Journal of Experimental Psychology: Human Perception and Performance

ISSN: 0096-1523

2021, Vol. 47, No. 7, 908–933
https://doi.org/10.1037/xhp0000914

On the Automaticity of Reactive Item-Specific Control as Evidenced by Its Efficiency Under Load

Jihyun Suh and Julie M. Bugg
Department of Psychological and Brain Sciences, Washington University in St. Louis

Traditionally cognitive control is described as slow-acting, effortful, and strategic. Against this backdrop, the notion of "automatic control" is an oxymoron. However, recent findings indicate control also operates quickly with adjustments occurring outside awareness, leaving open the possibility that control could be automatic under certain conditions. Harnessing one such finding, the item-specific proportion congruent (ISPC) effect (i.e., reduction in congruency effect for mostly incongruent compared with mostly congruent items), we systematically investigated the automaticity of reactive item-specific control by examining its efficiency under a concurrent load. In four experiments using a picture-word Stroop task, participants first performed a block of trials in which an ISPC manipulation was embedded to acquire the item-control associations. In later blocks, we manipulated working memory load within-subjects (verbal in Experiment 1, visuospatial in Experiment 2, and *n*-back updating in Experiments 3 and 4) and compared the ISPC effect between low- and high-load conditions. The results of all four experiments showed that the ISPC effect was robust regardless of working memory load. In Experiment 4, we additionally included diagnostic items to assess whether transfer of item-specific control settings was also automatic. The ISPC transfer effect was abolished under high working memory load. Collectively, the findings suggest that reactive item-specific control is triggered and executed in an automatic manner (regardless of the available attentional resources), but only for items that directly support learning of the item-control associations that underlie item-specific control. We propose several hypotheses to account for these findings and discuss theoretical implications for control.

Public Significance Statement

It is commonly believed that controlling one's attention, for example, to minimize the influence of distractors, is a goal-directed mental process that is deliberate and taxing. However, growing evidence indicates attention can be controlled reactively, such that it is triggered by environmental cues and executed in a seemingly automatic fashion. In the present study, we systematically investigated the automaticity of cognitive control, more specifically, whether a form of reactive control called item-specific control can continue to operate efficiently even in the presence of a concurrent task that consumes working memory resources. A robust and consistent pattern was found showing that item-specific control was not detrimentally affected by a high working memory load compared with a low load. However, we also found a boundary condition for the automaticity of reactive item-specific control. Our findings extend our theoretical understanding of reactive control and suggest it is possible to achieve high levels of cognitive control even under conditions in which attention is directed to a secondary, demanding task.

Keywords: automaticity, cognitive control, item-specific proportion congruency, Stroop, working memory load

Supplemental materials: https://doi.org/10.1037/xhp0000914.supp

Jihyun Suh https://orcid.org/0000-0001-6528-1489
Julie M. Bugg https://orcid.org/0000-0002-0969-186X

Portions of the data were presented at the 61st Annual Meeting of the Psychonomic Society in Montreal, Canada. This research was supported by National Institute on Aging AG057937. The data are available at https://osf.io/vq3mw/. We thank Guus Gare, Isaac Bindert, Lahari Vuppaladhadiam, Sarah Matney, Lauren Seidman, Henry Sosland, and Dylan Nelson for their assistance with data collection.

Correspondence concerning this article should be addressed to Jihyun Suh, Department of Psychological and Brain Sciences, Washington University in St. Louis, Campus Box 1125, St. Louis, MO 63130, United States. Email: jihyun.suh@wustl.edu

The past decade has witnessed a shift in cognitive control research—no longer is cognitive control exclusively conceived of in the traditional sense as slow acting, effortful, and strategic (Norman & Shallice, 1986; Posner & Snyder, 1975; Shiffrin & Schneider, 1977). Rather, it has been demonstrated that control can also be fast acting and reflect implicit adjustments that are based on information outside of awareness (for reviews see Bugg, 2012, 2017; Bugg & Crump, 2012; but see Diede & Bugg, 2017; for evidence that the adjustments may nonetheless involve some effort). This shift was in part precipitated by a class of effects known as item-specific proportion congruence (ISPC) effects (see also Blais et al., 2012; Crump et al., 2006; Crump & Milliken, 2009; Corballis & Gratton, 2003; for a related class known as context-specific proportion congruence effects). First observed by Jacoby et al. (2003), the ISPC effect refers to the reduction in the Stroop effect for items (i.e., stimuli defined by a specific feature) that are mostly incongruent throughout an experiment compared with items that are mostly congruent. Subsequent research has demonstrated that ISPC effects illustrate the operation of reactive control, a mechanism that acts poststimulus onset to modulate attention (Braver et al., 2007; see Bugg & Dey, 2018; Bugg et al., 2011; Bugg & Hutchison, 2013; Spinelli & Lupker, 2020; for a consensus view, see Braem et al., 2019). Reactive control is implicated because, in the ISPC paradigm, mostly incongruent and mostly congruent items are intermixed with 50% of trials comprising each item type. Accordingly, participants cannot anticipate which type of item will be presented on any given trial, and thus cannot proactively prepare different control settings in advance for mostly congruent compared with mostly incongruent items (see Shedden et al., 2013; for evidence that items can be distinguished as mostly congruent or mostly incongruent items ~ 150 ms poststimulus onset). The fact that Stroop effects differ across item types indicates that they are processed differently poststimulus onset, with less attention being allocated to the word dimension for mostly incongruent items. Critically, for present purposes, the finding of an ISPC effect led Jacoby et al. to *speculate* that processing of the irrelevant words may be subject to "automatic control." In the present study, we harness the ISPC effect to test the possibility that one form of reactive control (item-specific control) may be relatively automatic.

A process may be relatively automatic in a variety of senses (Bargh, 1989, 1994; Melnikoff & Bargh, 2018; Moors & De Houwer, 2006). For example, it can be unintentional (i.e., occur without instruction), stimulus-driven (i.e., resistant to top-down control), or efficient (i.e., require minimal attentional capacity as is implied if it continues to operate even when there is a concurrent load; Moors & De Houwer, 2006). Bringing extant evidence to bear on the question of whether reactive control, as indexed by the ISPC effect, is automatic, it seems clear that the differential weighting of the word and color dimension for mostly incongruent compared with mostly congruent items is unintentional given that participants in ISPC paradigms are not instructed nor aware that there are different item types (Bejjani et al., 2020), and yet the effect emerges reliably (for reviews see Bugg & Crump, 2012; Bugg, 2017). It also appears that the effect is stimulus-driven given that modulations of attention underlying the ISPC effect occur very rapidly poststimulus onset and furthermore given evidence that one cannot willfully produce an ISPC effect using top-down control (Entel et al., 2014). As for the third criterion, namely whether the processes underlying the ISPC effect are efficient meaning that they require minimal attentional capacity, one prior study has examined this question (Spinelli et al., 2020). The findings provided initial support for the view that reactive item-specific control is relatively automatic because the ISPC effect was found regardless of the magnitude of the concurrent load (no load, low load, or high load), but the criterion was examined under limited conditions, which we will describe in detail in the next section.

Working Memory Load and Cognitive Control

Working memory (WM) refers to "the mechanisms and processes that hold the mental representations currently most needed for an ongoing cognitive task available for cognitive processing" (Oberauer, 2019; p. 1). Traditionally, WM has been intimately linked to cognitive control, the goal-oriented biasing of attention in favor of goal-relevant information, under the assumption that such biasing entails active maintenance of task goals (Braver & Cohen, 2000; Conway et al., 2001; Conway & Engle, 1994; Engle, 2002; Engle & Kane, 2004; Kane & Engle, 2000, 2003; O'Reilly et al., 1999; Roberts et al., 1994). However, recent conceptualizations of cognitive control raise the question of whether certain control mechanisms may operate rather independently of WM (Braver et al., 2007; Bugg, 2012; Bugg & Crump, 2012). The dual mechanisms account posits two control mechanisms (Braver et al., 2007; cf. Kane & Engle, 2003). A resource-demanding proactive control mechanism maps onto traditional views of control and involves maintaining task goals in an active (sustained) state in anticipation of conflict (De Pisapia & Braver, 2006). In contrast, a reactive control mechanism entails the transient (re)activation of task goals once conflict is detected (see Gonthier et al., 2016; for evidence that proactive and reactive control are dissociable). A prominent account of reactive control is the episodic retrieval account (Crump & Milliken, 2009), which proposes that in paradigms like the ISPC paradigm participants learn associations between particular features (e.g., the color of a word in a color-word Stroop task; the to-be-named picture in a picture-word Stroop task) and the likelihood of conflict associated with these features. Consequently, when participants encounter such a feature on later trials, they reactively retrieve prior episodes involving that feature, and these episodes include the attentional control setting that tended to be used on trials comprising that feature in the past (e.g., for a color or picture that is mostly incongruent, they would retrieve a focused setting that minimizes processing of the word dimension).

