Original Article

QJEP

Quarterly Journal of Experimental Psychology
2022, Vol. 75(8) 1497–1513

DOI: 10.1177/17470218211055162
qjep.sagepub.com
SAGE

The dominance of item learning in the location-specific proportion congruence paradigm

Julie M Bugg, Jihyun Suh and Jackson S Colvett

Abstract

Prior research has shown that various cues are exploited to reactively adjust attention, and such adjustments depend on learning associations between cues and proportion congruence. This raises the intriguing question of what will be learned when more than one cue is available, a question that has implications for understanding which cue(s) will dominate in guiding reactive adjustments. Evidence from a picture-word Stroop task demonstrated that item learning dominated over location learning in a location-specific proportion congruence (LSPC) paradigm, a pattern that may explain the difficulty researchers have faced in replicating and reproducing the LSPC effect. One goal was to reproduce this pattern using a non-overlapping two-item set design that more closely matched prior studies, and another goal was to examine generalisability of the pattern to two other tasks. Using a prime-probe, colour–word Stroop task (Experiment 1), and a flanker task (Experiment 2), we again found clear dominance of item learning. In Experiment 3, we attempted to disrupt item learning and promote location learning by using a counting procedure that directed participants' attention to location. Once again, we found the same pattern of item dominance. In addition, in none of the experiments did we find evidence for conjunctive (location–item) learning. Collectively, the findings suggest item learning is neither design- or task-specific; rather, it is robust, reliable, and not easily disrupted. Discussion centres on factors dictating dominance of item- over location-based adjustments and implications for the broader literature on LSPC effects.

Keywords

Location-specific proportion congruence; reactive control; associative learning; attention

Received: 2 March 2021; revised: 27 July 2021; accepted: 24 September 2021

It is now well established that individuals can reactively adjust attention (extent to which they process goal-relevant versus goal-irrelevant information) based on information they have implicitly learned about associations between external cues and the likelihood of encountering distraction (i.e., conflict; for reviews, see the studies by Bugg, 2012; Bugg & Crump, 2012). This reactive (stimulus-driven) control of attention has been observed for a variety of cues including the location in which a stimulus appears (e.g., Crump & Milliken, 2009) and features of the stimulus such as the colour in which it is rendered (e.g., Bugg & Hutchison, 2013) or the picture it represents (e.g., Bugg & Dey, 2018; Bugg, Jacoby, & Chanani, 2011). This raises an intriguing theoretical question: when more than one such cue is available, which cue(s) will be attended and exploited to adaptively adjust attention on a trial-by-trial basis? In other words, what will participants learn in paradigms in which more than one association between a cue and the likelihood of distraction can be learned?

Recently Bugg, Suh, Colvett, & Lehmann (2020) investigated this question in the context of a location-specific proportion congruence (LSPC) paradigm. Before describing their findings, allow us to first introduce the LSPC paradigm. In this paradigm, participants encounter stimuli (e.g., Stroop, flanker) in two equiprobable locations on screen. Most stimuli that appear in one location (e.g., upper) are congruent, whereas most stimuli that appear in the other location (e.g., lower) are incongruent (Corballis & Gratton, 2003; Crump, Gong, & Milliken, 2006). The LSPC effect is the pattern whereby the congruency effect

Department of Psychological and Brain Sciences, Washington University in St. Louis, St. Louis, MO, USA

Corresponding author:
Julie M Bugg, Department of Psychological and Brain Sciences, Washington University in St. Louis, Campus Box 1125, St. Louis, MO 63130, USA.
Email: jbugg@wustl.edu

(i.e., difference in performance between incongruent and congruent trials) is reduced in the mostly incongruent (MI) location compared with the mostly congruent (MC) location. A key finding in the literature is that this pattern is observed for both inducer items and diagnostic items (Crump & Milliken, 2009). Inducer items (e.g., words BLUE and YELLOW appearing in colours blue or yellow) are MC in one location and MI in the other location. The presence of an LSPC effect for inducer items demonstrates that participants learned the association between each location and its proportion congruency (PC). In contrast, diagnostic items are comprised of words and colours from a separate two-item set (e.g., words RED and GREEN appearing in colours red or green) and are 50% congruent in both locations. The presence of an LSPC effect for diagnostic items demonstrates that learning about the PC of the inducer items led to abstract attentional adjustments representing location-specific control (Braem et al., 2019).

Bugg and colleagues' (2020) use of the LSPC paradigm to examine what is learned when there are multiple opportunities for learning was motivated, in part, by prior studies that had failed to replicate or reproduce Crump and Milliken's (2009) key finding. That is, these studies did not find evidence that participants used the location to guide control even though the paradigm was designed to promote exactly that (see the studies by Bugg et al., 2020; Crump, Brosowsky, & Milliken, 2017; Hutcheon & Spieler, 2017; for failures to reproduce the original finding). Strikingly, these studies not only failed to observe an LSPC effect for the diagnostic items, but they also failed to observe an LSPC effect for the inducer items.[1] As Bugg et al. (2020) noted, the lack of an LSPC effect for inducer items is uniquely informative as it implies that participants did not learn the critical association between locations and their PC. Without such learning, one should not expect to find an LSPC effect for diagnostic items (i.e., transfer of control from the inducer to the diagnostic items). Bugg and colleagues reasoned that participants instead may have learned associations between items and their PCs (i.e., item–PC learning; cf. Figure 2 of Hutcheon & Spieler, 2017, for suggestion that participants may learn other associations when items with differing PCs appear in a single location) as such learning would result in precisely the pattern that was observed in the prior studies (no LSPC effect for inducer or diagnostic items).

Multiple opportunities for learning and the importance of binning

Thus far, we have alluded to the possibility of location–PC learning and item–PC learning in LSPC paradigms. These are two of three possible associations participants could learn, with the third being associations between location–item conjunctions and PC (i.e., conjunctive learning). Next, we detail each type of learning along with the concept of binning, which we believe is central to understanding what type of learning dominates in LSPC paradigms.

A primary account of LSPC effects is the episodic retrieval account (Crump & Milliken, 2009). Drawing on prior theories (e.g., instance theory; Logan, 1988; event files theory, Hommel, 1998), this account posits that on each trial during a task, an episodic file is created that includes representations of the stimulus, response, contextual information such as the location in which the stimulus is presented, and the attentional setting that was used when processing the stimulus. To answer the question of when LSPC effects will be observed, one must understand how the episodic files (representing experiences on each trial) are organised within memory, a process we refer to as "binning" (Bugg et al., 2020). A key point is that there are multiple approaches to binning and consequently multiple levels at which learning can occur, as we will describe next.

The location–PC learning hypothesis posits that participants bin their experiences by location (Figure 1a), which LSPC research typically assumes. On this view, all experiences in the upper location would be dumped into one bin, whereas all experiences in the lower location would be dumped into a separate bin with each bin being associated with a unique PC (e.g., upper bin is MC, lower bin is MI). In keeping with the episodic retrieval account (Crump & Milliken, 2009), if such binning occurs, the attentional setting associated with a given bin (e.g., lower location bin) should be reinstated whenever a trial consistent with the bin occurs. For example, if the lower location is MI, a focused attentional setting should be retrieved whenever a stimulus appears in the lower location. Critically, if participants learn such location–PC associations, one should observe the pattern depicted in Figure 1a (upper part)—an LSPC effect for inducer and diagnostic items—as Crump and Milliken (2009) observed.

However, participants could also bin their experiences by item (Figure 1b). To be clear, by item, we are referring to a feature of the stimulus such as the colour in colour–word Stroop. The item–PC learning hypothesis posits that participants bin their experiences by item, which means that experiences with each possible item (e.g., if colours, then there are four possible colours and thus four possible items) are dumped into separate item-specific bins (e.g., every time you encounter a stimulus in blue, the corresponding episodic file gets dumped into the blue bin). In the standard design (Figure 1b, upper part), each item bin is 50% congruent because all items (e.g., colours) appear equally often as congruent and incongruent. This means that presentation of a stimulus on a given trial should trigger retrieval of the attentional setting associated with the item that appears (e.g., the colour blue). Because this setting is equivalent for all items (as all items are on average 50% congruent), there should be no difference in the

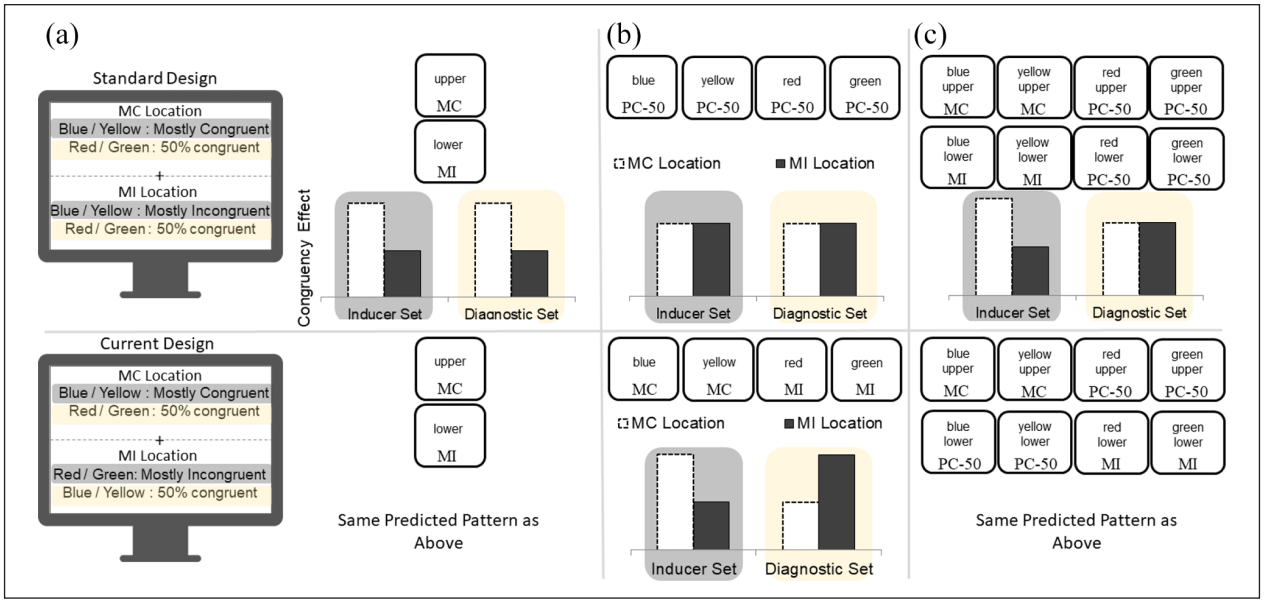

Figure 1. Three opportunities for learning in the location-specific proportion congruency paradigm: (a) location–PC learning, (b) item–PC learning, and (c) conjunctive learning.

