayes Factor, $BF_{10} = 1.65 \pm 4.55\%$, indicated "anecdotal" evidence for the presence of the interaction. There was no main effect of List Type either, $F(1, 46) = .17$, $MSE = .001$, $p = .685$, $\eta_p^2 = .004$.

Contingency-learning analysis

A contingency-learning effect emerged in the latencies, $t(47) = -2.13$, $p = .019$, $\eta_p^2 = .088$, with faster responses to high-contingency items (799 ms) than to low-contingency items (814 ms). The Bayes Factor, however, $BF_{-0} = 2.42 \pm .00\%$, indicated only "anecdotal" evidence for the presence of this effect.

In the error rates, although the high-contingency items produced slightly lower error rates (1.23%) than low-contingency items (2.19%), this effect did not approach significance, $t(47) = -.67$, $p = .254$, $\eta_p^2 = .009$, and the Bayes Factor, $BF_{-0} = .29 \pm .00\%$, indicated “moderate” evidence for the absence of this effect.

Transfer items (Proportion-Congruent analysis)

RTs. There was a main effect of Congruency, $F(1, 46) = 254.76$, $MSE = 2374$, $p < .001$, $\eta_p^2 = .844$, indicating faster responses to congruent than incongruent items, but no main effect of List Type, $F(1, 46) = .60$, $MSE = 3351$, $p = .443$, $\eta_p^2 = .013$. Importantly, Congruency and List Type interacted, $F(1, 46) = 17.93$, $MSE = 923$, $p < .001$, $\eta_p^2 = .276$, with the Bayes Factor, $BF_{10} = 5.39 \pm 2.66\%$, indicating “moderate” evidence for the presence of the interaction. The congruency effect was larger in the MC list (131 ms) than in the MI list (84 ms), the typical pattern of the PC effect.

Error rates. There was a significant main effect of Congruency, $F(1, 46) = 49.34$, $MSE = .001$, $p < .001$, $\eta_p^2 = .512$, with congruent items eliciting fewer errors than incongruent items, and a marginal main effect of List Type, $F(1, 46) = 3.78$, $MSE = .001$, $p = .058$, $\eta_p^2 = .074$, indicating a tendency for the MI list to produce a lower error rate than the MC list overall. There was also a significant interaction between Congruency and List Type, $F(1, 46) = 6.01$, $MSE < .001$, $p = .018$, $\eta_p^2 = .113$, although the Bayes Factor, $BF_{10} = 1.77 \pm 4.74\%$, indicated only “anecdotal” evidence for the presence of the interaction. In this case as well, the interaction reflected the typical PC effect, with a larger congruency effect in the MC list (3.27%) than in the MI list (1.84%).

Discussion

The results of Experiment 1 were straightforward. First, in line with previous research using the context/transfer paradigm (e.g., Blais & Bunge, 2010; Bugg, 2014; Bugg et al., 2008), the context items produced a (large) PC effect in the latencies (although not in the error rates), with a larger congruency effect in the MC list than in the MI list. As noted, no firm conclusions can be made about the source of this effect because the effect could be the result of any of a number of processes. An assumption that is commonly made, though, is that the effect results, at least in part, from a process of word-color contingency learning (e.g., Bugg, 2014; Schmidt, 2013a). The results of Experiment 1 support this assumption. In a design where contingency learning could be dissociated from other processes for context items in the MI list, we found faster latencies for high- than low-contingency items matched on all other relevant aspects, although this effect was not particularly large (15 ms despite a strong, 21-to-1 contingency manipulation; note that contingency-learning effects have been twice as large in weaker, e.g., 8-to-1, contingency manipulations: e.g., Forrin & MacLeod, 2017; Spinelli et al., in review) and its presence was not strongly favored by the Bayes Factor analysis. Nevertheless, the presence of that effect does suggest that contingency learning was engaged and likely contributed to the PC effect observed for the context items.

What is crucial here is the fact that contingency learning was observed *in the MI list*. According to the last-resort account of the PC effect (Bugg, 2014), in this type of situation (i.e., an MI list where context words allow contingency learning), contingency learning would not only be engaged (as our results confirm), but it would be engaged *instead of* proactive conflict adaptation. Therefore, no evidence for proactive conflict adaptation should have emerged for the items in the experiment specifically designed to evaluate the existence of this process, i.e., the transfer items. Because both the item-specific account (Blais et al., 2007) and the contingency-learning account (Schmidt, 2013a) of the PC effect deny the existence of proactive conflict adaptation, those accounts also make the same prediction.

In contrast with all three accounts, evidence for proactive conflict adaptation did emerge in Experiment 1 in the form of a PC effect on the transfer items: The congruency effect was larger in the MC list than in the MI list, with the Bayes Factor analyses favoring the presence of the interaction, especially for the latencies. Because the transfer items were identical in the two lists, there would have been no differences across lists for those items in terms of item-specific conflict frequency or contingency learning that could explain that result. Therefore, this result must reflect a process of proactive conflict adaptation as assumed by the conflict-monitoring model (Botvinick et al., 2001), a general process that would induce focused attention in a situation where conflict is frequent (i.e., the MI list), producing a reduced congruency effect in that situation, and relaxed attention in a situation where conflict is infrequent (i.e., the MC list), producing a larger congruency effect in that situation.

To our knowledge, besides Hutchison (2011), this is the first set of results in the color-word Stroop task that supports a role for proactive conflict adaptation in a situation in which contingency learning could be concurrently engaged as a solid alternative process. As noted, this type of evidence did not emerge in either Bugg's experiments (Bugg, 2014; Bugg et al., 2008) or Blais and Bunge's (2010). In the following experiments, we examined whether the evidence for proactive conflict adaptation that we obtained in Experiment 1 would remain in experiments that more closely resembled the situations examined by those researchers.

Experiment 2

Experiment 1 produced a PC effect for items in a situation in which neither item-specific conflict frequency nor contingency learning could have produced the effect even though contingency learning was a viable process in the experiment as a whole. This effect must thus reflect a different process, most likely the process of proactive conflict adaptation (although see footnote 4). Therefore, its presence contradicts the most recently developed accounts of the PC effect, accounts which assume that proactive conflict adaptation would not be engaged, either at all or in the type of situation examined in that experiment.

What is particularly surprising about that effect is the fact that it emerged in a similar paradigm as that used in Bugg's (2014; Bugg et al., 2008) and Blais and Bunge's (2010) experiments even though those experiments failed to obtain such an effect. However, it is important to note that there was a difference between the design used in Experiment 1 and that used in Bugg's as well as Blais and Bunge's experiments. Except for Bugg's (2014) Experiments 1a and 2b (see also Hutchison, 2011), all those experiments used a so-called two-item set design for the context words, a design in which each of those words was only combined with two colors, the congruent color and an incongruent color. That is, there were only two context words, with one of them appearing in the high-contingency color (i.e., the congruent color in the MC list and the incongruent color assigned to that word in the MI list) and only *one* word appearing in the low-contingency color (i.e., the assigned incongruent color in the MC list and the congruent color in the MI list). In contrast, Experiment 1 used a four-item set design in which each word was combined with four colors, the congruent color and three incongruent colors. Therefore, the context words allowed contingency learning in Experiment 1 because one of those words appeared in their high-contingency color (i.e., the congruent color in the MC list and one of the incongruent colors assigned to that word in the MI list) with there being *three* low-contingency word-color pairings (i.e., the incongruent colors assigned to that word in the MC list and the congruent color and the other two assigned incongruent colors in the MI list).

As a result, although the four-item set design used in Experiment 1 had the advantage of allowing a dissociation of contingency learning from other processes, it may have created a somewhat different situation than the alternative accounts of the PC effect, particularly the last-resort account, were intended to address. More specifically, according to the last-resort account, as noted, contingency learning would be the preferred process in an MI list *if engagement of that process would allow a minimization of the interference from task-irrelevant information*, i.e., the interference experienced on the frequent incongruent trials in that list (Bugg, 2014). That is, although contingency learning is considered an implicit process that is often out of conscious control (e.g., Schmidt et al., 2007), its usage may be modulated depending on the experimental situation (e.g., Schmidt, 2019; Schmidt et al., 2010).

Specifically, it could be argued that, although the context words in the MI list in Experiment 1 did allow contingency learning, engaging contingency learning would not necessarily minimize the overall interference on incongruent trials to any major degree. Considering the high-contingency incongruent context items in the MI list (e.g., RED in purple – see Table 2), it would certainly be the case that engaging a contingency-learning process would help reduce interference for the majority of the incongruent trials in that list (84 trials out of 140). The reason is that, although there is a conflict between the word (RED) and the color (purple) for that type of item, the color is the expected one for that word. On the other hand, for low-contingency incongruent context items (e.g., RED in white, WHITE in green; 8 trials), contingency learning would result in *increased* interference. The reason is that, for that type of item, there would be a conflict not only between the word (RED) and the color (white) but also between the expected (high-contingency) color (purple for the word RED) and the actual (low-contingency) color (white).

Similarly, for incongruent transfer items (e.g., YELLOW in black; 48 trials), items that are also low-contingency in the four-item set design used in Experiment 1, engaging contingency learning could also produce some additional interference. The reason is that the expected (high-contingency) color (e.g., yellow, the high-contingency color for the word YELLOW as it occurs on half of the trials when the word is yellow) would always conflict with the actual (low-contingency) colors (e.g., black). Thus, in the MI list used in Experiment 1, overall, contingency learning may not have been a particularly beneficial process to engage as a means for minimizing interference on the frequent incongruent trials in that list. Essentially, contingency learning would have helped on only on 60% of the incongruent trials (the 84 high-contingency incongruent context items) whereas it would have been misleading on the remaining 40% (the 8 low-contingency incongruent context items and the 48 low-contingency incongruent transfer items).

From this point of view, the fact that participants in Experiment 1 engaged proactive conflict adaptation, the "last resort" according to Bugg (2014), would not be especially surprising to proponents of that account, nor would the fact that the participants in Experiment 1 showed a somewhat small contingency-learning effect in the MI list. That is, because contingency learning in the MI list would not often help minimize interference on incongruent trials, participants may not adopt only that process and, instead, may also engage proactive conflict adaptation. By focusing attention to task-relevant information, the proactive conflict-adaptation process, unlike the contingency-learning process, would *always* help minimize interference on incongruent trials. Based on these considerations, the pattern of results obtained in Experiment 1 would be reasonably consistent with the last-resort account, although it would still be inconsistent with the item-specific and contingency-learning accounts.

In Experiment 2, we provided a stronger test of the last-resort account by creating a situation that would more closely resemble that examined by Bugg (2014; Bugg et al., 2008) and Blais and Bunge (2010) because, as in those experiments, a two-item set design was used. The frequency of the color-word combinations in one of the counterbalancings in the experiment is presented in Table 4 for the MC list and Table 5 for the MI list. As can be seen, as in Experiment 1, one set of words and their corresponding colors (e.g., the words RED, WHITE, GREEN, PURPLE, and their corresponding colors) was used for the context set and another set of words and their corresponding colors (e.g., the colors YELLOW, BLACK, BLUE, PINK, and their corresponding color names) was used for the transfer set. However, unlike Experiment 1, each word, both context and transfer, was only combined with two colors, the congruent color and an incongruent color (e.g., the word RED would only appear in the congruent color red and in the incongruent color white).

Unlike in Experiment 1, for the items for which a contingency can be learned (i.e., the context items), this two-item set design does not allow us to dissociate contingency learning from other processes. The reason is that, for those items, contingency and congruency proportion are perfectly confounded (e.g., a contingency between each word and its typical color can be learned not only in the MC list, but also in the MI list). Therefore, any difference obtained for those items (e.g., a PC effect) could be the result of either contingency learning or (item-specific or proactive) conflict-adaptation processes.