There is evidence that WM load impairs proactive control in the Stroop task (Kalanthroff et al., 2015), in line with the notion of a resource-demanding proactive control mechanism. In Kalanthroff et al. participants performed a Stroop task under a concurrent high (2-back memory task) or low (0-back memory task) WM load. They reported an increased interference effect (incongruent RT—neutral RT) and a reversed facilitation effect (neutral RT—congruent RT) under high WM load compared with low WM load, consistent with the authors' view that proactive control would be detrimentally affected when available WM resources were reduced (see also Cohen & Servan-Schreiber, 1992). The question we systematically examined in the present study is whether WM load also impairs reactive control or, alternatively, if reactive control can operate just as efficiently when WM is taxed by concurrent demands.

To date, only one study to our knowledge has examined the effects of WM load using ISPC effects as the index of reactive control (see Soutschek et al., 2013, for a study using congruency sequence effects). In that study, Spinelli et al. (2020) compared ISPC effects under no-load, low-load, and high-load conditions. In the low- and high-load conditions, participants named the ink color of a color word in the Stroop task while simultaneously holding in mind two or five digits, respectively. After they named the ink color, they then encountered a probe string and had to indicate whether the two- or five-digit probe string was the same as the one they were holding in mind. The no-load condition did not require subjects to remember any such string. Using between-subjects manipulations of load, they found a consistent pattern in vocal and manual Stroop tasks—the ISPC effect was not affected by the verbal WM load manipulation. Based on these findings, Spinelli et al. concluded that the ISPC effect is not affected by load and that item-specific control "may be especially useful when WM resources are scarce" (p. 1).

We sought to significantly expand the investigation of Spinelli et al. (2020) to address several limitations and theoretical questions, thereby providing a systematic examination of the efficiency of reactive item-specific control under load. One novel aspect of our study was that we examined the effects of WM load using a "confound-minimized"[1] variant of a picture-word Stroop ISPC paradigm (see Braem et al., 2019; for a consensus view) that consistently yields robust ISPC effects. This variant employs the overlapping sets design wherein the relevant dimension (picture) serves as the ISPC signal (Bugg et al., 2011; see also Bugg & Dey, 2018; Gonthier et al., 2016; see also Chiu et al., 2017, which includes evidence of dissociable neural activation patterns for this version compared with a contingency-confounded version). Participants named the animal in the picture while ignoring the superimposed animal word. Pictures of two animals (e.g., birds and cats) were mostly congruent across trials while pictures of the other two animals (e.g., dogs and fish) were mostly incongruent across trials, and the item sets overlapped (e.g., pictures of birds appeared not just with the words BIRD and CAT, but additionally with the words DOG and FISH). Spinelli et al. (2020) used a different variant that is referred to as the two-item set design. In the two-item set design, the two sets do not overlap (e.g., using the picture-word Stroop task as an example, pictures of the mostly congruent birds and cats would only appear with the words BIRD and CAT, and the mostly incongruent dogs and fish would only appear with the words DOG and FISH). Consequently ISPC is perfectly confounded with contingency in the two-item set design such that the ISPC effect can be attributed to contingency learning (i.e., use of the word to predict the single high contingency response on congruent trials for mostly congruent items and on incongruent trials for mostly incongruent items; Schmidt & Besner, 2008; see also Bugg & Hutchison, Experiment 3) rather than reactive control (but see Spinelli & Lupker, 2020 for evidence that contingency-learning does not entirely explain the ISPC effect even in the two-item set design). Although Spinelli et al. demonstrated dissociative effects of their load manipulation on the ISPC effect and a contingency-learning effect in a separate nonconflict task (i.e., the contingency learning paradigm) by showing that only the latter was detrimentally affected by load (see also Schmidt et al., 2010), it remains possible that the lack of an effect of WM load on the ISPC effect in their study could be specific to their design. Therefore, it is valuable to examine the research question using the overlapping sets design (as opposed to the two-item set design) such that theoretical conclusions concerning effects of load on the ISPC effect can be more confidently attributed to the operation of a reactive control mechanism under varying levels of load (see, e.g., Bugg & Hutchison, 2013, Experiment 2, for a transfer effect demonstrating clear evidence that this design induces item-specific control and not contingency learning [using words to predict highly contingent responses]).

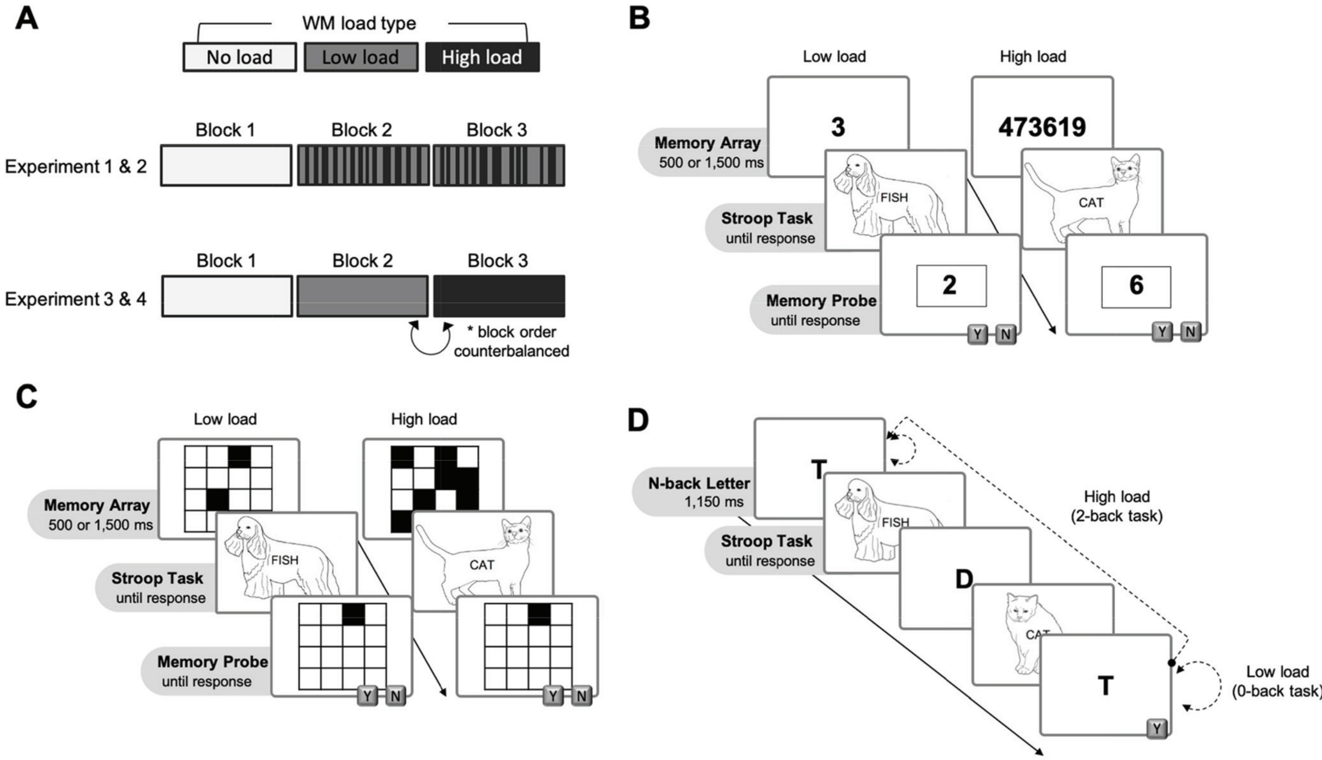
Second, in addition to examining a verbal WM load manipulation in Experiment 1 like Spinelli et al. (2020), we examined two other WM load manipulations to ensure that the pattern of results was not specific to one type of load. We will develop the theoretical motivation for our selection of the novel load types in the next section and in the introductions to Experiment 2 (visuospatial WM load) and Experiment 3 (n-back updating demands). Third, in all our experiments we implemented within-subject manipulations of WM load. Spinelli et al. (2020) used between-subjects manipulations, which are less powerful, with 20 subjects per group in the first two experiments and ~40 per group in the last experiment (the no load group was three times this size for purposes of looking at individual differences within that group). Another advantage of using a within-subjects manipulation is that any differences in the ISPC effect as a function of load can be readily attributed to the manipulation and not to between-subjects factors that could affect the ISPC effect (e.g., WM capacity; Hutchison, 2011).