A design in which the upper location is mostly congruent (MC), and the lower location is mostly incongruent (MI) is depicted in this figure. The colour–word Stroop task is used as the sample task in this figure, but the three types of learning, associated bins, and predicted patterns can be generalised to other four-choice tasks. The top panel illustrates the standard design of Crump and Milliken (2009) where the colours blue/yellow are serving as items in the inducer set (highlighted in grey) and the colours red/green are serving as items in the diagnostic set (highlighted in ivory). Inducer items blue/yellow are MC in the upper location but MI in the lower location. On the contrary, diagnostic items red/green are 50% congruent in both upper and lower locations. The bottom panel illustrates the current design (Bugg et al., 2020). Here, the colours that serve the role of inducer items differ across locations. Blue/yellow are MC in the upper location, whereas red/green are MI in the lower location (highlighted in grey). Similarly, the colours that serve the role of diagnostic items differ across locations—red/green and blue/yellow are 50% congruent in the upper and lower location, respectively (highlighted in ivory). Combining items within a location, the upper location is MC and the lower location is MI. The schematic illustrations in the figure depict hypothetical differences in the nature of the episodic bins that are formed and stored during the task depending on whether learning is based on (a) location, (b) item, or (c) a combination of location and item. The boxes represent the episodic bins that correspond to each type of learning. The PC of each bin refers to the overall PC collapsed across all episodic representations that would be stored in a bin during the task. Using the current design as the example, in (a), there are two bins—one for the upper location and one for the lower location and each bin is associated with the average PC of all stimuli that appear in a given location (MC and MI, respectively). In (b), there are four bins—one for each item (e.g., blue, red) and each bin is associated with the average PC of all trials on which the item is presented regardless of the location (e.g., blue is MC and red is MI, on average). In (c), there are eight bins representing all possible conjunctions of item and location with each bin (e.g., blue in the upper location) associated with the average PC of the specific conjunction (i.e., MC). Note the differences in the PC of the bins corresponding to item–PC learning in the current design (bottom panel) as compared to the standard design (top panel). Along with the schematic illustration of episodic representations, each panel shows predicted findings for the standard and current design corresponding to each learning hypothesis. The predicted difference in performance for the inducer items is highlighted in grey (corresponding to the highlighting of inducer items in grey in the design illustrations on the left part of the figure) whereas the predicted difference in performance for the diagnostic items is highlighted in ivory (corresponding to the highlighting of the diagnostic items in ivory in the design illustrations). Note the difference in the predicted pattern of results corresponding to item–PC learning for the standard and current designs. Note that the figures are idealised representations of the predicted data. PC: proportion congruence.

congruency effect between the upper and lower location for the inducer items. That is, even though blue and yellow are MC in the upper location and MI in the lower location, participants are retrieving the attentional setting associated with the item (e.g., the colour blue or yellow) meaning that the retrieved attentional setting is the same for blue (or yellow) in the upper location as it is for blue (or yellow) in the lower location (i.e., a setting corresponding to a 50% congruent PC level). In the standard design (i.e., Crump & Milliken, 2009), there also should be no difference between the two locations for the diagnostic items (for the same reason—the colours red and green on average are both 50% congruent). This is exactly the pattern that has been found in prior replication/reproduction attempts (Bugg et al., 2020; Crump et al., 2017; Hutcheon & Spieler, 2017).

Finally, participants could bin their experiences based on a conjunction of location and item (Figure 1c). The conjunctive learning hypothesis posits that both are considered during the binning process. For example, blue in the upper location and blue in the lower location would be organised into separate bins, with each bin associated with a unique PC in the case of inducer items (MC vs. MI, respectively) but the same PC in the case of diagnostic items (e.g., the bin for red items in the upper location would be 50% congruent and the same is true for the bin for red items in the lower location). Consequently, if

participants are learning the conjunctions, there should be an LSPC effect for inducer items but not diagnostic items (see Figure 1c, upper part).

Evidence for item–PC learning beyond null LSPC effects

Prior failures to reproduce/replicate the LSPC effect for inducer and diagnostic items are consistent with an explanation based on the dominance of item–PC learning (as opposed to location–PC learning) in the LSPC paradigm. However, prior to the study of Bugg et al. (2020), this explanation was theoretically plausible but based entirely on a set of null effects and open to an alternative interpretation.[2] To address these limitations, Bugg et al. (2020, Experiment 3) devised a novel variant of the LSPC paradigm. In this paradigm, all three forms of learning were again possible but importantly, evidence for item–PC learning would take the form of a specific three-way interaction and not a set of null effects.

The design is illustrated in the lower left panel of Figure 1 (where it is referred to as the current design given its use in this study). The key features are as follows: just as in the standard LSPC design, there are two locations with one being MC and the other MI (upper vs. lower location, respectively, in the figure). However, unlike the standard design, the inducer items differ across locations. For example, two items (e.g., birds and cats in Bugg et al., 2020, or the colours blue and yellow in the current Experiment 1 and Figure 1) are MC in the upper location and the two other items (e.g., dogs and fish, or red and green) are MI in the lower location. (Note that the remaining examples will refer to colours as in the figure and the current Experiment 1, and not animals as in Bugg et al., 2020.) The "diagnostic"[3] items also differ across locations, but the assignment of colours to locations is switched relative to inducer items. Using the same examples, blue and yellow would appear as 50% congruent items in the lower location and red and green would appear as 50% congruent items in the upper location. The predictions for this new design as a function of the type of learning that (hypothetically) dominates are illustrated in the bottom part of Figure 1.

If the location–PC learning hypothesis is supported, there should be an LSPC effect for both inducer and diagnostic items with a larger congruency effect in the MC (upper) location than the MI (lower) location. This prediction mirrors that of the standard design because as can be seen in Figure 1a, the bins are the same. Thus, stimuli that appear in the upper location (regardless of whether they are inducer or diagnostic items) should reinstate a more relaxed attentional setting than stimuli in the lower location, resulting in a larger congruency effect. Critically, and in contrast to the standard design, the item–PC learning hypothesis predicts that an LSPC effect should be found for inducer items, but a reversed LSPC effect should be found for diagnostic items (see Figure 1b, lower part). This prediction falls directly out of the hypothesis that participants bin based on items. To elaborate, in this design, two of the items (blue and yellow) on average are MC and two on average (green and red) are MI (as depicted by the bins in Figure 1b). Thus, whenever blue or yellow items appear, they reinstate the attentional setting associated with the blue or yellow bin, respectively, resulting in a more relaxed state relative to when green or red items appear. Looking at the predicted pattern for the inducer set (highlighted in grey in the figure), the congruency effect is thus larger for the inducer items blue/yellow in the MC [upper] location than the inducer items green/red in the MI [lower] location. As for the diagnostic set (highlighted in ivory in the figure), a reversed LSPC effect is expected because in this design, green/red (which are MI on average) are now in the MC [upper] location and blue/yellow (which are MC on average) are now in the MI [lower] location and, therefore, the congruency effect should now be smaller for the MC location than the MI location. Finally, the conjunctive learning hypothesis predicts an LSPC effect for inducer items (larger congruency effect in MC [upper] than MI [lower] location) but no difference in the congruency effect for diagnostic items (a null LSPC effect), just as in the standard design. This is because the contrast between inducer items in the upper versus lower location represents a contrast between items that are MC and items that are MI, respectively, whereas the contrast between diagnostic items represents a contrast between a set of items that is 50% congruent in the upper location and a set of items that is 50% congruent in the lower location.

Bugg et al. (2020, Experiment 3) implemented this design for the first time using a picture–word Stroop task (where four to-be-named pictures of animals served the role as the four colours in Figure 1). Consistent with the item–PC learning hypothesis and inconsistent with the location–PC learning hypothesis, the anticipated three-way interaction pattern was found showing an LSPC effect for the inducer items and a reversed LSPC effect for the diagnostic items. This pattern is also inconsistent with the conjunctive learning hypothesis. This led to the conclusion that participants may be inclined to learn about and exploit associations between items and their PC and not locations and their PC (or the conjunction) when both cues are available to guide control. That is, item–learning may dominate and this may, in part, explain why prior studies have had difficulty replicating/reproducing the LSPC effect.

Although the three-way interaction was robust, it is important to reproduce the findings of Bugg et al. (2020) as their study differed in some potentially important ways from extant LSPC studies. First, they employed an overlapping sets design wherein each item (e.g., bird) was paired with all possible distractor words (BIRD, CAT, DOG, FISH), a design that promotes item-level control (see e.g., Bugg & Dey, 2018; Bugg & Hutchison, 2013;

Bugg et al., 2011; Suh & Bugg, 2021). In contrast to this design, Crump and Milliken (2009) and the prior studies that failed to replicate/reproduce their findings all employed a non-overlapping two-item sets design wherein each item (e.g., blue) was only paired with words within the same item set (e.g., blue and yellow appeared only with the words BLUE and YELLOW, but not RED and GREEN). In this design, participants can learn the distractor-response contingencies for a given location resulting in a larger congruency effect for the MC location as compared with the MI location (i.e., LSPC effect for the inducer items). That is, given the use of non-overlapping two-item sets, participants can learn to produce the congruent response when an inducer word is shown in the upper, MC location (e.g., say "blue" when you encounter BLUE) and the incongruent response when an inducer word is shown in the lower, MI location (e.g., say "yellow" when you encounter BLUE). Of course, responses are equally contingent on the word across locations in the diagnostic set (e.g., the word RED appears equally often with red and green) and accordingly, use of contingency learning should not produce an LSPC effect for diagnostic items. Although there is no clear evidence for this pattern (an LSPC effect for the inducer items but no LSPC effect for the diagnostic items, which supports the conjunctive learning hypothesis) to date either in Crump and Milliken (2009) or the subsequent replication/reproduction attempts that used two-item sets (Crump et al., 2017; Hutcheon & Spieler, 2017), Schmidt and Lemercier (2019) did find evidence for conjunctive learning (which they referred to as compound contingency learning) using a variant of the context-specific PC design in which font was the contextual cue and high and low contingency items were differentially distributed across the two font contexts. The current study affords us the opportunity to determine whether evidence for conjunctive learning will be more apparent (relative to Bugg et al., 2020) when the non-overlapping two-item sets design is used. For example, participants might be more inclined to predict responses on inducer trials based on the word/location conjunctions that yield high contingency responses (congruent response in MC location and incongruent response in MI location) in this design than in the overlapping sets design used by Bugg and colleagues to test the item–PC learning hypothesis.