Although the impact of contingency learning is not directly measurable in the two-item set design, this design is well-suited to address the question of whether that process would prevent a proactive conflict-adaptation process from being implemented, as assumed by the last-resort account. The reason is that, in this design, engagement of contingency learning in the MI list would help minimize interference on the (high-contingency) incongruent context items (e.g., RED in white – see Table 5), which constitute the majority of the incongruent trials in that list (84 trials out of 132). At the same time, it would result in no cost for the incongruent transfer items (e.g., YELLOW in black; 48 trials) because no contingencies can be learned for those items (e.g., for the word YELLOW, the yellow and black colors are equally probable). In sum, contingency learning would help on 64% of the incongruent trials (the 84 high-contingency incongruent context items) whereas it would not affect performance on the remaining 36% (the 48 incongruent transfer items). As a result, this design would clearly meet the condition specified by the last-resort account that contingency learning must afford an overall minimization of interference in the MI list in order for this process to be preferred over proactive conflict adaptation. The implication is that this account would predict that no PC effect should be observed for transfer items in this type of design.

Table 4

Template for the Frequency of Color-Word Combinations in the MC List in Experiment 2

		Word							
		Context				Transfer			
Color		RED	WHITE	GREEN	PURPLE	YELLOW	BLACK	BLUE	PINK
Context	Red	21	3						
	White	3	21						
	Green			21	3				
	Purple			3	21				
Transfer	Yellow					12	12		
	Black					12	12		
	Blue							12	12
	Pink							12	12

Table 5

Template for the Frequency of Color-Word Combinations in the MI List in Experiment 2

		Word							
		Context				Transfer			
Color		RED	WHITE	GREEN	PURPLE	YELLOW	BLACK	BLUE	PINK
Context	Red	3	21						
	White	21	3						
	Green			3	21				
	Purple			21	3				
Transfer	Yellow					12	12		
	Black					12	12		
	Blue							12	12
	Pink							12	12

Method

Participants

As in Experiment 1, 48 participants took part in this experiment, a sample size that exceeds that needed to detect, with a power of .80, the effect sizes reported in Bugg's (2014) Experiments 1a and 2b. All participants were students at the University of Western Ontario (age 17–34 years) who participated for course credit. They were all native English speakers and had normal or corrected-to-normal vision.

Materials

The same color names and colors as in Experiment 1 were used. As in Experiment 1, the words RED, WHITE, GREEN, PURPLE and their corresponding colors formed one set, and the words YELLOW, BLACK, BLUE, PINK and their corresponding colors formed the other set, one (e.g., the former) serving as the context set and the other (e.g., the latter) as the transfer set. Unlike in Experiment 1, however, each word was only combined with its congruent color and one incongruent color from the same set (a two-item set design). Specifically, four subsets were created, one formed by the words RED and WHITE and their corresponding colors, one formed by the words GREEN and PURPLE and their corresponding colors, one formed by the words YELLOW and BLACK and their corresponding colors, and one formed by the words BLUE and PINK and their corresponding colors (see Tables 4 and 5).

In the MC list, each word in the context set (e.g., RED) appeared 21 times with its congruent color (e.g., red, the high-contingency color) and 3 times with the incongruent color in its subset (e.g., white, the low-contingency color). Overall, there were 84 congruent items and 12 incongruent items in the context set in the MC list, an item-specific congruency proportion of 87.5%. In the MI list, each word in the context set (e.g., RED) appeared 21 times with the incongruent color in its subset (e.g., white, the high-contingency color) and 3 times with its congruent color (e.g., red, the low-contingency color). Overall, there were 12 congruent items and 84 incongruent items in the context set in the MI list, an item-specific congruency proportion of 12.5%.

Each word in the transfer set (e.g., YELLOW) appeared 12 times in its congruent color (e.g., yellow) and 12 times in the incongruent color in its subset (e.g., black) in both lists. Overall, there were 48 congruent items and 48 incongruent items in the transfer set in both lists, an item-specific congruency proportion of 50%. Considering both context and transfer items, there were overall 132 congruent items and 60 incongruent items in the MC list (a list-wide congruency proportion of 68.75%) and 60 congruent items and 132 incongruent items in the MI list (a list-wide congruency proportion of 31.25%). Note that the congruency proportion for the MI list was slightly higher in this experiment (31.25%) than in Experiment 1 (27.08%) because, in this experiment, there were no low-contingency incongruent items among the context items in the MI list.

In this experiment as well, the contingency manipulations were parallel for words in the MC and MI lists because in each list, participants could learn a 7-to-1 contingency for context words (i.e., the (congruent or incongruent) high-contingency color was 7 times more likely than the (congruent or incongruent) low-contingency colors) and no contingency for transfer words (i.e., the (congruent) color was as likely as the (incongruent) color in the subset). Also, the absolute value of C was .90 for both lists, indicating that the absolute strength of color-word correlations was the same in the two lists.

As in Experiment 1, in both lists, the context set and the transfer set were randomly intermixed. The assignment of the two sets of words/colors to context and transfer items and the order with which the MC and MI lists were presented were counterbalanced across participants.

Procedure

The procedure was the same as in Experiment 1.

Results

As in Experiment 1, the waveforms of responses were manually inspected with CheckVocal (Protopapas, 2007) to determine the accuracy of the response and the correct placement of timing marks. Prior to the analyses, invalid trials due to technical failures and responses faster than 300 ms or slower than 2000 ms, the time limit (accounting for .8% of the data points), were discarded.

Also similar to Experiment 1, separate analyses were conducted for context and transfer items on latencies and errors. Both analyses were Proportion-Congruent analyses with Congruency (Congruent vs. Incongruent) and List Type (Mostly congruent vs. Mostly incongruent) as within-subject factors. As noted, for the context items, it was impossible to dissociate contingency learning from other processes in this experiment. Therefore, no contingency-learning analysis was conducted for the context items. The mean RTs and error rates are presented in Table 6.

Table 6

Mean RTs and Percentage Error Rates (and Corresponding 95% Confidence Intervals) for Context and Transfer Items in Experiment 2

	RTs		Error rates	
Item type	MC list	MI list	MC list	MI list
Context items				
Congruent	697 [675, 719]	756 [725, 788]	.12 [.02, .23]	.36 [-.15, .87]
Incongruent	867 [835, 898]	817 [787, 846]	5.38 [2.72, 8.05]	2.18 [1.42, 2.94]
Congruency Effect	170	61	5.46	1.82
Transfer items				
Congruent	709 [688, 731]	745 [717, 774]	.18 [0, .35]	.66 [.26, 1.06]
Incongruent	837 [809, 865]	833 [803, 863]	3.77 [2.37, 5.18]	2.27 [1.24, 3.31]
Congruency Effect	128	88	3.59	1.61

Context items (Proportion-Congruent analysis)

RTs. There was a main effect of Congruency, $F(1, 46) = 153.30$, $MSE = 4144$, $p < .001$, $\eta_p^2 = .765$, indicating faster responses to congruent than incongruent items, but no main effect of List Type, $F(1, 46) = .38$, $MSE = 2920$, $p = .543$, $\eta_p^2 = .008$. List Type interacted with Congruency, however, $F(1, 46) = 70.75$, $MSE = 2028$, $p < .001$, $\eta_p^2 = .601$, indicating that a typical PC effect was obtained. The congruency effect was larger in the MC list (170 ms) than in the MI list (61 ms). The Bayes Factor, $BF_{10} = 48990000 \pm 15.52\%$, indicated "extreme" evidence for the presence of this interaction.

Error rates. There was a main effect of Congruency, $F(1, 46) = 18.41$, $MSE = .003$, $p < .001$, $\eta_p^2 = .281$, indicating that congruent items elicited fewer errors than incongruent items. The main effect of List Type was also significant, $F(1, 46) = 6.40$, $MSE = .002$, $p = .015$, $\eta_p^2 = .120$, indicating lower error rates in the MI list than in the MC list overall. Congruency and List Type interacted as well, $F(1, 46) = 9.56$, $MSE = .001$, $p = .003$, $\eta_p^2 = .169$, with the Bayes Factor, $BF_{10} = 3.71 \pm 3.16\%$, indicating "moderate" evidence for the presence of this interaction. The interaction reflected the typical PC effect, with a larger congruency effect in the MC list (5.46%) than in the MI list (1.82%).

Transfer items (Proportion-Congruent analysis)

RTs. There was a significant main effect of Congruency, $F(1, 46) = 232.96$, $MSE = 2383$, $p < .001$, $\eta_p^2 = .832$, indicating faster responses to congruent than incongruent items, and a marginal main effect of List Type, $F(1, 46) = 3.86$, $MSE = 3191$, $p = .055$, $\eta_p^2 = .076$, indicating a tendency for faster latencies in the MC list than in the MI list overall. There was also a significant interaction between Congruency and List Type, $F(1, 46) = 26.42$, $MSE = 749$, $p < .001$, $\eta_p^2 = .360$, with the Bayes Factor, $BF_{10} = 12.17 \pm 6.46\%$, indicating "strong" evidence for the presence of this interaction. The interaction reflected the typical PC effect, i.e., a larger congruency effect in the MC list (128 ms) than in the MI list (88 ms).

Error rates. There was a significant main effect of Congruency, $F(1, 46) = 23.33$, $MSE = .001$, $p < .001$, $\eta_p^2 = .332$, with congruent items eliciting fewer errors than incongruent items, and a marginal main effect of List Type, $F(1, 46) = 3.38$, $MSE < .001$, $p = .072$, $\eta_p^2 = .067$, indicating a tendency for the MI list to produce lower error rates than the MC list overall. There was also a significant interaction between Congruency and List Type, $F(1, 46) = 10.75$, $MSE < .001$, $p = .002$, $\eta_p^2 = .186$, with the Bayes Factor, $BF_{10} = 3.91 \pm 5.53\%$, indicating "moderate" evidence for the presence of this interaction. In this case as well, the interaction reflected the typical PC effect, with a larger congruency effect in the MC list (3.59%) than in the MI list (1.61%).

Discussion

Experiment 2 replicated the results of Experiment 1 and extended them using a two-item rather than a four-item set design. As noted, although this design did not allow a dissociation of contingency learning from other processes, it is almost certain that contingency learning did take place. Consistent with this idea, a larger PC effect was obtained for the context items than for the transfer items, especially in the latencies, presumably because contingency learning contributed to that effect by facilitating responses both to the (high-contingency) congruent items in the MC list (thus increasing the congruency effect in that list) and to the (high-contingency) incongruent items in the MI list (thus reducing the congruency effect in that list).

More importantly, a PC effect was also obtained for transfer items in both the latencies and the error rates, with the Bayes Factor analyses showing good support for the presence of that effect. The PC effect is consistent with the idea that a proactive conflict-adaptation process was implemented, as assumed by the conflict-monitoring model (Botvinick et al., 2001). What this effect is not consistent with, however, is the idea that proactive conflict adaptation does not exist (the item-specific and contingency-learning accounts) or that it would not be implemented in the type of situation that we examined in Experiment 2 (the last-resort account).

Concerning the last-resort account in particular, our finding of a PC effect challenges the idea that, in a two-item set design, contingency learning should be an especially reliable option for the purpose of minimizing interference in the MI list and should thus be preferred over proactive conflict adaptation. Overall, it would appear that neither the fact that contingencies can be learned for context items in the MI list nor the nature of the two-item vs. four-item set design used are crucial determinants of whether proactive conflict adaptation would be engaged in a PC paradigm.