Fourth, we aimed to examine the effects of WM load on the *retrieval and execution* of reactive item-specific control settings poststimulus onset, rather than the learning of the associations between item types and their requisite control settings (e.g., the binding of an MI item [picture] to a control setting that minimizes processing of the word dimension). Consequently, unlike Spinelli et al. (2020), we employed designs in which all participants first experienced a block of 192 trials (216 trials in Experiment 4) without a WM load to give them an opportunity to learn the item-control associations prior to implementing the WM load manipulation. In a pilot study in our lab, we determined that the ISPC effect does not significantly grow if additional trials are presented beyond that block length. Thus, performance on this first block served as a manipulation check to confirm participants had learned the item-specific control settings. Then, in the remaining blocks of the task, we implemented a WM load manipulation by including low load and high load trials (within or between blocks

[1] In a recent consensus article, Braem et al. (2019) described designs researchers can use to study key markers of adaptive control (in addition to the ISPC effect, the markers were congruency sequence effects, list-wide and context-specific PC effects) while minimizing the influence of potential confounds, such as exploiting simple stimulus-response associations. They referred to these designs as "confound-minimized," which contrasts with confound-prone designs that they did not recommend like the two-item set ISPC design. The term *confound-minimized* as opposed to *confound-free* recognizes that effects are rarely process pure.

Figure 1

A Schematic Illustration of Experimental Design and Sample Displays Used in Experiments 1–4



Note. (A) illustration of working memory (WM) load manipulation in Experiments 1–4. (B) Sample display sequences of low and high verbal WM load trials in Experiment 1. The memory array was presented for 500 ms for the low load and 1,500 ms for the high load trials in Experiments 1 and 2. (C) Sample display sequences of low and high visuospatial WM load trials in Experiment 2. (D) A sample display sequence of the high (2-back task) and low (0-back task) WM load block in Experiments 3 and 4. In this example, the letter *T* serves as a 0-back task target (participants were assigned with a random letter).

depending on the experiment but always within-subjects; see Figure 1A). This design enabled us to better isolate the effects of WM load on reactive control from the effects of load on learning.

Fifth, we examined the question of whether effects of WM load on the ISPC effect may depend on whether attention can be modulated poststimulus onset based on direct prior experiences with stimuli (i.e., retrieval of a control setting that has been associated with a specific stimulus feature based on conflict experiences in the past) or whether generalization is required (i.e., retrieval of a control setting in response to a stimulus feature that is similar to but not exactly the same as stimuli from the past that were responsible for the learning of the control setting). Prior research has demonstrated that ISPC effects do transfer to novel exemplars. For example, if birds and cats are mostly congruent items and dogs and fish are mostly incongruent items, an ISPC effect is found when novel exemplars (i.e., new birds, cats, dogs, and fish that did not serve as the items that allowed participants to learn about ISPC) are encountered that are 50% congruent regardless of animal category. This suggests participants can use the category as a feature to modulate control on an item-by-item basis (Bugg et al., 2011; Bugg & Dey, 2018; Bugg & Hutchison, 2013). However, it remains unknown whether such "transfer" of reactive control may demand more WM resources.

By systematically investigating the effects of load on the ISPC effect, the goals are to significantly expand current theoretical understanding of reactive control and the plausibility of the concept of "automatic control" and inform the practical question of whether it is possible to achieve high levels of cognitive control under conditions of high load.

Overview of Experiments

We conducted a series of four experiments with the goal of examining the automaticity of reactive control, with automaticity here defined by the efficiency criterion. We approached this question from the angle of asking whether the ISPC effect is detrimentally affected by various concurrent WM loads (see Figure 1B–1D). In Experiment 1 we employed a Sternberg-like task manipulating verbal WM load and we found that item-specific control continued to operate as effectively under high load as under low load. In Experiment 2, we employed the same approach but used a visuospatial WM load surmising that this type of load may be more likely to detrimentally affect item-specific control because such control is dependent on rapid processing of the visual features of the stimuli that trigger adjustments in control (i.e., the picture stimulus). Once again, however, we found no effect of the load manipulation on item-specific control. Deviating from the approach of manipulating load via the maintenance component of WM and motivated by prior investigations indicating high updating demands detrimentally affect other markers of control (proactive control; Kalanthroff et al., 2015; reactive sequential control adjustments, Soutschek et al., 2013), Experiments 3 and 4 employed an *n*-back manipulation of updating demands. These two experiments were similar except Experiment 4 included transfer (i.e., diagnostic) trials, novel exemplars representing the animals participants learned control settings for based on the training (inducer) trials in the first block (cf. Bugg et al., 2011; Bugg & Dey, 2018). Including transfer trials allowed us to examine whether the apparent automaticity of item-specific control even in the face of high updating demands (as in Experiment 3) extends beyond stimuli for which one has had prior direct experience acquiring and applying attentional control settings.

Experiment 1

The purpose of Experiment 1 was to examine whether the retrieval and execution of learned item-control associations is influenced by a concurrent WM load, with an eye toward informing the question of whether these processes are automatic with respect to the efficiency criterion (i.e., processes require minimal attentional capacity). We compared the magnitude of the ISPC effect under a low and a high concurrent verbal WM load following learning of the item-control associations in an initial block. We hypothesized that a high load should interfere with item-specific control if the triggering of control processes poststimulus onset shares attentional resources with the concurrent WM task. This would be evidenced by a *reduction* in the magnitude of the ISPC effect under a high WM load compared with a low WM load.[2] In contrast, if item-specific control is automatic, which we define here as operating efficiently regardless of the additional load, we expected that the magnitude of the ISPC effect should be equivalent across the WM load conditions.

Method

Participants

Forty-eight undergraduate students (M_{age} = 19.79, SD_{age} = 1.27, 15 female) from Washington University in St. Louis participated to fulfill a credit as a partial requirement of a course. All participants were native English speakers with normal or corrected-to-normal vision and agreed to participate in the study based on written informed consent provided at the beginning of the study. The experimental procedures were approved by the Institutional Review Board at the Washington University in St. Louis.

We implemented a stopping rule of 48 participants for the present and all subsequent experiments which is larger than the sample sizes used by Spinelli et al. (2020) and larger than the typical sample size reported in previous studies that examined whether WM load affected cognitive control using within-subject designs (Kalanthroff et al., 2015; Soutschek et al., 2013). There has been only one prior study that examined the effect of WM load on the ISPC effect (Spinelli et al., 2020), but load was manipulated between-subjects and a null effect of load on the ISPC effect was found. Thus, a power analysis based on estimates of effect size from that study was not a viable option to calculate the expected sample size for our study. To confirm that our stopping rule was reasonable, we conducted a simulation-based power analysis (Lakens & Caldwell, 2019) using RT data from the within-subjects design of Kalanthroff et al. (2015)[3] who found a detrimental effect of load on proactive control in a Stroop task. With an alpha level of .05, a correlation between the measures of .5, and a partial eta squared of .29, the simulated data yielded a power of .8 with the sample size of 16 for the interaction representing the effect of WM load on cognitive control, which is much smaller than the sample sizes we achieved in our experiments. Nonetheless, given that a null is a theoretically informative outcome in the present experiments, we also report Bayesian analyses to quantify the evidence in support of the null. Bayes Factors (BF) are reported for theoretically important null effects in the form of BF_{01} (Bayes factor favoring H_0 over H_1). A Bayes factor between 1 and 3 means anecdotal and a value between 3 and 10 means substantial evidence for the null hypothesis (Jeffreys, 1961; Wagenmakers et al., 2018).

Apparatus

The stimuli were presented with the E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) on a 17-in. LCD monitor. The vocal response was detected and recorded from a microphone connected to a voice-key of PST serial response box (Psychological Software Tools, Pittsburgh, PA). The standard keyboard was used for the experimenter to code the response.

Stimuli and Procedure

Picture-Word Stroop Task. All stimuli were presented in black on a white background (see Figure 1B for sample displays). Each trial began with a fixation cross ("+," font size: 20) presented at the center of the screen for 1 s. A line-drawing picture of a bird, dog, cat, or fish with either a congruent or incongruent animal word was then presented. We used animal pictures that were previously used in ISPC studies (Bugg et al., 2011; Bugg & Chanani, 2011; Bugg & Dey, 2018). Participants were asked to name aloud the animal in the picture using general names (e.g., "bird" but not "woodpecker") while ignoring the word as quickly and accurately as possible. The proportion congruency was manipulated based on the categories of the animal pictures (i.e., the relevant dimension) in an overlapping sets design. Two animal categories were mostly (PC-75) congruent, and the remaining two animal categories were mostly incongruent (PC-25). For one group

[2] A larger ISPC effect reflects enhanced reactive control relative to a smaller ISPC effect.