Another implication of our decision to use the non-overlapping two-item sets design is that when we refer to item–PC learning (e.g., Figure 1), this now includes the possibility that participants are learning contingencies at the item level, and predicting responses based on the overall contingencies associated with a given word (collapsed across location). In the study by Bugg et al. (2020), the overlapping sets design was used such that item–PC learning referred to the learning and use of item-level control (i.e., modulating attention on an item-by-item basis, rather than predicting contingent responses).

A second unique aspect of the design of Bugg et al. (2020) was that they employed a picture–word Stroop task not previously used in the LSPC literature. The task did yield additional findings in the study by Bugg et al. that reproduced patterns observed previously in tasks more commonly used in this literature (e.g., colour–word Stroop; flanker). For instance, in their attempt to reproduce the findings of Crump and Milliken (2009), Bugg et al. found an LSPC effect for neither inducer nor diagnostic items. In addition, in an inducer–item only design, Bugg et al. did find an LSPC effect like Hutcheon and Spieler (2017) who used such a design in the colour–word Stroop task. Nonetheless, the question remains whether the evidence for the dominance of item-level learning in the LSPC paradigm is task-specific. This could be the case because, relative to stimuli used in other conflict tasks like Stroop or flanker, the picture stimuli in the picture–word Stroop task are more variable (there are multiple exemplars of each animal) and arguably more distinctive, both of which might attract attention to the items and disproportionately bias participants to learn about the items rather than the locations. In short, either or both deviations from the typical design may have contributed to the dominance of item–PC learning over location–PC learning in the study by Bugg et al. (2020).

This study

This study comprised three experiments. Experiments 1 and 2 were attempts to reproduce the three-way interaction observed by Bugg et al. (2020, Experiment 3), which demonstrates the dominance of item–PC learning over location–PC learning (i.e., use of item and not location cues to guide reactive adjustments trial-by-trial). Both experiments used two-item, non-overlapping sets of stimuli just as in Crump and Milliken (2009; see also Crump et al., 2017; Hutcheon & Spieler, 2017). The key difference across experiments was the task type. The prime-probe, colour–word Stroop task was employed in Experiment 1. This task was chosen because the original pattern (LSPC effects for inducer and diagnostic items) was observed with this task (Crump & Milliken, 2009), although it was later not reproduced (Crump et al., 2017, Experiment 2). An arrow flanker task was employed in Experiment 2. This task was chosen for two primary reasons. First, in comparison to the distinctiveness of picture–word Stroop stimuli (items like pictures of dogs and fish), the items in an arrow flanker task are arguably less distinct from each other (see Bugg, 2015). Second, flanker tasks have been used frequently in LSPC paradigms also by Crump et al. (2017, Experiment 4) who partially reproduced the Crump and Milliken (2009) study by finding an LSPC effect for diagnostic but not inducer items in a letter flanker task. In addition, a flanker task was selected because prior evidence suggests that spatial conflict may be a pre-requisite for other indicators of location–PC learning (Pickel, Pratt, & Weidler, 2019).

If item–PC learning dominates despite these methodological changes from Bugg et al. (2020), then the same three-way interaction should again be observed showing an LSPC effect for inducer items and a reversed LSPC effect for diagnostic items. To preview the results, this is precisely what we observed in Experiments 1 and 2. Thus, Experiment 3 was designed to potentially disrupt the dominance of item–PC learning. The experiment was identical to Experiment 2 except that participants were asked to count stimuli that were presented in one of the two locations. Prior research has shown that counting manipulations successfully shift attention to otherwise unattended (or less attended) cues (Brosowsky & Crump, 2020; Crump, Vaquero, & Milliken, 2008). If the dominance of item–PC learning can be overcome by drawing participants' attention to location, we expect that the findings will support either the location–PC learning hypothesis or the conjunctive learning hypothesis (which also relies on attention to location but does not fully ignore the items being learned).

Experiment 1

The purpose of Experiment 1 was to reproduce the findings of Bugg et al. (2020, Experiment 3) demonstrating the dominance of item–PC learning in an LSPC paradigm using a two-item, non-overlapping sets design and a prime-probe, colour–word Stroop task.

Method

Participants. A total 94 undergraduates (78 women; age $M = 19.63$, $SD = 1.23$) from Washington University in Saint Louis participated in this study. We used simulation-based power analysis (Lakens & Caldwell, 2019) to calculate the sample size. Based on the results of Bugg et al. (2020) Experiment 3, the desired sample size to detect the three-way interaction effect ($\eta_p^2 = .19$) with the power of 0.8 and alpha level of 0.05 was 42 for the current and subsequent experiments. We targeted 96 participants because we anticipated that contributions of item–PC learning may be smaller in the context of the colour–word Stroop task than the picture–word Stroop task given the differences noted above (i.e., less variability, less distinctive items), and thus we collected approximately one-third more data than in the study by Bugg et al. (2020, Experiment 3).[4] The final sample included data from 94 participants because in-person data collection was halted due to COVID-19. All participants earned class credit for participation. All participants were native English speakers and had normal or corrected-to-normal vision and colour vision.

Stimuli. The stimuli for this experiment were adapted from previous studies using a colour–word prime-probe task (e.g., Crump et al., 2006; Crump & Milliken, 2009). All stimuli were presented on a black background. There were four colour–word primes (BLUE, RED, GREEN, YELLOW) presented in white and 38-point Arial font. The target probe was a filled rectangle that was presented in blue, red, green, or yellow. The colour patches were 6×2 cm in size.

Design. As in the study by Bugg et al. (2020, Experiment 3), the design of Experiment 1 was a $2 \times 2 \times 2$ within-subjects design with factors of trial type (congruent or incongruent), location PC (MC or MI), and set type (inducer or diagnostic). Trials were congruent when the identity of the prime and probe matched (e.g., word YELLOW and yellow rectangle) and incongruent (e.g., word BLUE and yellow rectangle) when the identity of the prime and probe mismatched. Location PC refers to whether a given location was mostly comprised of congruent trials (i.e., MC location) or mostly comprised of incongruent trials (i.e., MI location).

There were two sets of items (i.e., set types). Inducer items were either MC or MI depending on the location, whereas the diagnostic items were 50% congruent in both locations. Critically, we used a non-overlapping two-item sets design where word/colour patches were paired and did not overlap (cf. standard design of Crump & Milliken, 2009) meaning that a given set of words (e.g., BLUE and YELLOW) only appeared with the corresponding colour patches (blue and yellow) but never with the colour patches in the opposite set (red and green; see Table 1 for pairings). Unlike the standard design but consistent with Bugg et al., 2020 (Experiment 3), the specific word/colour patches that served as the inducer items differed across locations. For example, if blue and yellow served as MC inducer items in the upper location, then red and green served as MI inducer items in the lower location.[5] Similarly, the specific word/colour patches that served as the diagnostic items also differed across locations, and these items appeared in the location opposite the corresponding inducer items. For example, blue and yellow served as 50% congruent diagnostic items in the lower location, whereas red and green served as 50% congruent diagnostic items in the upper location (see Figure 1). This was counterbalanced across participants, as was the assignment of PC (MC or MI) to location (upper or lower).

Note that combining inducer and diagnostic sets within a location resulted in locations that were 70% (MC location) or 30% congruent (MI location). Collapsing across all instances regardless of location, one set of items (e.g., blue and yellow in above example) was 70% congruent (75% congruent in inducer set and 50% congruent in diagnostic set), and the other set of items (e.g., red and green in above example) was 30% congruent (25% congruent in inducer set and 50% congruent in diagnostic set).

Collapsing across locations, 50% of trials were congruent and 50% were incongruent, and 50% of trials were

Table 1. Frequencies of stimulus presentation in Experiment 1.

Location (PC)	Set type	Colour	Word			
			BLUE	YELLOW	RED	GREEN
Upper (MC)	Inducer	Blue	72	24	0	0
		Yellow	24	72	0	0
	Diagnostic	Red	0	0	12	12
		Green	0	0	12	12
Lower (MI)	Inducer	Red	0	0	24	72
		Green	0	0	72	24
	Diagnostic	Blue	12	12	0	0
		Yellow	12	12	0	0

PC: proportion congruence; MC: mostly congruent; MI: mostly incongruent.
Numbers represent frequencies of stimulus presentation in Experiment 1. Underlined frequencies represent congruent trials. In this example, in the MC location, blue and yellow represent the inducer set and red and green represent the diagnostic set. In the MI location, red and green represent the inducer set, whereas blue and yellow represent the diagnostic set. This was counterbalanced across participants, as was assignment of location to PC.

presented in each location, meaning participants could not predict congruency or location on a trial-by-trial basis. Table 1 presents the stimulus frequencies.

Procedure. E-prime 2.0 software was used to present stimuli on a 17-inch liquid crystal display (LCD) monitor (Psychological Software Tools, Pittsburgh, PA), and vocal responses were detected by a microphone connected to a voice key using the PST serial response box (Psychological Software Tools, Pittsburgh, PA).

Participants consented to participate and were then given instructions for the task. Participants were instructed to name aloud the colour of the target patch rather than the distractor word as quickly as possible without sacrificing accuracy. Following the study by Crump and colleagues (2006; Crump & Milliken, 2009), each trial began with a fixation cross that appeared centrally for 1,000 ms, followed by a blank screen for 250 ms. Next, the irrelevant prime word was presented centrally for 100 ms. Immediately thereafter, the target probe was presented in either the upper or lower location. The location was the upper or lower side of the screen (6.5 cm from the centre of the display) and was centrally aligned on the horizontal dimension. The probe remained on screen until a vocal response was detected by a voice key. After the voice key was triggered, the experimenter coded the participant's response using a keyboard (i.e., indicated what colour word was spoken). Sounds that unintentionally triggered the voice key (e.g., "um," a cough) or vocal responses that were otherwise imperceptible or unintelligible were coded as "scratch trials" by the experimenter and excluded from analysis. After experimenter coding, the next trial began. Reaction time (RT) (ms) and error rate were recorded.