Experiment 3

Experiment 1 and 2 demonstrated that 1) proactive conflict adaptation is an ability that humans possess and 2) the possibility of concurrently engaging contingency learning does not prevent that ability from being engaged. In so doing, those experiments present a challenge to the item-specific, contingency-learning, and last-resort accounts of the PC effect. Our results also stand in stark contrast with those produced by previous empirical investigations that failed to obtain a PC effect on transfer items, at least in a two-item set design similar to the one that was used in our Experiment 2 (Blais & Bunge, 2010; Bugg, 2014; Bugg et al., 2008). Experiment 3 thus aimed to corroborate the robustness of the results of Experiment 2 by taking the design used in that experiment one step closer to that used in some of the previous investigations, particularly the seminal ones (Blais & Bunge, 2010; Bugg et al., 2008).

In Experiment 2, the two-item set design was implemented in a context/transfer paradigm in which eight colors (and the corresponding color names) were used, four for the context items and four for the transfer items (as in Experiment 1). In contrast, both Bugg et al. (2008) and Blais and Bunge (2010) implemented the two-item set design in a context/transfer paradigm in which four colors (and the corresponding color names) were used, two for the context items and two for the transfer items.

Although the context/transfer paradigm was later extended to larger stimulus sets (i.e., sets with eight or more colors: Bugg, 2014; see also Hutchison, 2011), it may be argued that a relatively small stimulus set may create the most appropriate situation for examining the predictions of the alternative accounts of the PC effect, particularly the last-resort account. According to this account, the possibility of engaging contingency learning to minimize interference, particularly in the MI list, is the crucial determinant of whether proactive conflict adaptation would be implemented. However, it is known that contingency learning has processing limitations: For example, word-color contingency learning in a simple color-identification task (i.e., a task where the color of neutral words, e.g., WIDE, must be identified) is impaired when limited-capacity resources are diverted from that task by use of a concurrent working-memory load manipulation (Schmidt et al., 2010; Spinelli et al., 2020).

Although the idea has not been explored yet, a potential implication is that contingency learning may also be impaired when there are numerous contingencies to be learned in a list of trials, because maintaining all those contingencies at once would be quite demanding. That is, learning that, e.g., the word RED is associated with a white color response, would be somewhat difficult in a situation with a relatively large stimulus set as our Experiment 2 because several other contingencies would have to be concurrently maintained (e.g., WHITE-red, GREEN-purple, and PURPLE-green; see Table 5). As a result of these considerable demands, contingency learning may not be a very attractive option after all and would not prevent proactive conflict adaptation from also being engaged. However, learning a RED-white contingency would be less difficult in a situation with a smaller stimulus set, for example, one in which only one other contingency would have to be concurrently maintained. As a result of reducing demands and helping to minimize the frequent interference in the MI list, contingency learning would be quite appealing in such a situation and may prevent proactive conflict adaptation from also being engaged, as assumed by the last-resort account.

Experiment 3 examined this situation by using the same two-item set design as in Experiment 2 but reducing the stimulus set from eight to four colors. To this end, two versions of the experiment were created. One version (presented to one group of participants) involved half of the colors (and their corresponding color names) used in Experiment 2 (and Experiment 1) and the other version (presented to another group of participants) involved the other half. The frequency of the color-word combinations in one of the counterbalancings of the two versions of the experiment is presented in Table 7 for the MC list and Table 8 for the MI list. Within each version, the four words used were divided into two two-item sets, one (e.g., the words RED, WHITE and their corresponding colors in Version 1) serving as the context set and the other (e.g., the words YELLOW, BLACK and their corresponding colors in Version 1) serving as the transfer set.

This design has many similarities with that used in the seminal experiments that introduced the context/transfer paradigm (Blais & Bunge, 2010; Bugg et al., 2008). Notably, those experiments failed to produce a PC effect on the transfer items, and those failures had a critical role in inspiring the development of the item-specific, contingency-learning, and last-resort accounts of the PC effect (see Blais & Bunge, 2010; Bugg, 2014; Schmidt, 2013a). By examining that instantiation of the context/transfer paradigm while avoiding potential issues with the infrequent presentation of transfer items and the arbitrary S-R mappings that might have affected those experiments, Experiment 3 can thus help establish whether the absence of a PC effect on the transfer items (and, by implication, the fact that no proactive conflict adaptation is engaged) is indeed the norm in that situation.

Table 7

Template for the Frequency of Color-Word Combinations in the MC List in Experiment 3

			Word							
			Context				Transfer			
Version	Color		RED	WHITE	GREEN	PURPLE	YELLOW	BLACK	BLUE	PINK
1	Context	Red	42	6						
		White	6	42						
	Transfer	Yellow					24	24		
		Black					24	24		
2	Context	Green			42	6				
		Purple			6	42				
	Transfer	Blue							24	24
		Pink							24	24

Table 8

Template for the Frequency of Color-Word Combinations in the MI List in Experiment 3

			Word							
			Context				Transfer			
Version	Color		RED	WHITE	GREEN	PURPLE	YELLOW	BLACK	BLUE	PINK
1	Context	Red	6	42						
		White	42	6						
	Transfer	Yellow					24	24		
		Black					24	24		
2	Context	Green			6	42				
		Purple			42	6				
	Transfer	Blue							24	24
		Pink							24	24

Method

Participants

Fifty participants took part in this experiment. No data were collected for two participants due to a technical malfunction, leaving a sample size of *N* = 48, as in Experiments 1 and 2. Again, this sample size exceeds the size needed in order to detect the effect sizes reported in Bugg's (2014) Experiments 1a and 2b with a power of .80. All participants were students at the University of Western Ontario (age 17–33 years) who participated for course credit. They were all native English speakers and had normal or corrected-to-normal vision.

Materials

The same color names and colors as in Experiment 2 were used, however, only four words/colors were presented to each participant. Two versions of the experiment were thus created. In Version 1, the words RED, WHITE and their corresponding colors formed one set and the words YELLOW, BLACK and their corresponding colors formed the other set. In Version 2, the words GREEN, PURPLE and their corresponding colors formed one set and the words BLUE, PINK and their corresponding colors formed the other set. One set (e.g., the words RED, WHITE and their corresponding colors) served as the context set and the other set (e.g., the words YELLOW, BLACK and their corresponding colors) served as the transfer set. As in Experiment 2, each word was only combined with the congruent color and one incongruent color from the same set (see Tables 7 and 8).

In the MC list, each word in the context set (e.g., RED) appeared 42 times with its congruent color (e.g., red, the high-contingency color) and 6 times with the incongruent color in its set (e.g., white, the low-contingency color). In the MI list, each word in the context set (e.g., RED) appeared 42 times with the incongruent color in its set (e.g., white, the high-contingency color) and 6 times with its congruent color (e.g., red, the low-contingency color). Each word in the transfer set (e.g., YELLOW) appeared 24 times with its congruent color (e.g., yellow) and 24 times with the incongruent color in its set (e.g., black) in both lists. The total numbers of congruent and incongruent items in each set and in each list overall, as well as the item-specific and list-wide congruency proportions, were the same as in Experiment 2. Finally, participants could learn the same contingencies in both lists as in Experiment 2. The absolute value of C was .78 for both lists, indicating that the absolute strength of color-word correlations was the same in the two lists in this experiment as well.

Procedure

The procedure was the same as in Experiments 1 and 2, with the exception that participants were only told about the colors that would appear in the version of the experiment that they were assigned to, and the 8-trial practice session only included those colors. The same number of participants (n = 24) were assigned to each version of the experiment.

Results

As in the previous experiments, the waveforms of responses were manually inspected with CheckVocal (Protopapas, 2007) to determine the accuracy of the response and the correct placement of timing marks. Prior to the analyses, invalid trials due to technical failures and responses faster than 300 ms or slower than 2000 ms, the time limit (accounting for .9% of the data points), were discarded.

The analyses were conducted as in Experiment 2. That is, separate Proportion-Congruent analyses were conducted for context and transfer items with Congruency (Congruent vs. Incongruent) and List Type (Mostly congruent vs. Mostly incongruent) as within-subject factors and latencies and error rates as dependent variables. (note 6) The mean RTs and error rates are presented in Table 9.

Table 9

Mean RTs and Percentage Error Rates (and Corresponding 95% Confidence Intervals) for Context and Transfer Items in Experiment 3

	RTs		Error rates	
Item type	MC list	MI list	MC list	MI list
<u>Context items</u>				
Congruent	662 [633, 690]	700 [664, 736]	.28 [.08, .47]	.69 [.02, 1.37]
Incongruent	809 [775, 842]	742 [711, 772]	8.02 [4.49, 11.55]	3.12 [2.03, 4.20]
Congruency Effect	147	42	7.74	2.43
<u>Transfer items</u>				
Congruent	681 [649, 713]	688 [656, 720]	.48 [.14, .82]	.18 [0, .35]
Incongruent	773 [743, 803]	748 [718, 778]	5.71 [2.95, 8.46]	3.68 [2.41, 4.96]
Congruency Effect	92	60	5.23	3.50

Context items (Proportion-Congruent analysis)

RTs. There was a main effect of Congruency, $F(1, 46) = 137.44$, $MSE = 3118$, $p < .001$, $\eta_p^2 = .745$, with faster responses to congruent than incongruent items. The main effect of List Type was also significant, $F(1, 46) = 4.75$, $MSE = 2100$, $p = .034$, $\eta_p^2 = .092$, with overall faster latencies in the MI list than in the MC list. Congruency and List Type interacted as well, $F(1, 46) = 98.77$, $MSE = 1344$, $p < .001$, $\eta_p^2 = .678$. The interaction reflected the typical PC effect, with a larger congruency effect in the MC list (147 ms) than in the MI list (42 ms). The Bayes Factor, $BF_{10} = 3516000000 \pm 2.87\%$, indicated "extreme" evidence for the presence of this interaction.

Error rates. Similar to the RT results, we found both a main effect of Congruency, $F(1, 46) = 24.56$, $MSE = .005$, $p < .001$, $\eta_p^2 = .343$, with congruent items eliciting fewer errors than incongruent items, and a main effect of List Type, $F(1, 46) = 7.45$, $MSE = .003$, $p = .009$, $\eta_p^2 = .137$, with the MI list producing a lower error rate than the MC list overall. Congruency and List Type interacted in this case as well, $F(1, 46) = 9.83$, $MSE = .003$, $p = .003$, $\eta_p^2 = .173$, with the Bayes Factor, $BF_{10} = 12.75 \pm 2.74\%$, indicating "strong" evidence for the presence of this interaction. Once again, the interaction reflected a typical PC effect, with a larger congruency effect in the MC list (7.74%) than in the MI list (2.43%).

Transfer items (Proportion-Congruent analysis)

RTs. There was a main effect of Congruency, $F(1, 46) = 134.48$, $MSE = 2052$, $p < .001$, $\eta_p^2 = .741$, with faster responses to congruent than incongruent items, but no main effect of List Type, $F(1, 46) = 2.39$, $MSE = 1589$, $p = .129$, $\eta_p^2 = .048$. Importantly, Congruency and List Type interacted, $F(1, 46) = 26.59$, $MSE = 445$, $p < .001$, $\eta_p^2 = .361$, indicating that the typical pattern of the PC effect was obtained once again, with a larger congruency effect in the MC list (92 ms) than in the MI list (60 ms). The Bayes Factor, $BF_{10} = 9.24 \pm 2.23\%$, indicated "moderate" evidence for the presence of this interaction.

Error rates. There was a main effect of Congruency, $F(1, 46) = 21.64$, $MSE = .004$, $p < .001$, $\eta_p^2 = .315$, with congruent items eliciting fewer errors than incongruent items, as well as a main effect of List Type, $F(1, 46) = 5.46$, $MSE = .001$, $p = .024$, $\eta_p^2 = .104$, with a lower error rate in the MI list than in the MC list overall. Numerically, the congruency effect was larger in the MC list (5.23%) than in the MI list (3.50%), however, the interaction between Congruency and List type did not reach significance, $F(1, 46) = 2.78$, $MSE = .001$, $p = .102$, $\eta_p^2 = .056$. The Bayes Factor, $BF_{10} = .37 \pm 7.16\%$, indicated "anecdotal" evidence for the absence of this interaction.