[3] We considered using Spinelli and colleagues' (2020) study for the power calculation, but the effect size estimates were not appropriate for our study because they manipulated WM load between subjects whereas our manipulations were entirely within subjects. We used Kalanthroff et al. (2015) because the descriptive data, which were needed for the simulation-based sample size calculation, were available whereas the data were not provided in Soutschek et al. (2013), the other most relevant prior study. Although Kalanthroff et al. examined effects of load on proactive control, we reasoned that if reactive control is susceptible to load just like proactive control, then using estimates from a study that showed a detrimental effect of load on proactive control should provide sufficient power to detect a similar effect on reactive control. We tripled the sample size because we thought it possible that effects of load on reactive control, if observed, could be smaller than those on proactive control and because we manipulated ISPC, which added one more variable to be accounted for compared with Kalanthroff et al.

of participants, pictures of birds and cats were mostly congruent (MC), and pictures of dogs and fish were mostly incongruent (MI). The other group of participants was assigned the opposite mapping such that the dogs and fish were MI and the birds and cats were MC (see Table 1).[4] For each of the animal categories, four unique exemplars were presented with equal probability. Items were randomly intermixed resulting in 50% congruent blocks of trials. The animal picture remained on the screen until the response was detected via the voice key. An experimenter then coded the vocal response using the keyboard during a blank screen following the vocal response (e.g., if the participant said "dog," the experimenter pressed the key corresponding to dog). Extraneous noise (e.g., cough) or undecipherable speech was coded as a "scratch" trial and excluded from analysis.

WM Task. The Sternberg task (Sternberg, 1966) was used to manipulate WM load by varying the number of to-be-remembered digits in a memory array. For the low WM load, a single digit (1–9) was presented for 500 ms. For the high WM load, an array of 6-digits[5] was presented for 1,500 ms. For both low and high WM loads, participants were asked to remember the digits in the memory array and informed that they would be tested on their memory after completing the Stroop task trial. After the Stroop task trial (i.e., after the experimenter coded the participants' response), a single probe digit (1–9) was presented until participants pressed a key on the response box (i.e., two-alternative forced choice) indicating whether the probe digit was in the memory array or not. The chance that a given probe digit was from the memory array was 50%. After the participant responded to the probe, the next Stroop stimulus was presented 1s later.

Design

A 3 (WM load type: no load vs. low load vs. high load) × 2 (ISPC: MC vs. MI) × 2 (trial type: congruent vs. incongruent) within-subject design was used. Each participant completed 16 practice trials (50% congruent) followed by an initial block of 192 trials without the concurrent WM load (see Figure 1A). The purpose of this block was to have participants acquire the item-specific control settings. Then, the concurrent WM task was introduced which was followed by eight practice trials with the concurrent low WM load task and another eight trials with the concurrent high WM load task (i.e., the order was fixed in the practice trials). During the practice trials, a feedback message ("correct" or "incorrect") was shown for the WM task accuracy. Next, participants completed two blocks of 192 Stroop trials (384 trials total) where low and high WM load trials were randomly presented with an equal probability (no feedback was shown). Note that this enabled us to control for order effects (e.g., time on task) between the low and high WM load conditions. There were brief rests at the halfway point of all blocks and the experiment lasted approximately 1 hr.

Table 1
Frequency of Picture and Word Pairs Used in Each Block of Experiment 1

	Picture			
Word	Bird	Cat	Dog	Fish
BIRD	*36*	4	12	12
CAT	4	*36*	12	12
DOG	4	4	*12*	12
FISH	4	4	12	*12*

Note. The italicized values indicate congruent pairs. Each frequency was derived from the repetition of four unique exemplars (i.e., for the BIRD–Bird pair, each of the four pictures was presented nine times). The frequencies described here represent one of two counterbalancing conditions where birds and cats are mostly congruent, and dogs and fish are mostly incongruent.

Results

WM Task

The mean percent probe accuracy for the high WM load condition (85%) was significantly lower than that of the low WM load condition (93%), $t(47) = 7.09$, $p < .001$, Cohen's d = 1.02. Participants were also slower to make a correct probe response in the high WM load condition (805 ms) compared with the low WM load condition (565 ms), $t(47) = 14.06$, $p < .001$, Cohen's $d = 2.03$. Both accuracy and RT suggest that the WM load manipulation was effective.

Stroop Task

As in previous studies using this task (e.g., Bugg et al., 2011; Bugg & Chanani, 2011; Bugg & Dey, 2018), trials faster than 200 ms or slower than 3,000 ms were eliminated from the analysis (less than 1% of trials were removed). For RT analysis, trials with incorrect Stroop responses were excluded. We also excluded trials with the incorrect memory probe responses (10.27% of trials) because errors imply that participants may have failed to maintain the information in WM.[6] Here, we report only theoretically important results.[7] The no load block was analyzed separately from the rest of the blocks where the WM load was manipulated because

[4] The design is set up such that the picture is the ISPC signal (e.g., bird and cat pictures are 75% congruent while dog and fish pictures are 25% congruent), which eliminates the contingency between the ISPC signal and responses (pictures are 100% predictive of the response for all items) that existed in the original two-item set design that Jacoby et al. (2003) developed (see also Spinelli et al., 2020). Regarding the predictiveness of the *word* in conveying the PC of the item, the words are slightly predictive in this design such that the MC words are 56% congruent whereas the MI words are 38% congruent (this contrasts with the two-item set design where the words are 75% and 25% congruent, respectively). This is a design feature that we have described and discussed in our prior work that has used the present design to examine item-specific control (Bugg et al., 2011; Bugg & Dey, 2018; Bugg & Hutchison, 2013; see also Bejjani et al., 2020). Notably, whether this feature is present or not (see Chiu et al., 2017, for a design in which the MC and MI words are both 50% congruent), the same behavioral pattern indicative of item-specific control emerges (i.e., asymmetrical ISPC effect primarily affecting incongruent trials), a pattern that is not consistent with a contingency learning account (i.e., using the word to predict responses). Additionally, using exactly this design, Bugg and Hutchison (2013) demonstrated transfer to stimuli that were paired with novel words, a pattern which is consistent with item-specific control but not contingency learning.

[5] For the high WM array, six digits were pseudo-randomly selected such that none of the two digits was consecutive in order (e. g., "473196" but not "453196," which includes consecutive digits 45).

[6] The same analysis was conducted on the trials including correct and incorrect probe responses and both results mirrored each other. See the online supplemental materials for the full summary of results.

[7] The full summary of results is available in the online supplemental materials.

the no load block was mainly designed to serve as a training phase for the learning of associations between specific items and their requisite control settings (i.e., ISPC). Table 2 summarizes mean RT and error rates.

Reaction Time

Manipulation Check: Item-Specific Learning Under No WM Load. A repeated-measures ANOVA with factors of ISPC (MC vs. MI) and trial type (congruent vs. incongruent) was conducted to test the initial learning of item-control associations during the first block. The incongruent trials were slower compared with congruent trials, $F(1, 47) = 343.36$, $p < .001$, $\eta_p^2 = .88$, $BF_{01} = .00$. Most importantly, a significant ISPC effect was found, $F(1, 47) = 29.24$, $p < .001$, $\eta_p^2 = .38$, $BF_{01} = .02$, showing a reduced Stroop effect in MI (85 ms) compared with MC (117 ms) items.

Effects of WM Load. A repeated-measures ANOVA with factors of concurrent WM load (low vs. high), ISPC (MC vs. MI), and trial type (congruent vs. incongruent) was conducted to test the effect of WM load on the ISPC effect. The main effect of trial type was significant, $F(1, 47) = 125.50$, $p < .001$, $\eta_p^2 = .73$, $BF_{01} = .00$, revealing that the congruent trials (723 ms) were faster than the incongruent trials (802 ms). In addition, the main effect of WM load was significant, $F(1, 47) = 10.87$, $p = .002$, $\eta_p^2 = .19$, $BF_{01} = .02$, suggesting that the concurrent WM load delayed the overall RT (high load condition = 775 ms on average compared with 750 ms in low load condition).

Importantly, a significant Trial Type × Item-Specific PC interaction was found, $F(1, 47) = 15.28$, $p < .001$, $\eta_p^2 = .25$, $BF_{01} = .12$, showing that the Stroop effect for MC items was greater (96 ms) than that of MI items (62 ms), replicating the typical ISPC effect. Most importantly, the three-way interaction was not significant, $F < 1$, $BF_{01} = 3.46$, highlighting the fact that the magnitude of the ISPC effect was not influenced by the concurrent WM load (see Figure 2).[8] The Bayes factor ($BF_{01} = 3.46$) also confirmed that the evidence substantially supported the null hypothesis.