Participants completed 16 practice trials. The PCs of the inducer and diagnostic sets in each location were consistent with that of the experimental lists. Practice trials were not included in data analyses. After completion of the practice trials, participants completed 480 experimental trials that were randomly presented without replacement according to the frequencies listed in Table 1. A brief break was provided after every 120 trials, resulting in four blocks. At the end of the experiment, participants were debriefed. The experiment lasted approximately 30 min.

Results

Responses slower than 3,000 ms or faster than 200 ms were excluded (less than 1% of all trials) as in our prior vocal Stroop research (Bugg et al., 2011, 2020). The mean RT and error rates are summarised in Table 2. For RT and error rate, a 2 (Location PC: MC vs. MI) × 2 (Trial Type: Congruent vs. Incongruent) × 2 (Set Type: Inducer vs. Diagnostic) repeated-measures analysis of variance (ANOVA) was conducted.[6]

Reaction time. Only correct responses were included in the RT analysis. The main effect of trial type, $F(1,93)=320.42$, $p<.001$, $\eta_p^2=.78$, $BF_{01}=0.00$, was significant showing that RT was faster for congruent ($M=545$ ms) compared with incongruent ($M=604$ ms) trials. In addition, a significant main effect of set type, $F(1,93)=21.12$, $p<.001$, $\eta_p^2=.19$, $BF_{01}=1.43$, revealed that the overall RT was faster for inducer ($M=571$ ms) than diagnostic sets ($M=578$ ms). The main effect of location PC, $F<1$, $BF_{01}=12.00$, and Location PC × Trial Type interaction, $F(1,93)=1.84$, $p=.18$, $\eta_p^2=.02$, $BF_{01}=5.49$, were not significant. However, most importantly, the three-way interaction was significant, $F(1,93)=21.47$, $p<.001$, $\eta_p^2=.19$, $BF_{01}=0.01$ (see Figure 2).

To better understand the three-way interaction, we decomposed it based on set type (inducer or diagnostic set). When decomposed by the set type, the congruency effect was significantly larger in the MC location ($M=68$ ms) than in the MI location ($M=48$ ms), $F(1,93)=29.25$, $p<.001$, $\eta_p^2=.24$, $BF_{01}=0.01$, indicating a standard LSPC effect

Table 2. Mean reaction time, error rate, and congruency effects in Experiment 1.

Set type	Location (PC)	DV	Trial type		Congruency effect
			Congruent	Incongruent	
Inducer	Upper (MC)	RT	537 (20)	605 (29)	68 (25)
		Error rate	.006 (.010)	.016 (.022)	.010 (.024)
	Lower (MI)	RT	548 (25)	595 (24)	48 (28)
		Error rate	.007 (.015)	.010 (.009)	.003 (.016)
Diagnostic	Upper (MC)	RT	552 (33)	605 (30)	54 (36)
		Error rate	.007 (.016)	.010 (.023)	.003 (.025)
	Lower (MI)	RT	544 (27)	610 (37)	66 (33)
		Error rate	.004 (.017)	.016 (.031)	.012 (.029)

PC: proportion congruence; DV = dependent variable; MC: mostly congruent; RT: reaction time; MI: mostly incongruent.
Values in parentheses indicate standard deviation of the mean. Note that items in the diagnostic set were 50% congruent in each location. For half of the participants, the upper location was MC and the lower location was MI (as depicted here); for the other half, this was reversed.

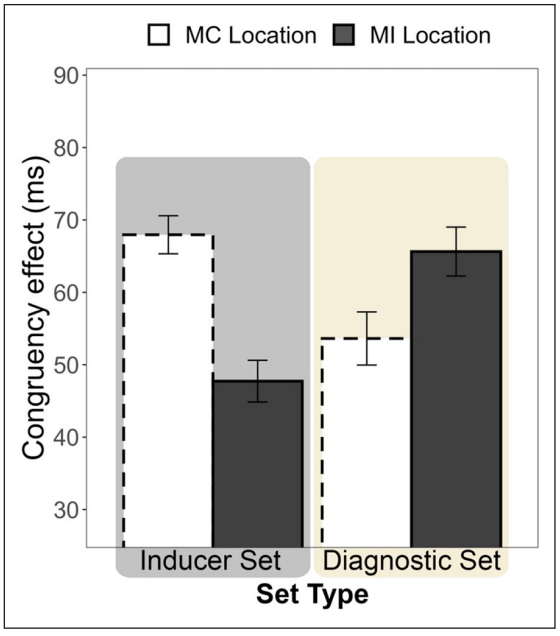

Figure 2. Mean congruency (Stroop) effects as a function of location-specific proportion congruence for the inducer and diagnostic sets in Experiment 1.
The error bars represent one within-subject standard error. The inducer set is highlighted in grey and the diagnostic set is highlighted in ivory, corresponding to Figure 1. Note that the same items (e.g., blue/yellow) played the role of the mostly congruent inducer items in the mostly congruent location (left-most bar) and the role of the 50% congruent diagnostic items in the mostly incongruent location (right-most bar). Conversely, a different set of items (e.g., red/green) played the role of the mostly incongruent inducer items in the mostly incongruent location (left middle bar) and the role of the 50% congruent diagnostic items in the mostly congruent location (right middle bar).

(i.e., Location × Trial Type interaction) for the inducer set. For the diagnostic set, however, the congruency effect was significantly larger in the MI location (M=66 ms) than the MC location (M=54 ms), $F(1,93)=5.05$, $p=.03$, $\eta_p^2=.05$, $BF_{01}=1.54$, indicating a reversal of the LSPC effect.

Error rate. The error rate was larger for incongruent (M=0.013) compared with congruent (M=0.006) trials, $F(1,93)=17.25$, $p<.001$, $\eta_p^2=.16$, $BF_{01}=0.00$. The main effects of Location PC, $F<1$, $BF_{01}=12.30$, and set type $F<1$, $BF_{01}=12.10$, were not significant. None of the two-way interactions was significant, Set Type × Location PC: $F(1,93)=2.18$, $p=.14$, $\eta_p^2=.02$, $BF_{01}=2.70$, Set Type × Trial Type: $F<1$, $BF_{01}=8.05$, Location PC × Trial Type: $F<1$, $BF_{01}=7.96$. However, the three-way interaction was significant, $F(1,93)=11.21$, $p=.001$, $\eta_p^2=.11$, $BF_{01}=0.14$.

Similar to the RT analysis, we decomposed the three-way interaction based on set type. For the inducer set, the congruency effect was larger in the MC location (M=0.010) compared with the MI location (M=0.003), $F(1,93)=7.57$, $p=.01$, $\eta_p^2=.08$, $BF_{01}=0.57$. For the diagnostic set, the congruency effect was larger in the MI location (M=0.012) than the MC location (M=0.003), $F(1,93)=5.53$, $p=.02$, $\eta_p^2=.06$, $BF_{01}=1.03$.

Discussion

The key finding from Experiment 1 was the three-way interaction showing a standard LSPC effect for the inducer items and a reversed LSPC effect for the diagnostic items, which uniquely supports the item–PC learning hypothesis. This finding provides initial evidence that the results of Bugg et al. (2020, Experiment 3) are neither specific to the overlapping sets design or the picture–word Stroop task.

Experiment 2

The purpose of Experiment 2 was to reproduce the findings of Bugg et al. (2020, Experiment 3) and the present Experiment 1 by demonstrating the dominance of item–PC learning in an LSPC paradigm using the same design as Experiment 1, but using an arrow flanker task.

Table 3. Frequencies of stimulus presentation in Experiments 2 and 3.

Location (PC)	Set type	Target	Flanker			
			<	>	^	v
Upper (MC)	Inducer	<	72	24	0	0
		>	24	72	0	0
	Diagnostic	^	0	0	12	12
		v	0	0	12	12
Lower (MI)	Inducer	^	0	0	24	72
		v	0	0	72	24
	Diagnostic	<	12	12	0	0
		>	12	12	0	0

PC: proportion congruence; MC: mostly congruent; MI: mostly incongruent.
Numbers represent frequencies of stimulus presentation in Experiments 2 and 3. Underlined frequencies represent congruent trials. In this example, in the MC location left and right represent the inducer set while up and down represent the diagnostic set. In the MI location, up and down represent the inducer set while left and right represent the diagnostic set. This was counterbalanced across participants, as was assignment of location to PC.

Method

Participants. A total 103 undergraduates[7] (67 women; age $M = 19.28$, $SD = 1.18$) from Washington University in Saint Louis participated in this study. All participants earned class credit for participation. All participants reported that they had normal or corrected-to-normal vision.

Stimuli. The stimuli were strings of seven arrows in a horizontal line that were presented in white on a black background. Arrows could point up, down, left, or right. The three outermost arrows surrounding the central arrow on each side were the flanker arrows and the central arrow was the target arrow.

Design. The design was equivalent to Experiment 1. Here, trials were congruent when the centre arrow matched the flanker arrows (e.g., <<<<<<<) and incongruent when the centre arrows mismatched the flanker arrows (e.g., <<<>><<<). Table 3 details the stimulus pairings and the stimulus frequencies.

Procedure. In this and the subsequent experiment, PsychoPy software (Peirce et al., 2019) was used to present stimuli and record RT and error rate. On each trial, a fixation cross appeared centrally for 1,000 ms, followed by a blank screen for 250 ms. Next, the flanker stimulus appeared in either the upper or lower location for 3,000 ms or until the participant responded. The upper and lower locations were presented in upper or lower halves of the screen (i.e., the exact distance from fixation varied based on participants' display size), and they were centrally aligned on the horizontal dimension. Participants responded to the identity of the central arrow using their right index finger on the up, down, left, and right arrow keys on their keyboard. Finally, there was a 200 ms interstimulus interval before the next trial began.

The experiment was hosted online on Pavlovia and participants used their own computer to participate. After providing consent, participants were instructed to respond to the identity of the centre arrow as quickly and accurately as possible. Participants first completed a 16-trial practice phase with feedback. The PCs of the inducer and diagnostic sets in each location were consistent with that of the experimental lists. Practice trials were not included in data analyses. After completion of the practice trials, participants were presented with 480 experimental trials that were randomly presented without replacement according to the frequencies listed in Table 2. A brief break was provided after every 120 trials, resulting in four blocks. At the end of the experiment, participants were debriefed. The experiment lasted approximately 30 min.