Discussion

The results of Experiment 3 were again straightforward. First, as in Experiments 1 and 2, a (large) PC effect emerged for the context items in both latencies and error rates. Although the two-item set design, as in Experiment 2, allows no definite conclusion regarding the source of the effect for those items, given the magnitude of the effect it is reasonable to assume a contribution of contingency learning. Second and more importantly, a PC effect also emerged for transfer items in the latencies (although not in the error rates), with the Bayes Factor analysis favoring the presence of that effect. Again, because neither item-specific conflict adaptation nor contingency learning could have produced the effect, it is, most likely, the result of a proactive conflict-adaptation process (although see footnote 4).

These results further replicate those obtained in Experiments 1 and 2 and extend them to a situation in which contingency learning was possible in the MI list (the situation examined in Experiment 1), largely helpful in minimizing the frequent interference in that list (the situation examined in Experiment 2), and, most importantly, engaging such a process would pose few demands because doing so would only require the maintenance of two (as opposed to four) contingencies in each list. Therefore, contingency learning should have been a particularly attractive option. As such, this situation would seem to be the ideal situation for the pattern predicted by the last-resort account to emerge – a PC effect on the context items, indicating contingency-learning engagement, but no PC effect on the transfer items, indicating that when contingency learning is engaged, proactive conflict adaptation is not. Such a pattern is indeed the one that Bugg et al. (2008) and Blais and Bunge (2010) reported in their initial experiments using the context/transfer paradigm. However, it is not the pattern that we found in Experiment 3 even though a very similar design was used as in those seminal experiments, i.e., a two-item set design with four colors, two used for the context items and two for the transfer items. Instead, a PC effect emerged for both context and transfer items, suggesting that, although contingency learning may have been engaged (for the context items), proactive conflict adaptation was engaged as well. In the General Discussion, we will discuss potential reasons for the discrepancy between our results and those previously reported in the literature.

Most importantly, the results of Experiment 3 converge with those of Experiments 1 and 2 in challenging the idea that the item-specific, contingency-learning, and last-resort accounts of the PC effect are the only viable accounts. Because that effect emerged in a situation in which: 1) neither item-specific conflict adaptation nor contingency learning could have produced it and 2) contingency learning would have been a reliable process to engage for the purpose of minimizing frequent interference, the effect is inconsistent with all those accounts. Instead, the effect is more in line with the original conflict-monitoring model (Botvinick et al., 2001), a model that interpreted the PC effect as the result of a proactive conflict-adaptation process that would be engaged whenever conflict would occur frequently, so as to better prepare for that conflict.

General Discussion

The Proportion-Congruent paradigm remains a valid paradigm for eliciting proactive conflict adaptation

An important function that might contribute to an efficient control system is one capable of regulating attention between task-relevant and task-irrelevant information in a preparatory fashion based on the frequency with which task-irrelevant information produces a conflict. Since the late 1970s, a central paradigm for the examination of such a proactive conflict-adaptation function has been the Proportion-Congruent (PC) paradigm. In its original implementation, this paradigm virtually always produces a PC effect, the finding that the congruency effect in Stroop-like tasks is larger in a Mostly-Congruent (MC) list, a list in which congruent items are frequent and incongruent items are infrequent, than in a Mostly-Incongruent (MI) list, a list in which incongruent items are frequent and congruent items are infrequent. The PC effect has traditionally been interpreted as a manifestation of a proactive conflict-adaptation process.

Within the widely popular conflict-monitoring model (Botvinick et al., 2001), in particular, the PC effect would reflect the fact that, in a situation in which conflict is frequent (such as in an MI list), the control system will be prepared for conflict because attention to task-relevant information will be proactively tightened. As a result, there is reduced interference from task-irrelevant information in that situation. In contrast, in a situation in which conflict is infrequent (such in an MC list), the control system will not be prepared for conflict because attention will be relaxed, causing increased interference in that situation when conflict does arise. This explanation has been extremely influential and has led to the adoption of the PC paradigm in many areas of research, from individual differences to cognitive neuroscience (for reviews, see Bugg & Crump, 2012; Chiu & Egner, 2019).

In recent years, however, the empirical grounds of this explanation have been shaken by the results reported by a few researchers using the context/transfer version of the PC paradigm, a version that allows a dissociation between proactive conflict adaptation and other processes that are typically confounded with it in the traditional version of the paradigm, i.e., item-specific conflict adaptation (a process whereby attention is regulated upon recognition of an item, rather than in a proactive, item-nonspecific fashion) and contingency learning (a process in which what is being learned is to associate stimuli with responses, rather than to regulate attention; Blais & Bunge, 2010; Bugg, 2014; Bugg et al., 2008; for reviews, see Schmidt, 2013a, 2019). As in the traditional PC paradigm, those researchers virtually always found a regular PC effect (i.e., a larger congruency effect in the MC list than in the MI list) for items for which item-specific conflict adaptation and contingency learning, in addition to proactive conflict adaptation, could have produced the PC effect (the context items). In contrast, for items for which proactive conflict adaptation was the only process that could have produced the PC effect (the transfer items), the PC effect emerged only when the design was set up so that no contingencies for the context items could be learned in the MI list. Otherwise, no PC effect emerged for the transfer items.

This null result is inconsistent with the conflict-monitoring account of the PC effect (Botvinick et al., 2001), an account that assumes that attention would be proactively enhanced in frequently conflicting situations (such as in an MI list) vs. relaxed in infrequently conflicting situations (such as in an MC list) regardless of the nature of the items involved. However, that null result could be accommodated by alternative accounts of the PC effect, accounts that assume that the PC effect reflects an item-specific conflict-adaptation process (the item-specific account: Blais et al., 2007) or a contingency-learning process (the contingency-learning account: Schmidt, 2013a) and would thus be observed only for items for which those processes function differently in the MC list vs. the MI list (i.e., the context items). Alternatively, that null result could be accommodated by hybrid accounts such as the last-resort account (Bugg, 2014) which assume that the PC effect typically reflects a contingency-learning process (and would thus be observed only for the context items) unless the situation prevents contingency learning from minimizing interference for the context items in the MI list. When that latter situation occurs, a proactive conflict-adaptation process would be used as a last resort (and the PC effect would thus be observed also for items for which proactive conflict adaptation would be the only process capable of producing such effect, i.e., the transfer items).

Importantly, although the alternative accounts of the PC effect have provided a reasonable explanation of the data pattern that has often emerged in the context/transfer PC paradigm experiments (but see Hutchison, 2011; Schmidt, 2017), those accounts imply a significant demotion of the role of proactive conflict adaptation in comparison to the role it plays in the conflict-monitoring model. Accepting any of those alternative accounts would also have widespread repercussions for past and future research in which the PC effect was or is to be taken as a marker of proactive conflict adaptation. Considerable past research framed within the conflict-monitoring model would need a reinterpretation, and, in the future, researchers would be required to adopt experimental designs that are somewhat complicated and, for some uses, impractical.

In the present research, in an effort to determine the validity of such concerns, we re-examined the empirical basis of the alternative accounts of the PC effect. To do so, we created similar situations in the color-word Stroop task as those that produced no evidence for a proactive conflict-adaptation process in past research. We also applied small methodological changes to certain aspects of that research that are not directly relevant to any account of the PC effect but are arguably suboptimal for the purposes of obtaining evidence for a proactive conflict-adaptation process. Specifically, we designed experiments in which two aspects of the prior studies were changed. First, context and transfer items were presented equally often, different from Bugg's studies (Bugg, 2014; Bugg et al., 2008) in which the transfer items were presented less frequently than the context items, a design choice that could have triggered processes of attention capture for transfer items, thus potentially making it difficult to observe a PC effect for those items. Second, vocal responses to colors were required, different from Blais and

Bunge's (2010) study in which manual responses were required, a design choice that could have imposed considerable working-memory demands for the maintenance of those arbitrary S-R mappings, thus potentially preventing the engagement of proactive conflict adaptation, a form of control that is thought to be resource-demanding (Braver, 2012).

Despite the fact that these were minor and seemingly irrelevant changes, evidence for a proactive conflict-adaptation process consistently emerged across three experiments in the form of a PC effect on the transfer items. This pattern of results is clearly inconsistent with the item-specific and contingency-learning accounts because, for transfer items, no item-specific conflict-adaptation process or contingency-learning process existed that could have produced that effect. This pattern of results is also inconsistent with the last-resort account because in all our experiments the design allowed use of contingency learning for minimizing at least a good portion of the frequent interference created for the context items in the MI list (the condition set by the last-resort account in order for contingency learning, rather than proactive conflict adaptation, to be the dominant process).

Specifically, in Experiment 1, a PC effect for the transfer items was found in the presence of a contingency-learning effect in the MI list observed independently from other processes, suggesting that although contingency learning was engaged, so was proactive conflict adaptation. In Experiment 2, a PC effect for the transfer items also emerged even though, compared to Experiment 1, engaging contingency learning was likely a more attractive option in the MI list because doing so would have resulted in a benefit on the (interfering) incongruent trials (a fact that was not true for Experiment 1 in which engaging contingency learning in the MI list would have resulted in a cost for a portion of the incongruent trials in that list). Finally, in Experiment 3, a PC effect on the transfer items emerged once again even though engaging contingency learning was further facilitated by the fact that, with a reduced stimulus set compared to Experiments 1 and 2, fewer contingencies would have had to be maintained. In sum, in contrast with the last-resort account (as well as the item-specific and contingency-learning accounts), the opportunity and ease of engaging contingency learning for minimizing the frequent interference in the MI list did not prevent a proactive conflict-adaptation process from being used in that situation.

Although the emergence of a PC effect on the transfer items in our experiments is inconsistent with the alternative accounts of the PC effect, it is easily accommodated by the conflict-monitoring account (Botvinick et al., 2001). According to that account, as noted, attention to task-relevant information for an upcoming item would be regulated based on previous experience with conflict from task-irrelevant information. Importantly, this regulation would occur in a proactive fashion, that is, before the upcoming item is recognized and independently from the identity of that individual item (and, therefore, independently from any item-specific control setting or S-R contingency learned up to that point in the task that could be used for that item). The conflict-monitoring account also makes no assumption that this proactive conflict-adaptation process would receive lower priority when the task allows other processes (such as contingency learning) to help deal with conflict. According to the conflict-monitoring account, therefore, the PC paradigm should elicit a proactive conflict-adaptation process in most situations. Further, provided that the paradigm is sensitive enough (a point to which we will return in the section "Provisional recommendations for future research using the Proportion-Congruent paradigm"), that process should result in a PC effect not only for items for which other processes such as item-specific conflict adaptation and contingency learning could produce that effect, but also for items for which proactive conflict adaptation would be the only process that could produce that effect.

The present results offer clear support for this idea, which is the idea that has guided the vast majority of research using the PC paradigm. As has commonly been assumed, our results suggest that, at least in the context of the Stroop task, the PC paradigm is a valid paradigm for eliciting proactive conflict adaptation, and the PC effect, the typical behavioral pattern observed in this paradigm, is at least a partial reflection of that process. Importantly, this conclusion would apply not only to context/transfer paradigms that allow a dissociation between the various processes that could contribute to the PC effect, but also to more traditional PC paradigms that do not allow such a dissociation. The reason is that it can be safely assumed that whatever processes are found to be engaged in the context/transfer paradigm, are likely also engaged in traditional PC paradigms even though they cannot be observed in isolation.