Z-Transformed RT. It is possible that the lack of the three-way interaction in the RT analysis was obscured by the baseline RT difference between the WM load conditions. Specifically, since the Stroop effect tends to be bigger with slower responses (Pratte, Rouder, Morey, & Feng, 2010); we might have overestimated the ISPC effect in high load trials (which were overall slower) while underestimating the effect in the low load trials. To rule out any confounding effect from the baseline RT difference, we *z*-scored individuals' RT based on WM load types and performed the same analysis as reported in the raw RT analysis. Consistent with the raw RT analysis, the main effect of trial type, $F(1, 47) = 127.19$, $p < .001$, $\eta_p^2 = .73$, $BF_{01} = .00$, was significant. Not surprisingly, the main effect of WM load ($F < 1$, $BF_{01} = 8.63$) was not significant given the transformation. The ISPC × Trial Type interaction, $F(1, 47) = 15.86$, $p < .001$, $\eta_p^2 = .25$, $BF_{01} = .00$, remained significant. Most importantly, the three-way interaction, again, was not significant, $F(1, 47) = 1.46$, $p = .23$, $\eta_p^2 = .03$, $BF_{01} = 7.02$.

Error Rate

Manipulation Check: Item-Specific Learning Under No WM Load. A repeated-measures ANOVA as a function of ISPC (MC vs. MI) and trial type (congruent vs. incongruent) showed that participants made more errors on incongruent trials (4.42%) compared with congruent trials (2.37%), $F(1, 47) = 15.72$, $p < .001$, $\eta_p^2 = .25$, $BF_{01} = .00$. Importantly, the Stroop effect was greater in MC items (3.20%) compared with the MI (.91%) items, $F(1, 47) = 4.94$, $p = .03$, $\eta_p^2 = .10$, $BF_{01} = .44$, suggesting that item-control learning had occurred during the initial block.

Effects of WM Load. The repeated-measures ANOVA on percent errors revealed that the main effect of trial type was marginally significant, $F(1, 47) = 3.56$, $p = .07$, $\eta_p^2 = .07$, $BF_{01} = 1.25$, demonstrating that the error rate was slightly higher in incongruent (3.23%) compared with congruent (2.57%) trials. The main effect of WM load was not significant, $F < 1$, $BF_{01} = 8.98$. In addition, a significant interaction between ISPC and trial type, $F(1, 47) = 3.98$, $p = .05$, $\eta_p^2 = .08$, $BF_{01} = 2.01$, indicated that the Stroop effect in error rate was smaller for MI (.15%) compared with MC (1.19%) items (i.e., an ISPC effect). Of our primary interest, the three-way interaction was not significant, $F < 1$, $BF_{01} = 4.26$.

Discussion

In Experiment 1, we investigated whether reactive control is influenced by a concurrent verbal WM load. Although the concurrent high WM load delayed the overall RT and led to more errors when judging WM probes than the low WM load, the magnitude of the ISPC effect was not modulated by the WM load. Bayesian analysis indicated substantial support for a null interaction, and the null interaction could not be explained by baseline differences in overall speed of responding on low and high WM load trials.

The present finding converges with Spinelli et al. (2020), who reported the lack of an effect of verbal WM load (i.e., maintaining a digit or letter array) on the ISPC effect in a color-word Stroop task. Notably, we did establish that the load manipulation was effective in decreasing performance on the WM task in the high compared with low load condition; thus, the lack of an effect of load cannot be attributed to the ineffectiveness of the load manipulation. Of theoretical importance, the present findings extend the prior results using the two-item set design to the confound-minimized, overlapping sets design, giving us more confidence in the conclusion that reactive control may be automatic with respect to the efficiency criterion. Additionally, we have extended the prior results to a design in which the effects of load on item-specific modulation of attention poststimulus onset were better isolated from learning of the item-control associations by allowing participants to engage in a no-load training phase prior to the implementation of the WM manipulation.

Despite the consistent findings, it should be noted that in both our study and that of Spinelli et al. (2020) a verbal WM load was used. According to the *specialized-load-account* (Kim et al., 2005;

[8] A 3 (WM Load: no load, low load, high load) × 2 (ISPC: MC, MI) × 2 (Trial Type: congruent, incongruent) repeated-measures ANOVA on RT, *z*-transformed RT, and error rate showed the lack of a 3-way interaction (see Table S11 in the online supplemental materials for the full reports) suggesting that the magnitude of the ISPC effect was not different between the WM load conditions including the no load condition. Note that this analysis should be interpreted with caution here and in subsequent footnotes since, by design, the low load condition always preceded the low and high load conditions.

Table 2
Mean Reaction Times (RTs) and Percent Errors for Congruent and Incongruent Trials for the Mostly Congruent and Mostly Incongruent Items Under No Load, Low WM Load, and High WM Load in Experiment 1

		No load		Low load		High load	
ISPC	Trial type	*M* (*SD*)	% err (*SD*)	*M* (*SD*)	% err (*SD*)	*M* (*SD*)	% err (*SD*)
Mostly congruent	Congruent	647 (75)	2.47 (2.44)	713 (116)	2.52 (2.45)	744 (137)	2.59 (3.04)
	Incongruent	764 (90)	5.67 (5.62)	815 (146)	3.91 (5.02)	834 (176)	3.57 (4.29)
	Stroop effect	117	3.20	102	1.39	90	0.98
Mostly incongruent	Congruent	643 (77)	2.26 (3.10)	706 (110)	2.30 (3.43)	729 (135)	2.85 (4.61)
	Incongruent	728 (91)	3.17 (2.61)	766 (113)	2.90 (3.04)	792 (140)	2.55 (3.43)
	Stroop effect	85	0.91	60	0.60	63	−0.30
ISPC effect		32	2.29	42	0.79	27	1.28

Note. ISPC = item-specific proportion congruence; WM = working memory.

Park et al., 2007), WM load operates in a domain-specific manner, such that verbal information in WM selectively interferes with processing of verbal information but no other form of information such as visuospatial information. In support of this view, Kim et al. (2005) showed that a concurrent WM load impaired the basic Stroop effect only when the items in WM overlapped with the task-relevant feature. Based on this account, it is possible that the lack of the effect of WM load on the ISPC effect in Experiment 1 (and Spinelli et al., 2020) was attributable to the fact that the verbal WM load failed to target a critical domain of information that is tightly linked to the retrieval of the item-specific control setting. In particular, if visuospatial features of the animal pictures (e.g., features that help identify animal categories such as the head position or presence of a wing or tail) trigger retrieval of the learned control settings, overloading visuospatial WM may be more likely to result in a reduction of the ISPC effect under high load by preventing an in-depth processing of control-triggering visuospatial information. To examine this possibility, Experiment 2 was conducted.

Experiment 2

The purpose of Experiment 2 was to further examine the automaticity of the retrieval and execution of item-specific control by using a visuospatial WM load instead of a verbal WM load. Especially, this change was made to rule out the possibility that the verbal WM load failed to buffer visuospatial information (i.e., contained within the pictures that get associated with low or high conflict probabilities) that would be a key trigger of the previously learned item-specific control settings. If the ISPC effect is not influenced by the concurrent visuospatial WM load, it would provide additional support bolstering the view that item-specific control occurs in an automatic manner regardless of the contents of WM. However, if the ISPC effect is reduced under the high visuospatial WM load compared with the low load, it would rather support the view that item-specific control depends on the availability of visuospatial WM resources and is thus not automatic.

Method

Participants

Forty-eight undergraduate students (M_{age} = 19.83, SD_{age} = 1.31, 29 female) from Washington University in St. Louis participated in the study to fulfill a credit as a partial requirement of a course. All participants were native English speakers with normal or corrected-to-normal vision.

Apparatus, Stimuli, Procedure, and Design

Apparatus, stimuli, procedure, and design were identical to those used in Experiment 1 except the concurrent WM task.

A modified version of the visuospatial WM task was used (Wood, 2011). At the beginning of each trial, participants were presented with a 4 × 4 grid (see Figure 1C). In the low WM load condition, two filled boxes were randomly located in the grid and remained on the screen for 500 ms. In the high WM load condition, eight filled boxes were presented at random locations for 1,500 ms. For the memory probe task, participants were given one filled box at a random location on the grid and requested to press a key on the response box (i.e., two-alternative forced choice) to indicate whether the same location was previously filled or not. For the probe response, the probability of the "yes" response was 50%.

Results

WM Task

The memory probe responses were substantially less accurate in high WM load trials (62%) compared with the low WM load trials (87%), $t(47) = 19.34$, $p < .001$, Cohen's $d = 2.79$. In addition, participants were significantly slower to make correct responses to high WM load probes (890 ms) compared with the low WM load probes (695 ms), $t(47) = 11.87$, $p < .001$, Cohen's $d = 1.73$. These differences indicate that the WM manipulation was successful.