Results

Two participants were excluded from analysis due to overall slow RT and an excessive number of errors (beyond 2.5 standard deviation from the mean of all participants[8]). Therefore, the final analysis included 101 participants. RTs slower than 2,000 ms or faster than 200 ms were excluded resulting in exclusion of less than 1% of all trials (Bugg, 2015; Weidler & Bugg, 2016). The mean RT and error rates are summarised in Table 4. For both RT and error rate, a 2 (Location PC: MC vs. MI) × 2 (Trial Type: Congruent vs. Incongruent) × 2 (Set Type: Inducer vs. Diagnostic) repeated-measures ANOVA was conducted.

Reaction time. Only correct responses were included in the RT analysis. The main effects of trial type, $F(1,100) = 1{,}079.78$, $p < .001$, $\eta_p^2 = .92$, BF01 = 0.00, and set type, $F(1,100) = 4.08$, $p = .05$, $\eta_p^2 = .04$, $BF_{01} = 11.81$, were significant showing slower RT for incongruent ($M = 805$ ms) than congruent ($M = 614$ ms) trials and slower RT for the diagnostic ($M = 711$ ms) compared with the inducer set ($M = 708$ ms). The main effect of Location PC was not significant, $F < 1$, $BF_{01} = 12.24$. The Location PC × Trial Type interaction was significant, $F(1,100) = 30.36$, $p < .001$, $\eta_p^2 = .23$, $BF_{01} = 0.25$, indicating that the congruency effect was larger in the MC location ($M = 202$ ms) compared with the MI location ($M = 179$ ms) when the inducer and diagnostic sets were collapsed together.[9] The Set Type × Location PC interaction, $F(1,100) = 3.57$, $p = .06$, $\eta_p^2 = .03$, $BF_{01} = 4.83$, and Set Type × Trial Type interaction, $F < 1$, $BF_{01} = 8.17$, were not significant. Most importantly, the three-way interaction was significant, $F(1,100) = 21.22$, $p < .001$, $\eta_p^2 = .18$, $BF_{01} = 0.00$ (see Figure 3).

The three-way interaction was decomposed based on set type and item set. For the inducer set, the congruency effect

Table 4. Mean reaction time, error rate, and congruency effects in Experiment 2 (no counting) and Experiment 3 (counting).

Counting	Set type	Location (PC)	DV	Trial type		Congruency effect
				Congruent	Incongruent	
No	Inducer	Upper (MC)	RT	597 (50)	826 (72)	228 (82)
			Error rate	.005 (.037)	.103 (.081)	.098 (.060)
		Lower (MI)	RT	630 (49)	779 (53)	149 (70)
			Error rate	.009 (.036)	.045 (.036)	.036 (.059)
	Diagnostic	Upper (MC)	RT	617 (52)	793 (78)	176 (91)
			Error rate	.007 (.037)	.054 (.055)	.047 (.069)
		Lower (MI)	RT	613 (52)	822 (62)	209 (75)
			Error rate	.004 (.041)	.093 (.088)	.090 (.069)
Yes	Inducer	Upper (MC)	RT	618 (59)	829 (81)	211 (76)
			Error rate	.009 (.037)	.094 (.115)	.085 (.090)
		Lower (MI)	RT	650 (55)	776 (52)	126 (66)
			Error rate	.006 (.044)	.040 (.046)	.034 (.065)
	Diagnostic	Upper (MC)	RT	648 (53)	807 (57)	159 (73)
			Error rate	.007 (.047)	.044 (.056)	.037 (.068)
		Lower (MI)	RT	641 (59)	838 (70)	197 (70)
			Error rate	.009 (.038)	.064 (.085)	.055 (.065)

PC: proportion congruence; DV: dependent variable; MC: mostly congruent; RT: reaction time; MI: mostly incongruent.
Values in parentheses indicate standard deviation of the mean. Note that items in the diagnostic set were 50% congruent in each location. For half of the participants, the upper location was MC and the lower location was MI (as depicted here); for the other half, this was reversed.

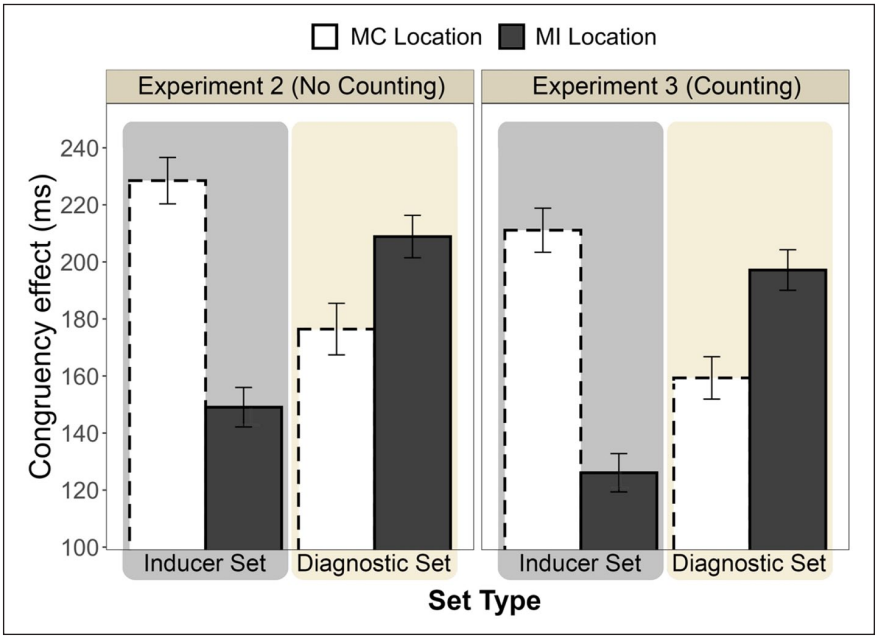

Figure 3. Mean congruency (flanker) effects as a function of location-specific proportion congruence for the inducer and diagnostic sets in Experiment 2 (no counting) and Experiment 3 (counting).
The error bars depict one within-subject standard errors. The inducer set is highlighted in grey and the diagnostic set is highlighted in ivory, corresponding to Figure 1. Note that the same items (e.g., left/right arrows) played the role of the mostly congruent inducer items in the mostly congruent location (leftmost bar) and the role of the 50% congruent diagnostic items in the mostly incongruent location (right-most bar). Conversely, a different set of items (e.g., up/down arrows) played the role of the mostly incongruent inducer items in the mostly incongruent location (left middle bar) and the role of the 50% congruent diagnostic items in the mostly congruent location (right middle bar).

was significantly larger in the MC location (M=228 ms) compared with the MI location (M=149 ms), $F(1,100)$= 42.54, $p<.001$, $\eta_p^2 = .30$, BF_{01}=0.00, indicating the standard LSPC effect (i.e., Location × Trial Type interaction). However, for the diagnostic set, the congruency effect was significantly larger in the MI location (M=209 ms) than the

MC location (M=176 ms), $F(1,100)=5.75$, $p=.02$, $\eta_p^2=.05$, $BF_{01}=0.37$, indicating a reversed LSPC effect.

Error rate. The main effect of trial type, $F(1,100)=117.20$, $p<.001$, $\eta_p^2=.54$, $BF_{01}=0.00$, was significant revealing that incongruent trials (M=0.074) were less accurate than congruent (M=0.007) trials. The main effects of set type, $F<1$, $BF_{01}=12.70$, and location PC, $F(1,100)=3.79$, $p=.06$, $\eta_p^2=.04$, $BF_{01}=8.57$, were not significant. The Location PC × Trial Type interaction was significant, $F(1,100)=4.54$, $p=.04$, $\eta_p^2=.04$, $BF_{01}=4.59$, showing a larger congruency effect in the MC location (M=0.073) compared with MI location (M=0.063) when the inducer and diagnostic sets were collapsed together. The Set Type × Trial Type interaction was not significant, $F<1$, $BF_{01}=9.45$, but the Set Type × Location PC interaction was significant, $F(1,100)=19.32$, $p<.001$, $\eta_p^2=.16$ =.16, $BF_{01}=0.00$. Most importantly, the three-way interaction was significant, $F(1,100)=32.05$, $p<.001$, $\eta_p^2=.24$, $BF_{01}=0.00$. Decomposing the three-way interaction, for the inducer set, the congruency effect was larger for the MC location (M=0.098) than the MI location (M=0.036), $F(1,100)=43.99$, $p<.001$, $\eta_p^2=.31$, $BF_{01}=0.00$. For the diagnostic set, this LSPC effect was reversed such that the congruency effect was larger for the MI location (M=0.090) compared with the MC location (M=0.047), $F(1,100)=14.80$, $p<.001$, $\eta_p^2=.13$, $BF_{01}=0.02$.

Discussion

Using an arrow-flanker task, Experiment 2 reproduced the key finding from Experiment 1, namely the three-way interaction showing a standard LSPC effect for the inducer items and a reversed LSPC effect for the diagnostic items, consistent with the item–PC learning hypothesis. This finding provides further evidence that the dominance of item–PC learning over location–PC learning in the LSPC paradigm is neither specific to the overlapping sets design or variants of the Stroop task.

Experiment 3

Having demonstrated that the dominance of item–PC learning over location–PC or conjunctive learning is neither design- nor task-specific but instead is quite robust and pervasive, the purpose of Experiment 3 was to try to disrupt this dominance by shifting attention to location and thereby facilitating location–PC or conjunctive (location–item)–PC learning. The experiment was identical to Experiment 2 except for the counting procedure. Given prior findings demonstrating that having participants count the number of stimuli in a select context (here, the upper or lower location) is effective in shifting their attention to the context and facilitating context–PC learning (see the study by Crump et al., 2008, for evidence of shape–PC learning when participants counted stimuli that followed a select shape but not when they did not count), one hypothesis was that participants should now consider location when organising and binning prior experiences resulting in LSPC effects for inducer and diagnostic items (evidence of location–PC learning) or an LSPC effect for inducer items but a null LSPC effect for diagnostic items (evidence of conjunctive [location–item] learning). The alternative hypothesis is that the tendency to learn about items and not locations is strong and, therefore, not mutable based on procedures such as counting. In this case, we should again find a three-way interaction (LSPC effect for inducer items and reversed LSPC effect for diagnostic items), supporting the item–PC learning hypothesis.