Based on these considerations, we disagree with the recently popularized idea that special PC paradigms (e.g., a context/transfer PC paradigm with at least six responses and no opportunity for contingency learning to be engaged in the MI list) would be needed to encourage proactive conflict adaptation (Braem et al., 2019) and the associated notion that past studies that did not use those special paradigms may not, in fact, have measured proactive conflict adaptation and would need a complete reinterpretation (see also Algom & Chajut, 2019; Schmidt, 2013a, 2019). In contrast, our results suggest that even in PC paradigms involving the Stroop task that were not specifically designed to encourage proactive conflict adaptation over other processes (and to dissociate it from other processes, as context/transfer PC paradigms do), proactive conflict adaptation is most likely to be a process underlying any PC effect that is obtained (although probably not the only process, a point that we address in the next section). These conclusions could also be extended to physiological data obtained from similar paradigms (e.g., De Pisapia & Braver, 2006). Those data likely captured processing that the researchers who collected the data typically intended to examine (e.g., neural activity associated with proactive conflict adaptation) along with processing that those researchers typically did not intend to examine (e.g., neural activity associated with item-specific conflict adaptation and contingency learning).

Multiple processes may be concurrently engaged in the Proportion-Congruent paradigm

Although the emergence of a PC effect on the transfer items in our experiments is consistent with the conflict-monitoring account of the PC effect, but not with the other accounts discussed here, it is important to acknowledge that the conflict-monitoring account is unlikely to provide a complete explanation of performance in our experiments. The reason is that this account is incapable of explaining item-specific conflict adaptation processes or S-R contingency learning (for a demonstration, see Blais et al., 2007). Yet, those processes (and potentially other processes, e.g., processes related to the frequency of individual stimuli, see, e.g., Hazeltine & Mordkoff, 2014) were likely engaged in our experiments. This conclusion is supported by both the fact that a contingency-learning effect, although small, emerged in Experiment 1 independent of other effects, and the fact that in all the experiments PC effects were larger for context items (items for which the PC effect could result from item-specific conflict adaptation and contingency learning, in addition to proactive conflict adaptation) than for transfer items (items for which the PC effect could not have resulted from those other processes). Such a pattern is most easily explained by the idea that the PC effect for context items was boosted by item-specific conflict adaptation, contingency learning, or a combination of the two. (note 7)

In fact, as a whole, the present results seem inconsistent with *all* accounts of the PC effect because each account assumes that only one process relevant to the PC effect (proactive conflict adaptation, item-specific conflict adaptation, or contingency learning) would be active at a time. In contrast, what the present results suggest, paralleling previous findings in the item-specific PC paradigm (Spinelli & Lupker, 2020a, 2020b), is that those processes are not necessarily alternatives to one another, but may co-exist, particularly in the (list-wide) PC paradigm (see also Hutchison, 2011). Although the idea that multiple processes may be concurrently engaged in the PC paradigm has received scarce attention so far (but see Schmidt, 2013a, 2019, for an account that does assume multiple processes in the PC paradigm, although not conflict-adaptation processes), a similar idea (the "multi-level learning account") has established itself in the context of the congruency-sequence paradigm (Egner, 2014). Extending this type of account to the PC paradigm would have important implications at both theoretical and methodological levels.

At the theoretical level, this idea would impose the need to develop a model that could account for multiple processes, both specific to conflict (proactive and item-specific conflict adaptation) and not specific to conflict (contingency learning), and both proactive (proactive conflict adaptation) and reactive (item-specific conflict adaptation and contingency learning). In this vein, Verguts and Notebaert (2008) have already developed a modified version of the conflict-monitoring model based on a single Hebbian learning rule which has been shown to be capable of accounting for proactive conflict adaptation, item-specific conflict adaptation, and contingency learning (see also Blais et al., 2007). However, this early modelling work based on a unified source for the three processes appears to be in contrast with: 1) more recent theorizing that makes the fundamental point that proactive and reactive control processes should be distinguished (Braver, 2012), and 2) empirical demonstrations that proactive conflict adaptation, item-specific conflict adaptation, and contingency learning show distinct properties both in terms of individual differences and in experimental manipulations (e.g., Entel et al., 2014; Gonthier et al., 2016; Spinelli et al., 2020; Spinelli & Lupker, 2020b; see also Schmidt & Besner, 2008). Further research will clearly be necessary in order to establish the most appropriate theoretical characterization of the processes involved in PC paradigms. (note 8)

An additional issue that future research will need to address is how those processes interact with each other. In this regard, although the present results are inconsistent with the last-resort account in that proactive conflict adaptation emerged even in situations in which contingency learning was a viable alternative, the present results are not necessarily inconsistent with the general idea that there may be some trade-off between proactive conflict-adaptation and contingency-learning processes. That is, although proactive conflict adaptation may never be completely abandoned, its use may be reduced in a situation in which much of the interference experienced in the task, particularly in an MI list, could be well dealt with by means of contingency learning. This idea would explain why Bugg (2014) found larger PC effects on transfer items in experiments in which contingency learning could not be used to minimize interference in the MI list than in experiments in which contingency learning could be used to that effect (although in the latter experiments, unlike in the present experiments, there was little evidence for a PC effect for the transfer items, a point which we will return to below).

Conversely, use of contingency learning and processes based on item-specific information in general, such as item-specific conflict adaptation, may be reduced in a situation in which at least some degree of proactive conflict adaptation is required, e.g., a list of trials in which some conflict is constantly experienced, requiring attention to task-relevant information to be more focused than in a list of trials in which little conflict is experienced. Consistent with this idea, Whitehead et al. (2018) found reduced word-color contingency-learning effects in a list of trials in which all the words used were incongruent words (a 100% conflicting list) compared to a list of trials in which all the words used were color-unrelated words (a 0% conflicting list), presumably because, in the former situation, attention to words was reduced in response to the high frequency of conflict in that list (although for a different explanation, see Whitehead et al., 2018; see also Levin & Tzelgov, 2016a). Similarly, Hutchison and colleagues found reduced item-specific PC effects in situations that promote proactive control, i.e., MI lists (Hutchison, 2011), individuals with a high working-memory capacity (Hutchison, 2011), and experiments in which participants are informed about the congruency of the upcoming trial (Hutchison et al., 2016). Overall, these results suggest interactivity between the processes involved in PC paradigms.

In addition to theoretical implications, a multiple-process view of the PC effect would also have important methodological implications. Specifically, that view would imply that although the PC paradigm does seem to elicit proactive conflict adaptation (for many researchers, the intended purpose of the PC paradigm), it may also elicit additional (and often, unintended) processes that would be engaged concurrently with proactive conflict adaptation, contributing to the overall data pattern. Thus, it is important to emphasize that the point made in the previous section, that special PC paradigms are not necessary in order to elicit proactive conflict adaptation, should not be intended to mean that future research should always use traditional PC paradigms, in particular, paradigms that do not distinguish context and transfer items. The reason is that the PC effect produced in those paradigms likely reflects a mixture of different processes (including, but not limited to, proactive conflict adaptation). In this sense, we do agree with Braem et al.'s (2019) recommendation that researchers interested in distinguishing proactive conflict adaptation from other processes that afford control over interference would do well to adopt paradigms, such as the context/transfer PC paradigm, that allow them to distinguish between processes. In the next section, we turn to more specific methodological recommendations that the present results suggest for future research using the PC paradigm in examinations of proactive conflict adaptation.

Provisional recommendations for future research using the Proportion-Congruent paradigm

Although the present experiments produced robust evidence for proactive conflict adaptation in the PC paradigm, previous failures to obtain such evidence in similar situations, particularly in Bugg's (Bugg, 2014; Bugg et al., 2008) and Blais and Bunge's (2010) experiments, suggest that that this process does not inevitably play a major role in situations of that sort. Examining the reasons for this discrepancy lies beyond the scope of the present research, however, the methodological differences between the present research and past studies can be used to speculate on those reasons and offer provisional recommendations for researchers interested in using the PC paradigm to examine proactive conflict adaptation in the future.

One methodological difference between the present research and past studies, particularly Bugg's (2014; Bugg et al., 2008), is that the stimuli in the context set were not displayed with a higher frequency than the stimuli in the transfer set. For example, in Experiment 2, the context color red (and the context word RED) appeared 48 times in each list overall, as did the transfer color yellow (and the transfer word YELLOW; see Tables 7 and 8). In contrast, in, e.g., Bugg et al.'s (2008) Experiment 1, each context color (and context word) appeared 96 times in each list, twice as frequently as each transfer color (and transfer word), which appeared 48 times in each list. In an article describing a consensus view among opposing researchers in the area of conflict adaptation, Braem et al. (2019) recently recommended a design choice of the latter sort because, when the words and/or colors in the context set are more frequent than the words and/or colors in the transfer set, they afford a stronger PC manipulation than they do when the context items and the transfer items are equally frequent.

What we would like to note, however, is that, in our lab, we failed to find a PC effect in picture-word interference experiments on the transfer items when the pictures used for those items had a lower presentation frequency (i.e., only once in the experiment) than the pictures used for the context items (i.e., pictures presented multiple times in the experiment). In contrast, a regular PC effect emerged on the transfer items when the stimulus components used for those items had a similar frequency as those used for the context items (i.e., both types of pictures were presented multiple times; Spinelli & Lupker, in preparation). Similarly, while Bugg (2014; Bugg et al., 2008) failed to find a PC effect on the transfer items in some of her experiments (in which the stimuli used in the context set were more frequent than the stimuli used in the transfer set), a regular PC effect emerged for the transfer items in all of the present experiments (in which the stimulus components used in the context set and transfer set were equally frequent). A particularly noteworthy contrast is that between Bugg et al.'s (2008) Experiment 1 and the present Experiment 3. Those experiments shared stimulus set size (4 colors and 4 words), type of design (the two-item set design), response modality (vocal), and had similar congruency proportions for the two lists (Bugg et al.: 66.67% for the MC list, 33.33% for the MI list; the present Experiment 3: 68.75% for the MC list, 31.25% for the MI list). However, only the present Experiment 3 produced a significant PC effect on the transfer items (for a similar contrast, compare Bugg's, 2014, Experiment 2a and the present Experiment 2).

As discussed, one possibility is that at least part of the discrepancy can be explained in terms of a process of attention capture that infrequent transfer items, being somewhat surprising among frequent context items, may trigger upon those transfer items' presentation. Because attention to the transfer items would be increased in both MC and MI lists as a result of this process, the process may override any existing differences in attentional states between the two lists, differences that those items are designed to reveal. Although this possibility still needs examination, the implication would be that whatever advantage a design with more frequent context than transfer items affords in terms of an increased congruency proportion may be counterbalanced by a disadvantage in the sensitivity to detect evidence for proactive conflict adaptation on the transfer items.

Another part of the discrepancy, as noted by Hutchison (2011) in discussing Bugg et al.'s (2008) results, may be explained by the idea that Bugg's (2014; Bugg et al., 2008) experiments might have been somewhat underpowered to detect a PC effect on the transfer items. Note that, in those experiments, List Type (MC vs. MI) was manipulated between-subjects, rather than within-subjects as in the present experiments, and the sample sizes (N = 32-36, considering young adults only) were a little smaller than those used in the present experiments (N = 48). A between-subject manipulation coupled with a relatively small sample size may not afford much power to detect an interaction involving the between-subject factor.