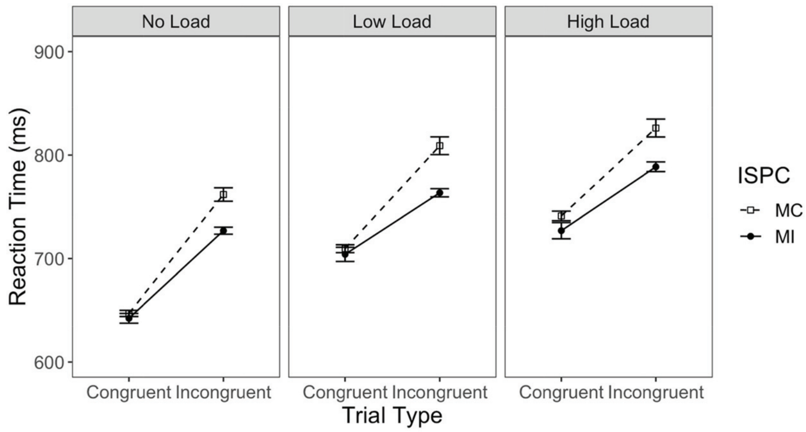
Stroop Task

The same trimming procedures were applied to Stroop performance as in Experiment 1 (less than 1% of trials were removed because of the RT trim), including the exclusion of incorrect trials for the analysis of RT. In addition, trials with incorrect probe responses on the WM task were excluded from the analysis of Stroop performance.[9] Table 3 summarizes the mean RT and error rate for each condition.

[9] The same analysis was conducted on the trials including correct and incorrect probe responses and both results mirrored each other. See the online supplemental materials for the full summary of results.

Figure 2

Reaction Time as a Function of WM Load, ISPC, and Trial Type in Experiment 1



Note. ISPC = item-specific proportion congruence; MC = mostly congruent; MI = mostly incongruent; WM = working memory. Error bars depict within-subject standard error.

Reaction Time

Manipulation Check: Item-Specific Learning Under No WM Load. A repeated-measures ANOVA as a function of ISPC (MC vs. MI) and trial type (congruent vs. incongruent) was conducted to check whether participants were able to acquire the item-control associations during the first block. The main effect of trial type, $F(1, 47) = 281.56$, $p < .001$, $\eta_p^2 = .86$, $BF_{01} = .00$, was significant suggesting that incongruent trials (756 ms) were slower compared with the congruent trials (646 ms). Importantly, the ISPC × Trial Type interaction was significant, $F(1, 47) = 28.07$, $p < .001$, $\eta_p^2 = .37$, $BF_{01} = .00$, showing a reduced Stroop effect in MI items (87 ms) compared with MC items (134 ms; i.e., ISPC effect). This interaction indicates that the item-control associations were learned in the first block in the absence of a WM load.

Effects of WM Load. A 2 (WM Load: low WM load vs. high WM load) × 2 (ISPC: MC vs. MI) × 2 (Trial Type: congruent vs. incongruent) repeated measures ANOVA revealed a significant Stroop effect showing overall faster RT in congruent (717 ms) compared with incongruent (817 ms) trials, $F(1, 47) = 188.61$, $p < .001$, $\eta_p^2 = .80$, $BF_{01} = .00$. The main effect of WM load was not significant, $F(1, 47) = 1.60$, $p = .21$, $\eta_p^2 = .03$, $BF_{01} = 6.24$, indicating that the marginal Stroop RT was not different between the low and high WM load.[10] The ISPC × Trial Type interaction was significant, $F(1, 47) = 38.64$, $p < .001$, $\eta_p^2 = .45$, $BF_{01} = .00$, confirming the ISPC effect (see Figure 3). The magnitude of the Stroop effect was larger with MC items (125 ms) compared with MI items (74 ms). Importantly, the three-way interaction again did not reach statistical significance, $F < 1$, $BF_{01} = 3.95$.[11]

Error Rate

Manipulation Check: Item-Specific Learning Under No WM Load. A repeated-measures ANOVA revealed that participants made fewer errors for congruent trials (2.07%) compared with incongruent trials (4.68%), $F(1, 47) = 22.21$, $p < .001$, $\eta_p^2 = .32$, $BF_{01} = .00$. In addition, the magnitude of the Stroop effect was reduced for MI items (.87%) compared with MC items (4.34%), $F(1, 47) = 12.91$, $p < .001$, $\eta_p^2 = .22$, $BF_{01} = .03$.

Effects of WM Load. A repeated-measures ANOVA with factors of WM load, ISPC, and trial type was conducted with the mean percent error. Participants made more errors for incongruent (3.34%) trials than congruent (1.45%) trials, $F(1, 47) = 31.24$, $p < .001$, $\eta_p^2 = .40$, $BF_{01} = .00$. The error rate was not modulated by WM load, $F < 1$, $BF_{01} = 8.81$. The ISPC × Trial Type interaction was not significant, $F < 1$, $BF_{01} = 6.16$, indicating no the ISPC effect in error rate during the load blocks.[12] Furthermore, the three-way interaction was not significant, indicating that the ISPC effect was not different between the low and high WM trials, $F < 1$, $BF_{01} = 4.68$.[13]

Discussion

The magnitude of the ISPC effect was examined when a concurrent visuospatial WM load was manipulated. Consistent with previous experiments that investigated a verbal WM load (Spinelli et al., 2020), including Experiment 1, we found that the ISPC effect was not influenced by the visuospatial WM load. This finding rules out the alternative explanation of the main finding of Experiment 1

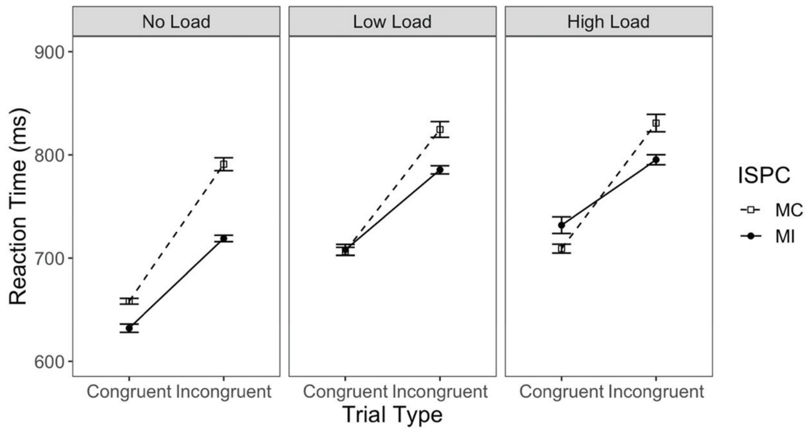
[10] Accordingly, unlike in Experiment 1, we did not conduct a *z*-score analysis.

[11] A 3 (WM Load: no load, low load, high load) × 2 (ISPC: MC, MI) × 2 (Trial Type: congruent, incongruent) repeated-measures ANOVA on RT showed the lack of three-way interactions (see Table S12 in the online supplemental materials for the full reports), suggesting that the magnitude of the ISPC effect in RT was not different between the WM load conditions, including the no load condition.

[12] It is not uncommon for the ISPC effect in accuracy to not be significant, including on diagnostic trials. This is a typical pattern that has been frequently reported in previous studies (Bugg et al., 2011, Bugg & Dey, 2018).

[13] A 3 (WM Load: no load, low load, high load) × 2 (ISPC: MC, MI) × 2 (Trial Type: congruent, incongruent) repeated-measures ANOVA on error rate revealed a significant three-way interaction (see Table S12 in the online supplemental materials for the full reports). This and the absence of the ISPC effect in error rate under WM load (as reported in the main text) together suggest that the robust ISPC effect in the no load block disappeared under low and high concurrent WM load.

Figure 3
Reaction Time as a Function of WM Load, ISPC, and Trial Type in Experiment 2



Note. ISPC = item-specific proportion congruence; MC = mostly congruent; MI = mostly incongruent; WM = working memory. Error bars depict one within-subject standard error.

which attributes the null effect of the load manipulation on the ISPC effect to the use of a verbal WM load, which we had reasoned might not share the same resources as those that process the critical information (i.e., visuospatial features of the animal categories) that triggers retrieval of item-specific control settings. Furthermore, the average WM task accuracy was lower in Experiment 2 (62% for high load) compared with Experiment 1 (85% for high load), suggesting that the null interaction reported in Experiment 1 was not simply attributable to the fact that the high verbal WM load was not demanding enough to add a bottleneck on ongoing cognitive processing. Therefore, the results of Experiment 2 provide additional support for the view that item-specific control is triggered and executed without attentional resources.

Clearly, we did not find any hint of an effect of the visuospatial WM load manipulation on the ISPC effect. One interesting observation, though, as is apparent from Table 3, is that the ISPC effect was present in error rate for the first (no load) block but the effect was much smaller in later blocks where the WM task was additionally performed (see the online supplemental materials for the full report). This might suggest that the presence of a concurrent visuospatial WM load weakened the ISPC effect in error rate. However, because the design was optimized for comparing the low and high WM load conditions, we cannot rule out that the ISPC effect in error rate might simply have weakened with time on task.