Method

Participants. A total 104[10] undergraduates (62 women; age M=19.36, SD=1.49) from Washington University in Saint Louis participated in this study. All participants earned class credit for participation. All participants reported that they had normal or corrected-to-normal vision.

Stimuli and design. The stimuli and design were identical to Experiment 2 (see Table 3).

Procedure. The procedure of Experiment 3 was equivalent to Experiment 2 except for the addition of a counting task. Participants were instructed to keep a running count of the number of trials that occurred in either the upper or lower location. Whether participants were counting the upper or lower location was counterbalanced across subjects. Five counting probes appeared during each 120-trial block of the flanker task (20 total) prompting participants to report their current count by typing the count into a response box. The number of trials between counting probes was random, with the restriction that probes occurred at least five trials apart from each other. Once participants reported their count, they were instructed to reset their count to zero and begin a new count on the next trial.

Results

Seven participants were excluded from analysis due to overall slow RT, or an excessive number of errors in the flanker task or counting task (beyond 2.5 standard deviation from the mean of all participants[11]). Data from 97 participants were used for the final analysis. Participants' estimated count numbers were accurate yielding the mean absolute error of 1.23 trials. As in the preceding experiment, data were trimmed (RTs slower than 2,000 ms or faster than 200 ms) resulting in exclusion of less than 1% of all trials. The mean RT and error rates are summarised in Table 4. For both RT and error rate, a 2 (Location PC: MC vs. MI) × 2 (Trial Type: Congruent vs. Incongruent)

× 2 (Set Type: Inducer vs. Diagnostic) repeated-measures ANOVA was conducted.

Reaction time. Only correct responses were included in the RT analysis. The main effects of trial type, $F(1,96)=737.99$, $p<.001$, $\eta_p^2=.88$, $BF_{01}=0.00$, and set type, $F(1,96)=50.86$, $p<.001$, $\eta_p^2=.35$, $BF_{01}=2.25$, were significant but not location PC, $F<1$, $BF_{01}=12.41$. The significant main effects showed that mean RT was slower for incongruent ($M=813$ ms) than congruent ($M=639$ ms) trials and slower for diagnostic ($M=734$ ms) than inducer ($M=718$ ms) items. The congruency effect was larger for the MC location ($M=185$ ms) than the MI location ($M=162$ ms), resulting in a Location PC × Trial Type interaction, $F(1,96)=20.15$, $p<.001$, $\eta_p^2=.17$, $BF_{01}=0.38$. And the effect was larger for the diagnostic set ($M=178$ ms) than the inducer set ($M=169$ ms), $F(1,96)=6.93$, $p=.01$, $\eta_p^2=.07$, $BF_{01}=4.93$. The Set Type × Location PC interaction, $F(1,96)=4.97$, $p=.03$, $\eta_p^2=.05$, $BF_{01}=3.67$, was also significant. Of our primary interest, the three-way interaction was significant, $F(1,96)=32.25$, $p<.001$, $\eta_p^2=.25$, $BF_{01}=0.00$ (see Figure 3). The three-way interaction was decomposed based on set type. For the inducer set, the congruency effect was significantly larger in the MC location ($M=211$ ms) than the MI location ($M=126$ ms), $F(1,96)=50.30$, $p<.001$, $\eta_p^2=.34$, $BF_{01}=0.00$, indicating the standard LSPC effect (i.e., Location × Trial Type interaction). For the diagnostic set, however, the congruency effect was significantly larger for the MI location ($M=197$ ms) than the MC location ($M=159$ ms), $F(1,96)=9.83$, $p=.002$, $\eta_p^2=.09$, $BF_{01}=0.12$, indicating a reversal of the LSPC effect.

Error rate. The main effects of trial type, $F(1,96)=59.31$, $p<.001$, $\eta_p^2=.38$, $BF_{01}=0.00$, set type, $F(1,96)=7.10$, $p=.009$, $\eta_p^2=.07$, $BF_{01}=6.64$, and location PC, $F(1,96)=11.73$, $p=.001$, $\eta_p^2=.11$, $BF_{01}=3.00$, were significant, indicating that participants made more errors for incongruent ($M=0.061$) than congruent ($M=0.008$) trials, inducer ($M=0.037$) than diagnostic ($M=0.031$) sets, and MC ($M=0.039$) than MI ($M=0.030$) locations. The Location PC × Trial Type interaction, $F(1,96)=11.23$, $p=.001$, $\eta_p^2=.10$, $BF_{01}=2.16$, was significant showing that the congruency effect was larger in the MC location ($M=0.061$) than the MI location ($M=0.045$). The Set Type × Trial Type interaction was significant, $F(1,96)=10.04$, $p=.002$, $\eta_p^2=.09$, $BF_{01}=3.20$, indicating that the congruency effect was larger for the inducer ($M=0.060$) than diagnostic set ($M=0.045$). The Set Type × Location PC interaction was also significant, $F(1,96)=7.91$, $p=.006$, $\eta_p^2=.08$, $BF_{01}=0.00$. Most importantly, the three-way interaction was significant, $F(1,96)=10.43$, $p=.002$, $\eta_p^2=.10$, $BF_{01}=0.01$. The three-way interaction was decomposed based on the set type. For the inducer set, the congruency effect was larger in the MC location ($M=0.085$) than MI location ($M=0.034$), $F(1,96)=14.95$, $p<.001$, $\eta_p^2=.13$, $BF_{01}=0.02$. For the diagnostic set, the pattern was reversed such that the congruency effect was larger in the MI location ($M=0.055$) than the MC location ($M=0.037$) but this difference did not reach significance, $F(1,96)=3.29$, $p=.07$, $\eta_p^2=.03$, $BF_{01}=2.38$.

Combined analysis. We combined datasets from Experiment 2 and 3 to directly examine the effect of the counting manipulation (no counting in Experiment 2 vs. counting in Experiment 3). For both RT and error rate, a 2 (Counting: Counting vs. No Counting) × a 2 (Location PC: MC vs. MI) × 2 (Trial Type: Congruent vs. Incongruent) × 2 (Set Type: Inducer vs. Diagnostic) mixed ANOVA was conducted with counting as the only between-subjects factor. Only results related to the counting factor are reported.

For reaction time, the main effect of counting was not significant, $F(1,196)=1.84$, $p=.18$, $\eta_p^2=.01$, $BF_{01}=3.65$, showing that RT was not different between counting and no counting groups. A significant Counting × Trial Type interaction, $F(1,196)=4.04$, $p=.05$, $\eta_p^2=.02$, $BF_{01}=0.34$, revealed that the congruency effect was larger in the no counting group ($M=191$ ms) than counting group ($M=173$ ms). The Counting × Set Type interaction was significant, $F(1,196)=18.25$, $p<.001$, $\eta_p^2=.09$, $BF_{01}=7.71$, showing that the difference between marginal RT for inducer and diagnostic sets was bigger in the counting group ($M_{inducer}=718$ ms, $M_{diagnostic}=733$ ms) relative to the same difference in the no counting group ($M_{inducer}=708$ ms, $M_{diagnostic}=711$ ms). There was not a Counting × Location PC interaction, $F<1$, $BF_{01}=13.25$. There were no three-way interactions with counting and most importantly the four-way interaction was not significant, $F<1$, $BF_{01}=5.05$.

For error rate, the main effect of counting was not significant, $F(1,196)=1.04$, $p=.31$, $\eta_p^2=.01$, $BF_{01}=7.03$, suggesting that the marginal error rate was not different between counting and no counting groups. There were no two-way interactions with counting. The Set Type × Trial Type × Count interaction was significant, $F(1,196)=6.68$, $p=.01$, $\eta_p^2=.03$, $BF_{01}=3.70$, indicating that the congruency effect was larger for the inducer set than the diagnostic set for the counting group ($M_{inducer}=0.060$, $M_{diagnostic}=0.046$) but not for the no counting group ($M_{inducer}=0.067$, $M_{diagnostic}=0.068$). The four-way interaction was not significant, $F(1,196)=1.66$, $p=.20$, $\eta_p^2=.01$, $BF_{01}=2.83$.

Discussion

Despite participants accurately counting the number of stimuli that appeared in a location, a procedure designed to shift participants' attention towards location (i.e., bin based on location or weight location more heavily than items; cf. Crump et al., 2008), the same three-way interaction was observed as in the previous experiments. This evidence

suggests that the tendency to learn about items instead of locations or location–item conjunctions in the LSPC paradigm is reliable and not easily disrupted. Furthermore, the cross-experimental analysis confirmed that there was not a reduced tendency to learn about the items in Experiment 2 (no counting) compared with Experiment 3 (counting); in fact, the effect size for the three-way interaction was nominally larger in Experiment 3.

General discussion

The primary take home message of this study is simple: in an LSPC paradigm, we repeatedly found that participants learned about items and not locations or location–item conjunctions. This pattern is consistent with Bugg et al. (2020) but extends their findings in three significant ways: First, the present design utilised non-overlapping two-item stimulus sets, a feature commonly used in the LSPC literature and a feature that was present in the study by Crump and Milliken (2009) and several subsequent replication/reproduction attempts (Crump et al., 2017; Hutcheon & Spieler, 2017). Second, this study pitted item–PC learning and location–PC learning (as well as conjunctive learning) against each other in the context of two tasks not examined by Bugg et al. (2020). One was the prime-probe, colour–word Stroop task used by Crump and Milliken (2009) in their original report of an LSPC effect for inducer and diagnostic items. The second was an arrow flanker task, a commonly used task in the LSPC literature that we reasoned might be more likely to produce evidence of location–PC learning given that items are less distinctive and given prior results showing the importance of spatial conflict in other location-based patterns (Pickel et al., 2019). The fact that the three-way interaction demonstrating item–PC learning was present regardless of the design or task indicates that the dominance of item–PC learning is neither specific to the overlapping sets design or the picture–word Stroop task used by Bugg et al. A third extension is that in this study, we attempted to disrupt the dominance of item–PC learning, which Bugg et al. did not attempt. However, having participants count the stimuli appearing in a location did not override this dominance—once again, item–PC learning prevailed. These results converge in demonstrating that the tendency to learn about items in an LSPC paradigm is robust, reliable, and not easily disrupted.