In order to examine this idea, we re-analyzed the results of the present experiments from the first block only, essentially making the List Type factor a between-subject factor as in Bugg's experiments. In these analyses, the PC effects in the latencies for transfer items were significant, however, they were reduced in size in all experiments (Experiment 1: $F(1, 46) = 11.83$, $MSE = 1248$, $p = .001$, $\eta_p^2 = .205$; Experiment 2:

$F(1, 46) = 7.80$, $MSE = 1353$, $p = .008$, $\eta_p^2 = .145$; Experiment 3: $F(1, 46) = 4.09$, $MSE = 1284$, $p = .049$, $\eta_p^2 = .082$). Most importantly, a power analysis based on those effect sizes suggests that sample sizes of *N* = 36, *N* = 52, and *N* = 92 would be required for Experiments 1, 2, and 3, respectively, to achieve a power of .80 to detect those effects in a similar design (i.e., a design in which List Type is manipulated between-subjects). With the exception of Experiment 1, those sample sizes were larger than the sample sizes that we used in the present experiments (*N* = 48) or those that Bugg (2014; Bugg et al., 2008) used for her experiments (*N* = 32-36).

In discussing Bugg et al.'s (2008) results, Hutchison (2011) also noted that Washington University undergraduates, the population typically sampled in Bugg's experiments (including Bugg, 2014, and Bugg et al., 2008), tend to score higher in attention control tasks than their peers at other institutions (Hutchison et al., 2013). The reason that this observation is relevant is that, as demonstrated by Kane and Engle (2003) and replicated by Hutchison (2011) in a situation controlled for item-specific conflict-adaptation and contingency-learning processes, (list-wide) PC effects tend to be smaller in individuals with higher working-memory capacity (a construct strongly associated with attention control, see Unsworth et al., 2021). The typical explanation is that, because proactive control is easier to engage for individuals with a higher working-memory capacity, those individuals would actively maintain attention focused on task-relevant information not only in contexts which support that focusing of attention, such as MI lists, but also in contexts that do not do so, such as MC lists. As a result, those individuals would tend to produce smaller PC effects (but for failures to observe this pattern using a different analytical approach, see Meier & Kane, 2013; Spinelli et al., 2021). Crucially, if it is assumed that the undergraduates sampled in Bugg's (2014; Bugg et al., 2008) experiments had, on average, higher working-memory capacity than their peers at other institutions, the implication is that those individuals would have been unlikely to produce large PC effects.

When considering this hypothesis and the idea that Bugg's (2014; Bugg et al., 2008) experiments might have been a bit underpowered, the failure to observe PC effects for transfer items in many of those experiments is not particularly surprising. Some provisional recommendations that these observations suggest for creating a better chance of obtaining a PC effect on transfer items in future research are that, first, List Type should be manipulated within-subjects or, should a between-subject manipulation be required, fairly large sample sizes should be used in such a manipulation. Second, if there is reason to believe that the individuals typically sampled in a laboratory have quite high working-memory capacity, it may be advisable to arrange for a more heterogeneous sample.

Apart from potential differences in power and participants' characteristics, the other main methodological difference between the present research and past studies, particularly Blais and Bunge's (2010) study, lies in the response modality used, or more specifically, as we suspect, in the different working-memory demands associated with vocal vs. manual responses in the context of the color-word Stroop task. Our use of vocal responses in this task likely posed few such demands because the association between colors and the pronunciation of their names is overlearned (i.e., it represents a situation with high S-R compatibility). In contrast, Blais and Bunge's (2010) use of manual (keypress) responses to multiple colors (8 to 12 in total, although only 4 colors were presented per block) likely imposed substantial demands on working memory because it tasked participants with learning and maintaining, without the assistance of feedback, several arbitrary associations between colors and the manual responses to those colors (i.e., it represents a situation with low S-R compatibility). Because working memory is thought to have a crucial role in proactive control processes (Braver, 2012; see also Kane & Engle, 2003), it seems reasonable that evidence for proactive conflict adaptation might be unlikely to emerge in a situation in which working memory is taxed.

Consistent with this idea, we recently replicated the null PC effect obtained by Blais and Bunge (2010) for transfer items in an experiment similar to the present Experiment 3, but in which manual rather than vocal responses were required (for another replication in the context of the picture-word interference task, see Bejjani et al., 2020). However, the PC effect on transfer items re-emerged when the task was switched from a manual-response color-word Stroop task to a manual-response counting Stroop task (Bush et al., 1998). This task involves responding to the number of digits in an array while ignoring the digit identity, a response that, crucially, is typically made with a spatially-compatible key (e.g., press the leftmost key if the array includes one digit, a response that is compatible with humans' tendency to associate smaller magnitudes with left-hand side responses: Dehaene et al., 1993). These results suggest that the crucial factor may not be so much response modality but rather S-R compatibility. That is, when dealing with a situation with high S-R compatibility (e.g., a vocal-response color-word Stroop task or a manual-response counting Stroop task), working memory would not be particularly taxed, allowing the engagement of proactive conflict adaptation in that type of situation. In contrast, when dealing with a situation with low S-R compatibility (e.g., a manual-response color-word Stroop task or a manual-response picture-word interference task), working memory would be taxed by the process of learning and maintaining those arbitrary S-R associations, preventing proactive conflict adaptation from playing a major role in that situation.

Another potential explanation for why proactive conflict adaptation may not play a major role in low S-R compatibility situations may have to do with the fact that, in those situations, task-relevant information does not produce a strong response tendency. For example, while in a vocal color-word Stroop task the word BLUE in an incongruent stimulus would produce a strong tendency to say "blue", in a manual color-word Stroop task that word would not produce a strong tendency to press the key designated for the blue response (for a review of supporting evidence, see MacLeod, 1991). As a result, in low S-R compatibility situations, unlike in high S-R compatibility situations, participants would experience little response conflict from incongruent stimuli, i.e., little conflict between the response associated with the task-relevant stimulus, e.g., the color, and that associated with the task-irrelevant stimulus, e.g., the word.

Importantly, in the conflict-monitoring account (Botvinick et al., 2001), response conflict is the type of conflict that is being monitored and it is used as a signal to adapt attention (but for an alternative interpretation of the type of conflict that is relevant to the conflict-monitoring model, see Entel & Tzelgov, 2018; Levin & Tzelgov, 2016b). What is possible, therefore, is that one reason why proactive conflict adaptation does not emerge in low S-R compatibility situations is that there is little response conflict in the first place that the control system needs to adapt to in those situations. Thus, even in an MI list in which incongruent stimuli are frequent, those stimuli would still not produce much response conflict when the S-R compatibility is low and would, therefore, often fail to signal to the control system that there is a need for adaptation. Although some evidence suggests that response conflict may not be the main type of conflict that humans adapt to (Shichel & Tzelgov, 2018), future research should examine this idea more fully.

In general, although we are still examining these ideas, a provisional recommendation that the extant results suggest for future research aimed at examining proactive conflict adaptation in the PC paradigm is that the paradigm should be implemented in a high S-R compatibility situation, or that the potential role of S-R compatibility should, at least, be considered when designing such experiments. For example, if high S-R compatibility cannot be achieved, researchers may consider providing participants with feedback, as feedback may allow participants to learn the relevant S-R associations more readily and, hence, rely less heavily on their working memory to maintain the instructed S-R mappings, better enabling the engagement of proactive conflict adaptation (Bejjani et al., 2020).

Overall, several details in the research design, from the relative frequency of context and transfer stimuli to the response modality, may have a role in determining whether a PC effect on transfer items would emerge. It is important to note, however, that this role and the provisional recommendations that we offered for future research using the PC paradigm are largely based on speculation on our part. Whether those speculative explanations are correct is a matter that future research should attempt to establish.

The role of timing processes in the Proportion-Congruent paradigm

Up to this point, we have considered the emergence of a PC effect on the transfer items in a context/transfer PC paradigm as relatively unambiguous evidence for the engagement of a proactive conflict-adaptation process, an interpretation that offers due consideration to the concerns raised by the alternative accounts of the PC effect, particularly the item-specific account (Blais et al., 2007) and the contingency-learning account (Schmidt & Besner, 2008). In recent years, however, the contingency-learning account has evolved into a more general learning/memory account that features other general processes that could account for the PC effect, even on contingency-controlled items such as the transfer items in a context/transfer PC paradigm (Schmidt, 2013b, 2013c; Schmidt et al., 2016).

Temporal learning is the most relevant of such processes (for the other processes, see footnote 3; see also Spinelli & Lupker, 2020c). Temporal learning refers to a process whereby, in speeded tasks, participants form temporal expectancies for the emission of a response based on previous experience in the task, expectancies that will influence the timing of subsequent responses. For example, in a list of trials in which there are many easy-to-process stimuli (e.g., pictures that are easy to name), the temporal expectancy will be generally fast. In contrast, in a list of trials in which there are many hard-to-process stimuli (e.g., pictures that are difficult to name), the temporal expectancy will be generally slow. Although there is ample evidence for timing processes of this sort taking place in, e.g., naming tasks (e.g., Lupker et al., 1997; Lupker et al., 2003), Schmidt (2013b) was the first to propose that those processes may be capable of producing a PC effect in Stroop-like tasks.

Schmidt's (2013b) account made the novel assumption that responding would speed up if processing of the current stimulus is nearly complete around the point in time at which the response is expected to be emitted (i.e., around the temporal expectancy). Thus, in a situation that creates a fast temporal expectancy, as does an MC list in which congruent (easy-to-process) stimuli are prevalent, congruent stimuli would speed up because their processing would be nearly complete around that temporal expectancy. In contrast, incongruent stimuli would not speed up because their processing would typically not be fast enough to meet the (fast) temporal expectancy in that situation. As a result, the congruency effect will be large in an MC list. Conversely, in a situation that creates a slower temporal expectancy, as does an MI list in which incongruent (hard-to-process) stimuli are prevalent, the incongruent stimuli would be the ones to speed up because their processing would be nearly complete around the typical temporal expectancy. In contrast, congruent stimuli would not speed up because their processing would typically have been completed earlier than the (slow) temporal expectancy. As a result, the congruency effect will be smaller in an MI list, producing the typical pattern of the PC effect.

Notably, because this temporal-learning process would operate on all types of items, including the transfer items in a context/transfer PC paradigm, this process could produce a PC-like effect on those items without proactive conflict adaptation being necessarily involved (for a simulation demonstrating this idea, see Schmidt, 2013b). Thus, according to this explanation, a PC effect on transfer items would still not be unambiguous evidence for proactive conflict adaptation. Even more problematic, because MC and MI lists differ intrinsically in temporal expectancies, controlling for temporal learning would prove particularly challenging in the PC paradigm. At a minimum, doing so would seem to require complex analytical and experimental procedures, procedures for which there is at present no consensus among researchers in the field (see Cohen-Shikora et al., 2019; Schmidt, 2013b, 2017, 2020; Spinelli et al., 2019; Spinelli & Lupker, in press).

A thorough discussion of the merits of the temporal-learning account of the PC effect is beyond the scope of the present research (for such discussions, see Cohen-Shikora et al., 2019; Spinelli & Lupker, 2020c, in press; Spinelli et al., 2019). Briefly, our position on this point is that although timing processes may be engaged in the PC paradigm, it is unlikely that those processes would be of the sort envisioned in Schmidt's (2013b) temporal-learning account (Spinelli & Lupker, 2020c). While Schmidt (2013b) contended that the temporal-learning mechanism he described is general and would function not only in Stroop-like tasks but in any task including easy-to-process and hard-to-process stimuli (but for a revision, see Schmidt, 2020), there are a number of tasks where this idea is contradicted by the extant data (e.g., Chateau & Lupker, 2003; Lupker et al., 1997; Lupker et al., 2003; Kinoshita & Mozer, 2006; Rastle, et al., 2003; Spinelli et al., 2019; but for a task that seems to be an exception, see Schmidt, 2013b, 2014, 2016). The reason is that, in those tasks, hard-to-process stimuli are *slower* in situations that should create slower temporal expectancies (e.g., lists including mainly or solely hard-to-process items), rather than faster as the temporal-learning account predicts. (note 9) Those types of results are more easily reconcilable with the time-criterion account (Lupker et al., 1997), an account that assumes that for both easy- and hard-to-process stimuli, response emission will be adjusted toward the temporal expectancy. More specifically, the time-criterion account would explain performance in non-Stroop tasks such as simple picture naming because, according to that account, response emission will be anticipated with a fast temporal expectancy (explaining why latencies for all stimuli tend to speed up in lists with mainly or solely easy-to-process stimuli) and delayed with a slow temporal expectancy (explaining why latencies for all stimuli tend to slow down in lists with mainly or solely hard-to-process stimuli).