Taken together, the evidence thus far suggests that item-specific control persists regardless of whether the concurrent WM load is verbal (Experiment 1) or visuospatial (Experiment 2). However, it is important to note that, in Experiments 1 and 2, the WM load manipulation involved varying the number of items (or number of spatial locations) maintained in WM (i.e., a Sternberg-like task). Therefore, the null interaction between WM load and the ISPC effect may indicate that the processes required to actively maintain items in WM (i.e., storage processes) might not be necessary to trigger retrieval and execution of item-specific control settings. These findings do not, however, rule out that WM updating processes may be necessary. Theoretically this seems like a valid alternative given that the ISPC effect depends on the flexible updating of control settings upon stimulus onset (e.g., quickly heighten attentional control when an MI item is shown but not an MC item).

Indeed, prior work has shown that the updating of information in WM and not the maintenance of information in WM interferes

Table 3
Mean Reaction Times (RTs) and Percent Errors for Congruent and Incongruent Trials for Mostly Congruent and Mostly Incongruent Items Under No Load, Low WM Load, and High WM Load in Experiment 2

		No load		Low load		High load	
ISPC	Trial type	*M* (*SD*)	% err (*SD*)	*M* (*SD*)	% err (*SD*)	*M* (*SD*)	% err (*SD*)
Mostly congruent	Congruent	659 (93)	1.96 (1.89)	712 (138)	1.56 (2.26)	713 (139)	1.60 (2.76)
	Incongruent	793 (109)	6.30 (6.71)	838 (167)	3.37 (4.66)	836 (159)	3.72 (5.27)
	Stroop effect	134	4.34	126	1.81	123	2.12
Mostly incongruent	Congruent	633 (82)	2.19 (3.24)	711 (121)	1.43 (2.97)	733 (153)	1.20 (3.56)
	Incongruent	720 (87)	3.06 (3.04)	793 (134)	3.05 (3.20)	800 (137)	3.21 (3.32)
	Stroop effect	87	0.87	82	1.62	67	2.01
ISPC effect		47	3.47	44	0.19	56	0.11

Note. ISPC = item-specific proportion congruence; WM = working memory.

with postconflict (reactive) adjustments in control in the form of the congruency sequence effect (Soutschek et al., 2013; see also Kalanthroff et al., 2015, who found that proactive control is detrimentally affected by demands on updating). To test this possibility, in Experiment 3, we examined the automaticity of reactive control as evidenced by the ISPC effect when the attentional system was occupied by a concurrent n-back task.

Experiment 3

It has been suggested that it is not the maintenance but rather the updating of information in WM that interferes with cognitive control (Kalanthroff et al., 2015; Soutschek et al., 2013). Of most relevance to the present study, Soutschek et al. (2013) reported that the congruency sequence effect in a Stroop task was impaired when participants were performing a concurrent 1-back or 2-back task (compared with a 0-back task) requiring participants to report whether the present letter was identical to the one that was presented in one or two trials back. The congruency sequence effect is the pattern whereby the Stroop effect is reduced following an incongruent trial compared with a congruent trial. Quite interestingly, and in contrast, they did not observe such impairment when participants were simply asked to maintain a verbal load in WM (as in our Experiment 1; although not the primary focus of their study, see also Moss et al., 2020, who found that congruency sequence effects were not detrimentally affected by a load imposed during the retention interval of a change detection task). The findings of Soutschek et al. raise the possibility that item-specific control could be interrupted when attentional resources are occupied by updating WM contents. To examine this possibility, we also adopted the n-back paradigm and used a concurrent 0-back task and 2-back task as low and high WM loads (Kalanthroff et al., 2015; Soutschek et al., 2013), respectively. We expected that the magnitude of the ISPC effect would be reduced in the high compared with the low load condition if updating WM contents shares key resources with the processes that support item-specific retrieval and execution of control settings. However, if item-specific control occurs in an automatic manner, there should be no effect of WM load on the magnitude of the ISPC effect.

Method

Participants

Forty-eight undergraduate students ($M_{\text{age}} = 19.31$, $SD_{\text{age}} = 1.17$, 35 female) from Washington University in St. Louis participated in the study to fulfill a credit as a partial requirement of a course. All participants were native English speakers with normal or corrected-to-normal vision.

Apparatus, Stimuli, and Procedure

Apparatus, stimuli, and procedure were identical to those used in Experiment 1 except the WM load manipulation. In Experiment 3, instead of the verbal WM task (i.e., digit memory), an *n*-back WM task was administered.

We employed the *n*-back task used by Kalanthroff et al. (2015) that was shown to disrupt proactive control in a Stroop task. A single letter, randomly selected from a set of B, D, G, P, and T, was presented at the beginning of each trial (see Figure 1D) for 1,150 ms. For the low WM (0-back task) load condition, participants were asked to press a key on the serial response box if the letter was the same as the one that they were assigned (e.g., the letter T) at the beginning of the experiment. The 0-back target letter was randomly assigned for each participant and remained the same for each participant throughout the experiment. For the high WM (2-back task) load, participants were asked to press a key if the letter that was presented two trials back was the same as the current letter.

Design

A 3 (WM load: no WM load vs. low WM load [0-back] vs. high WM load [2-back]) × 2 (ISPC: MC vs. MI) × 2 (trial type: congruent vs. incongruent) within-subject design was used. Participants completed 16 practice trials followed by 192 trials without the concurrent WM load (no WM load), just as in the previous experiments. Then, the concurrent n-back (i.e., 0-back or 2-back) task was introduced. In this experiment, WM load was manipulated between blocks following Kalanthroff et al. (2015; see also Soutschek et al., 2013), and the block order (low load first vs. high load first) was counterbalanced across participants (see Figure 1A). Each block consisted of 192 trials and was preceded by 16 practice trials that acquainted participants with the *n*-back task for that block. Accuracy feedback for the WM task (e.g., "Correct," "Incorrect") was presented during the practice blocks only. There was a brief rest at the halfway point of all blocks. Each participant completed 576 trials total which lasted approximately 1 hr.

Results

WM Task

The memory probe responses were significantly less accurate in the high WM task (2-back task; 86%) compared with the low WM task (0-back task; 100%), $t(47) = 14.92$, $p < .001$, Cohen's $d = 2.16$. The memory probe RT was significantly slower in the high WM task (606 ms) compared with the low WM task (514 ms), $t(47) = 6.66$, $p < .001$, Cohen's $d = .96$.

Stroop Task

The same trimming procedures were applied as in Experiment 1 (less than 1% of trials were removed for the RT trim), including the exclusion of incorrect trials and incorrect WM probe responses from the RT analysis.[14] Table 4 shows mean RT and error rate by condition.

Reaction Time

Manipulation Check: Item-Specific Learning Under No WM Load. A 2 (ISPC: MC vs. MI) × 2 (Trial Type: congruent vs. incongruent) repeated-measures ANOVA was conducted revealing that the congruent trials (674 ms) were faster than the incongruent trials (785 ms), $F(1, 47) = 250.78$, $p < .001$, $\eta_p^2 = .84$, $BF_{01} = .00$. The ISPC effect was significant, $F(1, 47) = 7.37$, $p = .009$, $\eta_p^2 = .14$, $BF_{01} = .31$, with a greater Stroop effect in MC items (129 ms) than MI items (94 ms).

[14] The same analysis on trials with incorrect WM probe responses yielded the same results. Further details can be found in the online supplemental materials.

Table 4

Mean Reaction Times (RTs) and Percent Errors for Congruent and Incongruent Trials for Mostly Congruent and Mostly Incongruent Items Under No Load, Low WM Load, and High WM Load in Experiment 3

		No load		Low load		High load	
ISPC	Trial type	*M* (*SD*)	% err (*SD*)	*M* (*SD*)	% err (*SD*)	*M* (*SD*)	% err (*SD*)
Mostly congruent	Congruent	673 (73)	4.47 (4.23)	710 (167)	3.17 (2.59)	824 (178)	2.95 (2.52)
	Incongruent	802 (109)	6.75 (5.97)	816 (173)	5.38 (5.36)	964 (185)	6.66 (5.46)
	Stroop effect	129	2.28	106	2.21	140	3.71
Mostly incongruent	Congruent	674 (90)	4.20 (5.69)	703 (170)	2.87 (4.05)	817 (179)	2.85 (3.76)
	Incongruent	768 (101)	5.42 (4.13)	777 (160)	3.56 (3.18)	898 (182)	4.45 (3.03)
	Stroop effect	94	1.22	74	0.69	81	1.60
ISPC effect		35	1.06	32	1.52	59	2.11

Note. ISPC = item-specific proportion congruence; WM = working memory.

Effects of WM Load. A repeated-measures ANOVA was conducted as a function of WM load (low WM load vs. high WM load), ISPC (MC vs. MI), and trial type (congruent vs. incongruent). The main effect of trial type was significant, $F(1, 47) = 138.13$, $p < .001$, $\eta_p^2 = .75$, $BF_{01} = .00$, confirming a typical Stroop effect featuring faster RT in congruent trials (763 ms) compared with incongruent trials (864 ms). Importantly, we also found a significant main effect of WM load, $F(1, 47) = 95.28$, $p < .001$, $\eta_p^2 = .67$, $BF_{01} = .00$, indicating that marginal RT was slower in the high WM load block (876 ms) compared with the low WM load block (752 ms).