Because we used the non-overlapping two-item sets design in the present experiments, it is important to acknowledge that what participants may have learned about items is not associations with PC per se, but instead were associations between specific distractors (words or flankers) and a high contingency response (i.e., item-specific S-R contingencies; see Schmidt & Besner, 2008). In the standard LSPC design (Figure 1, upper panel), an item is MC in one context and MI in the other, leading to the item being unbiased (50% congruent) across the entire experiment. In our experiments, however, items are biased across the entire experiment. Consider again Table 1. In the counterbalance described in the table, the word BLUE is MC in the upper context and unbiased in the lower context; collapsing across both locations, the PC of the word BLUE is MC (i.e., 70%) with the colour blue being the high contingency response. In contrast, the word RED is MI in the lower context and unbiased in the upper context; collapsing across both locations, the PC of the word RED is MI (i.e., 30%) with the colour green being the high contingency response. (Similarly, the word YELLOW like BLUE is MC and associated with a high contingency congruent response and the word GREEN like RED is MI and associated with a high contingency incongruent response.) Thus, the dominant type of item learning that participants used in the present experiments may be item-specific contingency learning. While this implies a different mechanism—item-specific contingency learning involves predicting high contingency responses associated with specific stimuli and not abstract attentional adjustments as in the case of item–PC learning, this does not detract from the overarching conclusion that item learning dominates over location learning. It simply means that item learning in the LSPC paradigm could reflect the learning of associations between items and their PC (as in the study by Bugg et al., 2020) and/or between items and the response that was most contingent on the item (as was possible in the present design).

As noted in the introduction, this study also provided another opportunity to examine whether conjunctive learning is a dominant type of learning in the LSPC paradigm. In none of the experiments were the data consistent with the conjunctive learning hypothesis. This aligns with the results of Bugg et al. (2020), including their Experiment 4, which directly tested the conjunctive learning hypothesis. However, Bugg et al. considered conjunctive learning exclusively to be learning that supports control given use of the overlapping sets design. The present non-overlapping two-item sets design afforded participants the opportunity to learn location–item contingencies (i.e., compound contingency learning; Schmidt & Lemercier, 2019), that is, the predictability of specific responses based on the word [Experiment 1] or based on the flanker [Experiments 2 and 3] dimension in a particular location (see Figure 1, right panel). It is unclear why conjunctive learning (either to guide control or to predict specific responses) was not a dominant form of learning in this study (see also the studies by Crump et al., 2017; Crump & Milliken, 2009; Hutcheon & Spieler, 2017). Possibly, a different pattern of results would be found if a different contextual cue were used, such as font, given the evidence for compound-cue contingency learning in a font-based PC paradigm (Schmidt & Lemercier, 2019). Alternatively, the design used by Schmidt and Lemercier (2019), which was unique

from this study (and Crump & Milliken, 2009, and the prior replication/reproduction attempts), may have somehow encouraged conjunctive learning.

An important implication of the present findings is the possibility that the robust tendency to learn about items (be it learning PC per se or contingent responses) may contribute to the difficulty researchers have faced in replicating Crump and Milliken's (2009) finding of LSPC effects for inducer and diagnostic items. Understanding this difficulty is important as the findings of Crump and Milliken represent a key piece of evidence demonstrating location-specific reactive control. As detailed in the Introduction section (see also Figure 1, top panel), in the standard design used by Crump and Milliken (see also the studies by Crump et al., 2017; Hutcheon & Spieler, 2017) learning about items should lead to no LSPC effect for either inducer or diagnostic items, exactly as prior replication/reproduction attempts have found. The present findings extend the study of Bugg et al. (2020) by demonstrating that the dominance of item–PC learning is not specific to the overlapping sets design or the picture–word Stroop task they employed. Rather, this tendency is also found when the non-overlapping two-item sets design is used in conjunction with the colour–word prime-probe Stroop task or flanker task.

In sum, the current findings (see also the study by Bugg et al., 2020) provide direct evidence for the dominance of item–PC learning in the inducer/diagnostic design used herein; the absence of LSPC effects in prior studies that used the standard inducer/diagnostic design (see upper panel of Figure 1) can also be accounted for by item–PC learning. However, it is important to note that these findings do not imply that location is never used to guide control in the LSPC paradigm or that location could not dominate. Indeed, there is evidence for location-based adjustments that cannot be accounted for by an item–PC learning account. For example, LSPC effects are reliably found when only inducer items are included in the design (i.e., all items are MC when presented in the upper location and those same items are MI when presented in the lower location; Crump et al., 2006; see also the studies by Bugg et al., 2020; Hutcheon & Spieler, 2017), and LSPC effects in inducer-only paradigms reliably transfer to novel locations (Weidler & Bugg, 2016; Weidler et al., 2020). Although it is not certain (as these designs do not include diagnostic items that are 50% congruent in both locations) whether these patterns reflect pure location–PC learning or learning of location–item conjunctions, the key point for present purposes is that location information is contributing to the control adjustments. The clearest evidence of location–PC learning is found in studies that have shown an LSPC effect for diagnostic items in the inducer/diagnostic design where inducer items are 100% congruent or 100% incongruent (Crump et al., 2017; Crump & Milliken, 2009). Additional research is needed to pinpoint why location–PC learning dominates in this case, but item learning dominates when the inducer/diagnostic design with MC and MI inducer items is used as in this study (see also studies by Crump et al., 2017; Crump & Milliken, 2009; Hutcheon & Spieler, 2017; see Footnote 1).

Limitations and future directions

It might be suggested that although the present design and task were better matched to those of prior relevant studies such as by Crump and Milliken (2009) than was the design of Bugg et al. (2020), the design still has unique elements that limit generalisability of the present findings. Specifically, unlike the standard design, the inducer items differed across locations (e.g., blue and yellow were MC in the upper location and red and green were MI in the lower location), as did the diagnostic items (e.g., blue and yellow were 50% congruent in the lower location and red and green were 50% congruent in the upper location). Consequently, the diagnostic items shared features with the inducer items. The reason for altering the design in this manner was to set up experiments wherein evidence for item–PC learning (in addition to evidence for either location–PC or conjunctive learning) would not be dependent on a set of null effects, effects which could also be explained by list-level learning in prior designs. At present, we do not see another way around this challenge, and critically, the design still afforded participants the opportunity to learn about any of the three potential cues (locations, items, or location–item conjunctions) just as in the standard design. Thus, we believe the findings this design yields inform our understanding of the relative dominance of the three types of learning in the LSPC paradigm.

In addition, one might find it problematic that the present design does not include standard diagnostic items (50% congruent and no feature overlap with the inducer items) and thus does not afford the opportunity to examine location-specific control independent of known confounds (i.e., using confound-minimised diagnostic items; Braem et al., 2019). This is a legitimate point. However, the purpose of the present experiments was to understand the learning that occurs within an LSPC paradigm and not to seek out evidence for control independent of known confounds. In the standard design, the (contingency- and frequency-biased) inducer items serve the purpose of promoting the learning of location–PC associations but as the prior studies have demonstrated, such learning largely is not occurring. As we have noted previously (Bugg et al., 2020), without such learning one cannot expect to find LSPC effects for standard diagnostic items. Thus, our approach was to take a step back and start with the question of what *is* learned in an LSPC paradigm before potentially moving to the question of whether the learning induces location-specific control. If future studies are successful in shifting the dominance from items to locations in the

current design, then follow-up studies could be performed that include a third two-item set that shares no features with either the inducer or diagnostic items to assess location-specific control independent of known confounds.

Another potential limitation is that we chose to extend the findings of Bugg et al. (2020) to a second Stroop task (colour–word instead of picture–word) and a flanker task, but not to a Simon task. On the view that conflict originates from stimulus processing in the former task but response conflict in the Simon task (Egner, 2008; but see the study by Hübner & Töbel, 2019 for an alternative view), and the assumption that the dominance of item learning may differ depending on the type of conflict, then we cannot generalise our findings to the Simon task. However, of the three task types (Stroop, flanker, Simon), the Simon task is the least frequently used task in the LSPC literature with only one prior study to our knowledge examining an LSPC effect in the Simon task (Hübner & Mishra, 2016) and no prior studies using the inducer/diagnostic design.

From our perspective, a central unresolved question that merits further attention is the question of how to encourage participants to learn about and use the location–PC associations, which is what the LSPC paradigm was originally intended to do. Counting was not effective in achieving this goal, which suggests that it is not simply a matter of drawing participants' attention to location. Quite possibly this is because the task goal of naming the ink colour or responding to the central arrow's direction makes items, which are defined in whole (Bugg et al., 2020) or part (present experiments) by the relevant dimension (that is the ink colour or central arrow's direction), the most regularly attended and salient feature of the episodic file even when participants are counting (cf. Bugg & Dey, 2018). Therefore, it may be more appropriate to ask the question of how to discourage participants from learning about and using the item–PC associations with the assumption being that if you do so successfully, participants may default to learning about the locations. This is a valuable direction for future research, and it would be informative to see how this approach affects item and location learning in both the current design and the standard inducer/diagnostic design of Crump and Milliken (2009).

One possible and perhaps counter-intuitive approach is to implement a working memory load. Previous studies have shown that when attentional resources are scarce such as under high working memory load, attention is controlled in an efficient manner by exploiting implicitly learned associations (e.g., between items and PC; see the studies by Spinelli, Krishna, Perry, & Lupker, 2020; Suh & Bugg, 2021; see also the studies by Annac, Zang, Müller, & Geyer, 2019; Vickery, Sussman, & Jiang, 2010 for further evidence in spatial context learning paradigms). If we can assume that location–PC learning, and specifically the binning that underlies it, is the most efficient approach in the current paradigm, then it is possible that location–PC learning may dominate under high working memory load. This assumption is bolstered by the fact that location–PC learning depends on organising experiences during the task into just two bins and consequently representing just two location–PC associations whereas binning based on items (or location–item conjunctions) potentially requires more working memory resources at least in the initial stages of learning because one must represent experiences in each of four bins (or eight for conjunctive learning; see Figure 1). However, it is also possible that item–PC learning is more efficient for other reasons (which may contribute to its dominance) and if that is true, then the dominance of item–PC learning may become even stronger under high load compared with low load.