In the context of Stroop-like tasks, the more standard time-criterion account would also help explain the fact that, in the typical pattern of the PC effect (a pattern that the present experiments replicated), latencies for congruent (easy-to-process) stimuli tend to be faster in the MC list (the list that should lead to an adjustment toward a fast temporal expectancy) than in the MI list (the list that should lead to an adjustment toward a slow temporal expectancy). What the time-criterion account would not be able to explain, however, is the fact that in the typical pattern of the PC effect, latencies for incongruent (hard-to-process) stimuli tend to be *faster* in the MI list (the list that should lead to an adjustment toward a slow temporal expectancy) than in the MC list (the list that should lead to an adjustment toward a fast temporal expectancy). An additional process would seem to be required to explain that pattern. Proactive conflict adaptation would be such a process, because, by causing more focused attention to task-relevant information in the MI list, interference from task-irrelevant information (on incongruent items) would be handled better in that list than in an MC list in which attention is presumed to be relaxed.

Conclusion

The process whereby information in the environment is used to prepare for conflict is a core property of the conflict-monitoring model (Botvinick et al., 2001) as well as other popular theories of cognitive control (e.g., Braver, 2012; Kane & Engle, 2003), and the PC paradigm has had, and continues to have, a central role in the examination of that proactive conflict-adaptation process. In the present research, we found evidence that reinforced that role in experiments that: 1) controlled for processes which are normally confounded with proactive conflict adaptation and 2) created situations that, in theory, may have discouraged the use of that process. Although future research would do well to consider the potential role that additional processes (i.e., in addition to proactive conflict adaptation) may have in the PC paradigm and, consequently, to adopt solutions that allow the examination of proactive conflict adaptation independently from the other processes, the PC paradigm remains a valid tool for eliciting that process.

Footnotes

1. Subsequently, unless otherwise noted, the terms “PC effect” and “PC paradigm” will be used to refer to the list-wide PC effect and the list-wide PC paradigm, respectively, i.e., an effect and a paradigm that involve a comparison between an MC list and an MI list.
2. On the other hand, in the literature, transfer items with infrequent stimulus components (relative to the stimulus components used for the context items) *have* produced a PC effect in some experiments (e.g., Bugg’s, 2014, Experiments 1a and 2b; Bugg & Chanani, 2011; Gonthier et al., 2016). Thus, the infrequent presentation of the stimulus components used for transfer items is unlikely to be the only reason why PC effects on transfer items are not always obtained.
3. The situation that we examined also controls for two other ways in which MC and MI lists often differ, color-word correlation and stimulus informativeness. Color-word correlation refers to the degree to which values on the color and word dimensions of the stimuli are correlated with one another in a list (Melara & Algom, 2003). The related notion of stimulus informativeness refers to the degree to which words allow contingency learning in a list (Schmidt, 2019). Compared to MI lists, MC lists are often higher in both color-word correlation (i.e., words can often be used to anticipate their colors) and stimulus informativeness (i.e., words can often be used to anticipate their high-contingency color). In those situations, it is possible that the larger congruency effect typically observed in the MC list relative to the MI list may result, in whole or in part, from the increased attention that word stimuli in that list would receive as a result of being highly informative (Algom & Chajut, 2019; Schmidt, 2019). In the situation examined in the present research, however, neither color-word correlation nor stimulus informativeness could have played a role in producing a PC effect because the MC and MI lists were equated on those aspects, i.e., color and words values were correlated to the same degree and words allowed contingency learning to the same degree in the two lists (see also Spinelli & Lupker, 2020c).

4. Note that the reasoning that a PC effect for transfer items would only be compatible with a proactive conflict-adaptation process ignores the argument that timing processes may also contribute to, or even produce, that PC effect (Schmidt, 2013b). We will discuss that argument in the General Discussion. To preview, we do not find that argument particularly compelling.
5. In this and the following experiments, the Proportion-Congruent analyses were repeated including the order in which the lists were presented to the participant (MC first vs. MI first) as an additional (between-subject) factor. For Experiment 1, the contingency-learning analysis was also repeated including the order in which the MI list (the list containing the items being contrasted in that analysis) was presented to participants (as the first list, i.e., the MI-first order, vs. as the second list, i.e., the MC-first order) as an additional (between-subject) factor. A discussion of the potential relevance of order effects for PC and contingency-learning effects, along with the results of those analyses, is reported in the Supplementary Materials. To preview, however, the order in which the lists were presented did not significantly modulate the effects of interest in most cases. Thus, for simplicity, we report the analyses without the Order factor in the main text.
6. The analyses were repeated including the version of the experiment (i.e., Version 1 using RED, WHITE, YELLOW, BLACK and their corresponding colors vs. Version 2 using GREEN, PURPLE, BLUE, PINK and their corresponding colors) as an additional (between-subject) factor. This factor had virtually no impact on the pattern of results.
7. This pattern, which is commonly observed in context/transfer PC paradigms, was confirmed for each experiment with an analysis involving both context and transfer items in which Item Type (Context vs. Transfer) was included as an additional within-subject factor. For latencies, a larger PC effect for context than for transfer items emerged for all experiments in the form of a three-way interaction between Item Type, List Type, and Congruency (Experiment 1: $F(1, 46) = 20.68$, $MSE = 1691$, $p < .001$, $\eta_p^2 = .306$; Experiment 2: $F(1, 46) = 24.13$, $MSE = 1176$, $p < .001$, $\eta_p^2 = .339$; Experiment 3: F(1,

46) = 47.29, *MSE* = 691, $p < .001$, $\eta_p^2 = .502$). For error rates, a significantly larger PC effect for context than for transfer items was only observed in Experiment 3 ($F(1, 46) = 6.01$, *MSE* = .001, $p = .018$, $\eta_p^2 = .113$) (for Experiment 1: $F(1, 46) = .16$, *MSE* = .001, $p = .689$, $\eta_p^2 = .003$; for Experiment 2: $F(1, 46) = 1.95$, *MSE* = .001, $p = .169$, $\eta_p^2 = .040$).

8. A reviewer of a previous version of this manuscript suggested that a model that would allow multiple processes to co-exist in PC paradigms is a horse-race model such as Logan's (1988) in which the relevant competing processes would be an algorithmic process of color identification and a memory retrieval process of S-R contingencies. For transfer items, the color-identification process would be completed earlier in an MI list than in an MC list because attention would be proactively focused on task-relevant information, resulting in decreased interference from task-irrelevant information in the former list (i.e., a PC effect). For context items, although the color-identification process would still be completed earlier in an MI list than in an MC list for the same reason, the S-R memory retrieval process might occasionally be completed even earlier (especially after an attentional lapse has occurred, inviting caution), producing even faster responses on high-contingency trials. The result would be a more pronounced PC effect for context items than for transfer items. Although this model, as with Verguts and Notebaert's (2008), would also require a distinction between the proactive and reactive processes modulating the speed of the color-identification process, we find this idea interesting and worthy of further consideration.
9. Even in the context of Stroop-like tasks, the type of tasks that the temporal-learning account was originally designed to explain (Schmidt, 2013b), some findings appear irreconcilable with that account. For example, De Jong et al. (1999) reported that in a spatial Stroop task, participants produced faster responses in a list in which the inter-stimulus interval (ITI) was short (200 ms) than in a list in which the ITI was longer (2000 ms). From a temporal-learning point of view, the implication would be that participants formed a faster temporal expectancy in the short-ITI list than

in the long-ITI list. Presumably, then, the congruency effect should be, if anything, larger in the short-ITI list in which congruent stimuli should be more likely to meet the fast temporal expectancy than incongruent stimuli, compared to the long-ITI list in which incongruent stimuli should be more likely to meet the slow temporal expectancy than congruent stimuli. De Jong et al. (1999), however, obtained the opposite pattern: a *smaller* congruency effect in the short-ITI list than in the long-ITI list (see also Jackson & Balota, 2013; Parris, 2014; for similar evidence from a different manipulation, see Bugg et al., 2015). Although the ITI manipulation appears to create differences in response speed, the crucial factor in Schmidt's (2013b) temporal-learning account of the PC effect, that account would seem to have no real way to explain the pattern that that manipulation produced. As in the case of the PC effect, that pattern seems better explained with a control process (e.g., a process whereby attention to task-relevant information is increased when the ITI is short, reducing interference in that situation; see De Jong et al., 1999).

References

Abrahamse, E., Ruitenberg, M., Duthoo, W., Sabbe, B., Morrens, M., & van Dijck, J. (2016). Conflict adaptation in schizophrenia: reviewing past and previewing future efforts. *Cognitive Neuropsychiatry*, *21*, 197-212. https://doi.org/10.1080/13546805.2016.1167679

Algom, D., & Chajut, E. (2019). Reclaiming the Stroop Effect Back From Control to Input-Driven Attention and Perception. *Frontiers in Psychology*, 10. https://doi.org/10.3389/fpsyg.2019.01683

Bejjani, C., Tan, S., & Egner, T. (2020). Performance feedback promotes proactive but not reactive adaptation of conflict-control. *Journal of Experimental Psychology: Human Perception and Performance*, *46*, 369-387. https://doi.org/10.1037/xhp0000720

Bélanger, S., Belleville, S., & Gauthier, S. (2010). Inhibition impairments in Alzheimer's disease, mild cognitive impairment and healthy aging: Effect of congruency proportion in a Stroop task. *Neuropsychologia*, *48*, 581-590. https://doi.org/10.1016/j.neuropsychologia.2009.10.021

Blais, C., & Bunge, S. (2010). Behavioral and neural evidence for item-specific performance monitoring. *Journal of Cognitive Neuroscience*, *22*, 2758-2767. https://doi.org/10.1162/jocn.2009.21365

Blais, C., Robidoux, S., Risko, E.F., & Besner, D. (2007). Item-specific adaptation and the conflict-monitoring hypothesis: A computational model. *Psychological Review*, *114*, 1076-1086. https://doi.org/10.1037/0033-295x.114.4.1076

Bonnin, C. A., Houeto, J. L., Gil, R., & Bouquet, C. A. (2010). Adjustments of conflict monitoring in Parkinson's disease. *Neuropsychology*, *24*, 542-546. https://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. https://doi.org/10.1037/0033-295X.108.3.624

Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring adaptive control in conflict tasks. *Trends in Cognitive Sciences*, *23*, 769-783. https://doi.org/10.1016/j.tics.2019.07.002

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences*, *16*, 106-113. https://doi.org/10.1016/j.tics.2011.12.010

Bugg, J. (2014). Conflict-triggered top-down control: Default mode, last resort, or no such thing? *Journal of Experimental Psychology: Learning Memory and Cognition*, *40*, 567-587. https://doi.org/10.1037/a0035032

Bugg, J. M., & Chanani, S. (2011). List-wide control is not entirely elusive: Evidence from picture–word Stroop. *Psychonomic Bulletin & Review, 18*, 930-936. https://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. https://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. https://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. https://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. https://doi.org/10.3758/MC.36.8.1484