Conclusion

We found that participants learned item–PC rather than location–PC (or conjunctive) associations in an LSPC paradigm with the non-overlapping two-item sets design in both prime-probe Stroop and flanker tasks. These findings demonstrate that the tendency to learn about items as opposed to locations is neither design- or task-specific; rather, it is robust and reliable across designs and tasks, including the picture–word Stroop task and overlapping sets design used previously (Bugg et al., 2020). In addition, we found the same tendency even when participants were encouraged to prioritise location information by counting the number of occasions that a stimulus was presented in either the upper or lower location. Collectively, our findings provide additional and strong evidence supporting the dominance of item learning over location learning and conjunctive learning when all three learning opportunities are available within a single task context. Future studies that attempt to seek evidence of location-specific control should be mindful of this dominance when considering the design of such studies and the interpretation of the results.

Declaration of conflicting interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors thank Spencer Lehmann and Caroline Moss for assistance with data collection. J.M.B. and J.S. were supported by NIA AG057937.

ORCID iD

Julie M Bugg https://orcid.org/0000-0002-0969-186X

Data accessibility statement

The data are available at https://osf.io/b8w5t/.

Notes

1. Note that patterns like this one are specific to the design in which the inducer locations are mostly congruent (MC) or mostly incongruent (MI); when those locations are 100% congruent or 100% incongruent, location-specific control is reliably evidenced (i.e., there is a location-specific proportion congruence [LSPC] effect for diagnostic items [no LSPC effect can be calculated for the inducer items]). It remains a mystery as to why the designs yield such different patterns.
2. It is possible that participants learned about the overall list (experimental context) and not specific items, which would result in the same set of null effects (i.e., no LSPC effect for inducer or diagnostic items because a uniform [list-wide] attentional setting is applied to all items).
3. Note that the diagnostic items are 50% congruent but share features with the inducer items and thus are not intended to examine control independent of known confounds in this design (see Braem et al., 2019).
4. In addition, it is notable that Crump et al. (2017) performed Monte-Carlo simulations for context-specific proportion congruence (PC) studies (e.g., LSPC) using different effect sizes (10, 20, and 30 ms), task types (e.g., Stroop, flanker) and sample sizes (25, 50, 100, 150). They concluded that sample sizes in the range of 100–150 will more reliably find such PC effects if they exist. The sample sizes in all three of the present experiments (94, 103, 104, respectively) are consistent with that recommendation.
5. Hereafter, the terms blue and yellow (or red and green) refer to both the colour patches and the words.
6. Experiments 2 and 3 were later conducted using an online platform and, therefore, we anticipated more outliers and thus planned to exclude participants whose performance deviated from the mean of all participants by 2.5 *SD*. This is not something we routinely do in our laboratory studies with flanker or Stroop. Thus, the analysis plan for Experiment 1 did not include such participant exclusions (see also Bugg et al., 2020). To confirm that the results of Experiment 1 did not depend on the inclusion of any potential outliers, we applied the same trims as later applied in Experiments 2 and 3. Seven participants were excluded from analysis due to overall slow reaction time (RT) and an excessive number of errors (beyond 2.5 *SD* from the mean of all participants), leaving 87 participants. Mean RT of excluded participants was 786 ms compared with 560 ms for included participants; mean error rate of excluded participants was .057 compared with .008 for included participants. Results were almost identical and none of our conclusions changed depending on the analytical approach.
7. We posted 106 experimental slots online (110% of the target sample size 96) and 103 participants completed the study.
8. Mean RT of excluded participants was 975 ms compared with 698 ms for included participants; mean error rate of excluded participants was .719 compared with .037 for included participants.
9. A reviewer raised the possibility that the Location PC × Trial Type interaction from the omnibus analysis of variance (ANOVA) for Experiments 2 and 3 might implicate location–PC learning or conjunctive learning. This interaction means that, when collapsed across inducer and diagnostic sets, on average there was a larger congruency effect in the MC location compared with the MI location. We suspected that this interaction could be an artefact of averaging across the sets with the means following the pattern of the inducer set, which expectedly (given the comparison between PC-75 and PC-25 cells) produced a larger difference between the MC and MI locations than the diagnostic set (which is a comparison of two PC-50 cells) and produced a larger congruency effect in the MC than MI location. One way to address this possibility is to analyse the Location PC × Trial Type interaction separately for the inducer set and diagnostic set (predictions remain the same as those illustrated in Figure 1) with all three hypotheses predicting a Location × Trial Type interaction for the inducer set whereby the MC location yields a larger congruency effect than the MI location, but the three hypotheses making unique predictions for the diagnostic set—a Location × Trial Type interaction in the same direction (MC > MI location) for the location–PC learning hypothesis, a Location × Trial Type interaction in the opposite direction (MC < MI location) for the item–PC learning hypothesis, and no Location × Trial Type interaction for the conjunctive learning hypothesis (MC = MI location). The results of these analyses are reported in each results section when decomposing the three-way interaction. The results uniquely support the item–PC learning hypothesis. The Location × Trial Type interaction was significant for the inducer items (as all three hypotheses expected). A Location × Trial Type interaction was also found for the diagnostic items, and it was in the opposite direction of the interaction for the inducer items, which cannot be explained by any hypothesis except the item–PC learning hypothesis. Consequently, we believe these results make it difficult to interpret the two-way interaction from the omnibus ANOVA where inducer and diagnostic items were combined as support for the location–PC learning or conjunctive learning hypothesis.
10. We posted 106 experimental slots online (110% of the target sample size 96) and 104 participants completed the study.
11. Mean RT of excluded participants was 1,041 ms compared with 712 ms for included participants; mean error rate of excluded participants was .644 compared with .036 for included participants

References

Annac, E., Zang, X., Müller, H. J., & Geyer, T. (2019). A secondary task is not always costly: Context-based guidance of visual search survives interference from a demanding working memory task. *British Journal of Psychology*, *110*(2), 381–399.

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

Brosowsky, N. P., & Crump, M. J. C. (2021). Contextual recruitment of selective attention can be updated via changes in task

relevance. *Canadian Journal of Experimental Psychology/ Revue canadienne de psychologie expérimentale*, *75*(1), 19–34. https://doi.org/10.1037/cep0000221

Bugg, J. M. (2012). Dissociating levels of cognitive control: The case of Stroop interference. *Current Directions in Psychological Science*, *21*(5), 302–309.

Bugg, J. M. (2015). The relative attractiveness of distractors and targets affects the coming and going of item-specific control: Evidence from flanker tasks. *Attention, Perception, & Psychophysics*, *77*(2), 373–389.

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*, Article 367.

Bugg, J. M., & Dey, A. (2018). When stimulus-driven control settings compete: On the dominance of categories as cues for control. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(12), 1905–1932.

Bugg, J. M., & Hutchison, K. A. (2013). Converging evidence for control of color–word Stroop I interference at the item level. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(2), 433–449.

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*(3), 844–859.

Bugg, J. M., Suh, J., Colvett, J. S., & Lehmann, S. G. (2020). What can be learned in a context- specific proportion congruence paradigm? Implications for reproducibility. *Journal of Experimental Psychology: Human Perception and Performance*, *46*, 1029–1050.

Corballis, P. M., & Gratton, G. (2003). Independent control of processing strategies for different locations in the visual field. *Biological Psychology*, *64*, 191–209.

Crump, M. J., Gong, Z., & Milliken, B. (2006). The context-specific proportion congruent Stroop effect: Location as a contextual cue. *Psychonomic Bulletin & Review*, *13*(2), 316–321.

Crump, M. J., Vaquero, J. M., & Milliken, B. (2008). Context-specific learning and control: The roles of awareness, task relevance, and relative salience. *Consciousness and Cognition*, *17*(1), 22–36.

Crump, M. J. C., Brosowsky, N. P., & Milliken B. (2017). Reproducing the Location-Based Context-Specific Proportion Congruent Effect for Frequency Unbiased Items: A Reply to Hutcheon and Spieler (2016). *Quarterly Journal of Experimental Psychology*, *70*(9), 1792–1807. https://doi.org/10.1080/17470218.2016.1206130

Crump, M. J. C., & Milliken, B. (2009). The flexibility of context-specific control: Evidence for context-driven generalization of item-specific control. *The Quarterly Journal of Experimental Psychology*, *62*, 1523–1532.

Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences*, *12*(10), 374–380.

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition*, *5*(1-2), 183–216.

Hübner, R., & Mishra, S. (2016). Location-specific attentional control is also possible in the Simon task. *Psychonomic Bulletin & Review*, *23*(6), 1867–1872.

Hübner, R., & Töbel, L. (2019). Conflict resolution in the Eriksen flanker task: Similarities and differences to the Simon task. *PLOS ONE*, *14*(3), Article e0214203.

Hutcheon, T. G., & Spieler, D. H. (2017). Limits on the generalizability of context-driven control. *The Quarterly Journal of Experimental Psychology*, *70*, 1292–1304.

Lakens, D., & Caldwell, A. R. (2019, May 28). *Simulation-based power-analysis for factorial ANOVA designs*. https://doi.org/10.31234/osf.io/baxsf

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492–527.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.

Pickel, L., Pratt, J., & Weidler, B. J. (2019). The transfer of location-based control requires location-based conflict. *Attention, Perception, & Psychophysics*, *81*(8), 2788–2797.

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*(3), 514–523.

Schmidt, J. R., & Lemercier, C. (2019). Context-specific proportion congruent effects: Compound-cue contingency learning in disguise. *Quarterly Journal of Experimental Psychology*, *72*(5), 1119–1130.

Spinelli, G., Krishna, K., Perry, J. R., & Lupker, S. J. (2020). Working memory load dissociates contingency learning and item-specific proportion congruence effects. *Journal of Experimental Psychology: Learning Memory, and Cognition*, *46*(11), 2007–2033.

Suh, J., & Bugg, J. M. (2021). On the automaticity of reactive item-specific control as evidenced by its efficiency under load. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(7), 908–933. https://doi.org/10.1037/xhp0000914

Vickery, T. J., Sussman, R. S., & Jiang, Y. V. (2010). Spatial context learning survives interference from working memory load. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1358–1371.

Weidler, B. J., & Bugg, J. M. (2016). Transfer of location-specific control to untrained locations. *Quarterly Journal of Experimental Psychology*, *69*, 2202–2217.

Weidler, B. J., Dey, A., & Bugg, J. M. (2020). Attentional control transfers beyond the reference frame. *Psychological Research*, *84*, 217–230.