Bugg, J. M., Diede, N. T., Cohen-Shikora, E. R., & Selmeczy, D. (2015). Expectations and experience: Dissociable bases for cognitive control?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1349-1373. https://doi.org/10.1037/xlm0000106

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L. (1998). The counting Stroop: an interference task specialized for functional neuroimaging—validation study with functional MRI. *Human Brain Mapping*, *6*, 270-282. https://doi.org/10.1002/(SICI)1097-0193(1998)6:4<270::AID-HBM6>3.0.CO;2-0

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. https://doi.org/10.1037/0096-1523.29.1.139

Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *40*, 343-367. https://doi.org/10.1037/h0080103

Chiu, Y. C., & Egner, T. (2019). Cortical and subcortical contributions to context-control learning. *Neuroscience & Biobehavioral Reviews*, *99*, 33-41. https://doi.org/10.1016/j.neubiorev.2019.01.019

Cohen-Shikora, E. R., Diede, N. T., & Bugg, J. M. (2018). The flexibility of cognitive control: Age equivalence with experience guiding the way. *Psychology and Aging*, *33*, 924-939. https://doi.org/10.1037/pag0000280

Cohen-Shikora, E. R., Suh, J., & Bugg, J. M. (2019). Assessing the temporal learning account of the list-wide proportion congruence effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*, 1703-1723. https://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. https://doi.org/10.3758/BF03193850

De Jong, R., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations: Dissociable causes of interference effects in conflict situations. *Acta Psychologica*, *101*, 379-394. https://doi.org/10.1016/S0001-6918(99)00012-8

De Pisapia, N., & Braver, T. S. (2006). A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions. *Neurocomputing*, *69*, 1322-1326. https://doi.org/10.1016/j.neucom.2005.12.100

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*, 371-396. https://doi.org/10.1037/0096-3445.122.3.371

Egner, T. (2014). Creatures of habit (and control): a multi-level learning perspective on the modulation of congruency effects. *Frontiers in Psychology*, *5*, 1247. https://doi.org/10.3389/fpsyg.2014.01247

Entel, O., & Tzelgov, J. (2018). Focusing on task conflict in the Stroop effect. *Psychological Research*, *82*, 284-295. https://doi.org/10.1007/s00426-016-0832-8

Entel, O., Tzelgov, J., & Bereby-Meyer, Y. (2014). Proportion congruency effects: Instructions may be enough. *Frontiers in Psychology*, *5*, 1108. https://doi.org/10.3389/fpsyg.2014.01108

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. https://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. https://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. https://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. https://doi.org/10.3758/s13421-016-0591-1

Hazeltine, E., & Mordkoff, J. T. (2014). Resolved but not forgotten: Stroop conflict dredges up the past. *Frontiers in Psychology*, *5*, 1327. https://doi.org/10.3389/fpsyg.2014.01327

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. https://doi.org/10.1037/a0023437

Hutchison, K. A., Balota, D. A., Neely, J. H., Cortese, M. J., Cohen-Shikora, E. R., Tse, C. S., ... & Buchanan, E. (2013). The semantic priming project. *Behavior Research Methods*, *45*, 1099-1114.

Hutchison, K. A., Bugg, J. M., Lim, Y. B., & Olsen, M. R. (2016). Congruency precues moderate item-specific proportion congruency effects. *Attention, Perception, & Psychophysics*, *78*, 1087-1103. https://doi.org/10.3758/s13414-016-1066-y

Jackson, J. D., & Balota, D. A. (2013). Age-related changes in attentional selection: Quality of task set or degradation of task set across time?. *Psychology and Aging*, *28*, 744-753. https://doi.org/10.1037/a0033159

Jacoby, L. L., Lindsay, D. S., & Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin & Review*, *10*, 638-644. https://doi.org/10.3758/BF03196526

JASP Team (2020). *JASP* (Version 0.14.1) [Computer software].

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Johnston, W. A., Hawley, K. J., Plewe, S. H., Elliott, J. M., & DeWitt, M. J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General*, *119*, 397-411. https://doi.org/10.1037/0096-3445.119.4.397

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*, 47-70. https://doi.org/10.1037/0096-3445.132.1.47

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. https://doi.org/10.3758/BF03193591

Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian data analysis for cognitive science: A practical course*. New York, NY: Cambridge University Press.

Levin, Y., & Tzelgov, J. (2016a). Contingency learning is not affected by conflict experience: Evidence from a task conflict-free, item-specific Stroop paradigm. *Acta Psychologica*, *164*, 39-45. https://doi.org/10.1016/j.actpsy.2015.12.009

Levin, Y., & Tzelgov, J. (2016b). What Klein's "semantic gradient" does and does not really show: decomposing Stroop interference into task and informational conflict components. *Frontiers in Psychology*, *7*, 249. https://doi.org/10.3389/fpsyg.2016.00249

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

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

Logan, G. D., Zbrodoff, N. J., & Williamson, J. (1984). Strategies in the color-word Stroop task. *Bulletin of the Psychonomic Society*, *22*, 135-138. https://doi.org/10.3758/BF03333784

Lowe, D. G., & Mitterer, J. O. (1982). Selective and divided attention in a Stroop task. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *36*, 684-700. https://doi.org/10.1037/h0080661

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. https://doi.org/10.1037/0278-7393.23.3.570

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. https://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. https://doi.org/10.1037/0033-2909.109.2.163

Meier, M. E., & Kane, M. J. (2013). Working memory capacity and Stroop interference: Global versus local indices of executive control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 748-759. http://dx.doi.org/10.1037/a0029200

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

Parris, B. A. (2014). Task conflict in the Stroop task: When Stroop interference decreases as Stroop facilitation increases in a low task conflict context. *Frontiers in Psychology*, *5*, 1182. https://doi.org/10.3389/fpsyg.2014.01182

Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*, 859-862. https://doi.org/10.3758/BF03192979

Rastle, K., Kinoshita, S., Lupker, S. J., & Coltheart, M. (2003). Cross-task strategic effects. *Memory & Cognition*, *31*, 867-876. https://doi.org/10.3758/BF03196441

Schmidt, J. R. (2013a). Questioning conflict adaptation: proportion congruent and Gratton effects reconsidered. *Psychonomic Bulletin & Review*, *20*, 615-630. https://doi.org/10.3758/s13423-012-0373-0

Schmidt, J. R. (2013b). Temporal learning and list-level proportion congruency: conflict adaptation or learning when to respond?. *PLoS One*, *8*, e82320. https://doi.org/10.1371/journal.pone.0082320

Schmidt, J. R. (2013c). The Parallel Episodic Processing (PEP) model: Dissociating contingency and conflict adaptation in the item-specific proportion congruent paradigm. *Acta Psychologica*, *142*, 119-126. https://doi.org/10.1016/j.actpsy.2012.11.004

Schmidt, J. R. (2014). List-level transfer effects in temporal learning: Further complications for the list-level proportion congruent effect. *Journal of Cognitive Psychology*, *26*, 373-385. https://doi.org/10.1080/20445911.2014.896367

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. https://doi.org/10.3389/fpsyg.2016.00634

Schmidt, J. R. (2017). Time-out for conflict monitoring theory: Preventing rhythmic biases eliminates the list-level proportion congruent effect. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *71*, 52-62. https://doi.org/10.1037/cep0000106

Schmidt, J. R. (2019). Evidence against conflict monitoring and adaptation: An updated review. *Psychonomic Bulletin & Review*, *26*, 753-771. https://doi.org/10.3758/s13423-018-1520-z

Schmidt, J. R. (2020). When data transformations are appropriate or even necessary: A response to Cohen-Shikora, Suh and Bugg (2019). *Timing & Time Perception*. Advance online publication. https://doi.org/10.1163/22134468-bja10019

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

Schmidt, J. R., De Houwer, J., & Besner, D. (2010). Contingency learning and unlearning in the blink of an eye: A resource dependent process. *Consciousness and Cognition*, *19*, 235-250. https://doi.org/10.1016/j.concog.2009.12.016

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

Shichel, I., & Tzelgov, J. (2018). Modulation of conflicts in the Stroop effect. *Acta Psychologica*, *189*, 93-102. https://doi.org/10.1016/j.actpsy.2017.10.007

Spinelli, G. (2021, March 26). Spinelli & Lupker JEPG. Retrieved from osf.io/rnqh6

Spinelli, G., Goldsmith, S. F., Lupker, S. J., & Morton, J. B. (2021). Bilingualism and executive attention: Evidence from studies of proactive and reactive control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication.

Spinelli, G., Krishna, K., Perry, J. R., & Lupker, S. J. (2020). Working memory load dissociates contingency learning and item-specific proportion-congruent effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*, 2007–2033. https://doi.org/10.1037/xlm0000934

Spinelli, G., & Lupker, S. J. (2020a). Dissociating adaptation to word-specific and color-specific conflict frequency in the Stroop task. In S. Denison., M. Mack, Y. Xu, & B.C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 2951-2956). Cognitive Science Society.

Spinelli, G., & Lupker, S. J. (2020b). Item-specific control of attention in the Stroop task: Contingency learning is not the whole story in the item-specific proportion-congruent effect. *Memory & Cognition*, *48,* 426-435. https://doi.org/10.3758/s13421-019-00980-y

Spinelli, G., & Lupker, S. J. (2020c). Proactive control in the Stroop task: A conflict-frequency manipulation free of item-specific, contingency-learning, and color-word correlation confounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* Advance online publication. https://doi.org/10.1037/xlm0000820

Spinelli, G., & Lupker, S. J. (in preparation). Adaptation to conflict frequency is observed for relatively familiar, but not for relatively unfamiliar stimuli, in the picture-word interference task.

Spinelli, G., & Lupker, S. J. (in press). Conflict-monitoring theory in overtime: Is temporal learning a viable explanation for the congruency sequence effect? *Journal of Experimental Psychology: Human Perception and Performance.*

Spinelli, G., Perry, J. R., Lupker, S. J. (2019). Adaptation to conflict frequency without contingency and temporal learning: Evidence from the picture-word interference task. *Journal of Experimental Psychology: Human Perception and Performance*, *45*, 995-1014. http://dx.doi.org/10.1037/xhp0000656

Spinelli, G., Perry, J. R., Lupker, S. J. (in review). The role of stimulus-response compatibility and feedback in color-word contingency learning.

Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643-662. https://doi.org/10.1037/h0054651

Surrey, C., Kretschmer-Trendowicz, A., Altgassen, M., & Fischer, R. (2019). Contextual recruitment of cognitive control in preadolescent children and young adults. *Journal of Experimental Child Psychology*, *183*, 189-207. https://doi.org/10.1016/j.jecp.2019.02.002

Unsworth, N., Miller, A. L., & Robison, M. K. (2021). Are individual differences in attention control related to working memory capacity? A latent variable mega-analysis. *Journal of Experimental Psychology: General*, *150*, 1332-1357. https://doi.org/10.1037/xge0001000

Verguts, T., & Notebaert, W. (2008). Hebbian learning of cognitive control: dealing with specific and nonspecific adaptation. *Psychological Review*, *115*, 518-525. http://dx.doi.org/10.1037/0033-295X.115.2.518

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. https://doi.org/10.1111/1469-8986.3720179

Whitehead, P. S., Brewer, G. A., Patwary, N., & Blais, C. (2018). Contingency learning is reduced for high conflict stimuli. *Acta Psychologica*, *189*, 12-18. https://doi.org/10.1016/j.actpsy.2016.09.002

Wilk, H. A., & Morton, J. B. (2012). Developmental changes in patterns of brain activity associated with moment-to-moment adjustments in control. *Neuroimage*, *63*, 475-484. https://doi.org/10.1016/j.neuroimage.2012.06.069