A significant Trial Type × ISPC interaction confirmed the ISPC effect, $F(1, 47) = 22.75$, $p < .001$, $\eta_p^2 = .33$, $BF_{01} = .55$, with a larger Stroop effect in MC (124 ms) compared with the MI items (77 ms). Most importantly, the three-way interaction was again not significant, $F(1, 47) = 3.38$, $p = .07$, $\eta_p^2 = .07$, $BF_{01} = 3.39$, suggesting that the magnitude of the ISPC effect was not influenced by WM load (see Figure 4).[15]

***Z*-Transformed RT.** Individual's RT was *z*-scored based on WM load condition to examine the effect of WM load on item-specific control on the baseline corrected RT. The main effect of trial type, $F(1, 47) = 193.70$, $p < .001$, $\eta_p^2 = .80$, $BF_{01} = .00$, was highly significant while the main effect of WM load was not (as expected given the transformation), $F(1, 47) = 1.77$, $p = .19$, $\eta_p^2 = .04$, $BF_{01} = 8.69$. The ISPC by trial type interaction remained significant, $F(1, 47) = 31.16$, $p < .001$, $\eta_p^2 = .40$, $BF_{01} = .00$, featuring the same pattern as reported in the raw RT analysis. Importantly, the three-way interaction also remained nonsignificant, $F(1, 47) = 1.89$, $p = .18$, $\eta_p^2 = .04$, $BF_{01} = 3.32$.

Error Rate

Manipulation Check: Item-Specific Learning Under No WM Load. A 2 (ISPC: MC vs. MI) × 2 (Trial Type: congruent vs. incongruent) repeated-measures ANOVA revealed that the error rate was higher for incongruent trials (6.09%) than congruent trials (4.33%), $F(1, 47) = 9.60$, $p = .003$, $\eta_p^2 = .17$, $BF_{01} = .06$. However, the Item-Specific PC × Trial Type interaction did not reach statistical significance, $F(1, 47) = 1.14$, $p = .29$, $\eta_p^2 = .02$, $BF_{01} = 2.50$.

Effects of WM Load. A repeated-measures ANOVA as a function of WM load (low load [0-back] vs. high load [2-back]), ISPC (MC vs. MI), and trial type (congruent vs. incongruent) was conducted on the percent error. The main effect of trial type was significant, $F(1, 47) = 24.87$, $p < .001$, $\eta_p^2 = .35$, $BF_{01} = .00$, showing that participants made more errors for the incongruent (5.01%) compared with the congruent (2.96%) trials. The main effect of WM load was not significant, $F(1, 47) = 1.29$, $p = .26$, $\eta_p^2 = .03$, $BF_{01} = 4.13$. The Stroop effect in error rate was greater for MC items (2.96%) compared with MI items (1.15%), $F(1, 47) = 8.44$, $p = .01$, $\eta_p^2 = .15$, $BF_{01} = .38$. Finally, the three-way interaction was again not significant, $F < 1$, $BF_{01} = 4.47$.

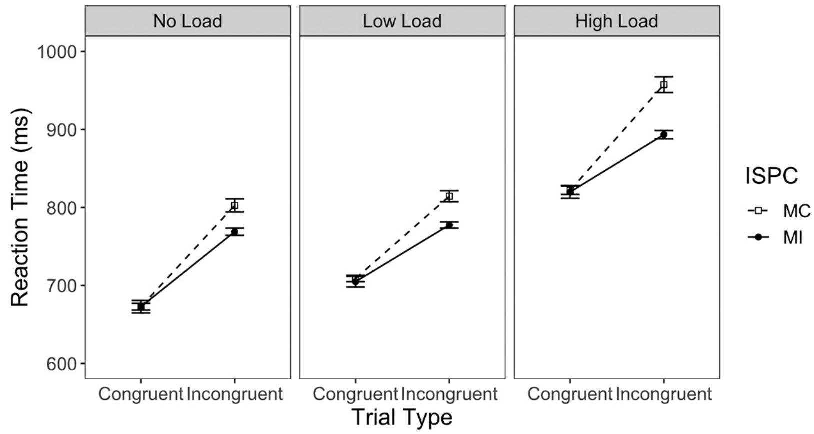
Discussion

The goal of Experiment 3 was to examine whether updating items in WM interferes with item-specific, reactive control. To do so, we adopted the n-back paradigm to modulate WM load. Therefore, participants had to update zero items under the low WM load and two items on every trial under the high WM load while they concurrently performed the Stroop task. The high WM load condition was more challenging than the low WM load condition, as indicated by accuracy and RT on the WM task, yet the magnitude of the ISPC effect was not modulated by WM load. A larger Stroop effect was found for MC (PC-75) items compared with MI (PC-25) items replicating a typical ISPC effect, in both conditions. The Bayesian analyses again indicated substantial evidence for the null effect in both RT and accuracy. Using the z-score analyses, we also demonstrated that the lack of an effect of WM load on the ISPC effect could not be explained by the baseline differences in RT for the WM load conditions. Our key finding replicated the results of Experiment 1 and 2 and strengthens the view that reactive control, as indicated by the ISPC effect, occurs automatically.

It is also notable that in addition to the Bayesian analyses, the pattern of results argues against an account suggesting the lack of a three-way interaction, and in particular a reduction in the ISPC effect with increasing load, reflects a type II error. Although the ISPC effect was statistically not different between the WM load conditions, the observed ISPC effect was in fact nominally *largest* for the high WM load (59 ms) condition (the no and low WM load conditions had ISPC effects of 35 ms and 32 ms, respectively).

[15] A 3 (WM Load: no load, low load, high load) × 2 (ISPC: MC, MI) × 2 (Trial Type: congruent, incongruent) repeated-measures ANOVA on RT, *z*-transformed RT, and error rate showed the lack of three-way interaction (see Table S13 in the online supplemental materials for the full reports), suggesting that the magnitude of the ISPC effect was not different between the WM load conditions, including the no load condition.

Figure 4
Reaction Time as a Function of WM Load, ISPC, and Trial Type in Experiment 3



Note. ISPC = item-specific proportion congruence; MC = mostly congruent; MI = mostly incongruent; WM = working memory. Error bars depict within-subject standard errors.

The discrepancy between the findings of Soutschek et al. (2013), and our results are interesting from a theoretical perspective especially given the fact that item-specific control, like sequential adjustments in control based on congruency, is considered to be a reactive form of control that is triggered post stimulus onset. We reserve further discussion of this issue for the General Discussion.

Experiment 4

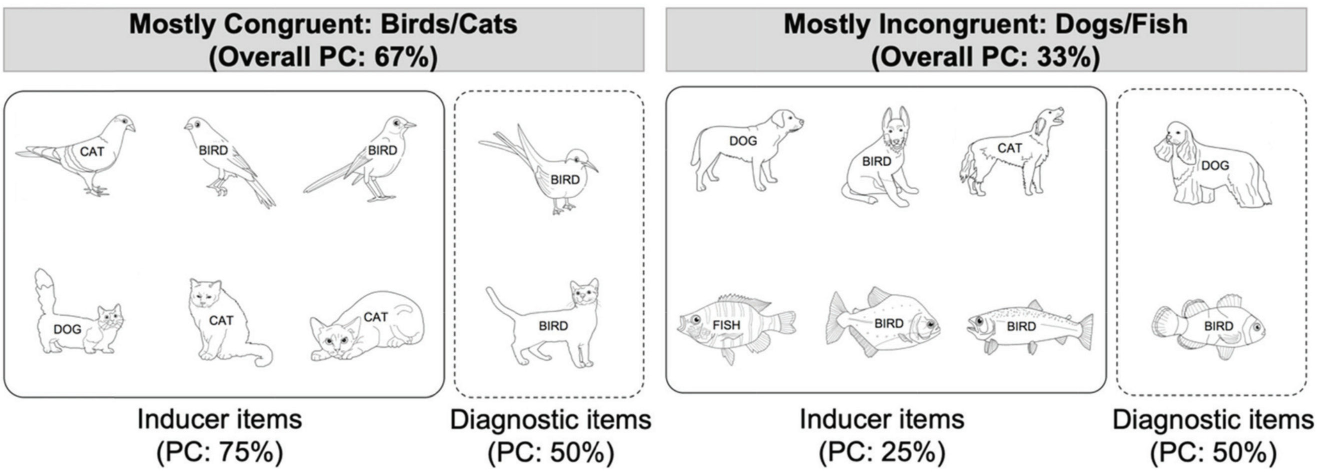
In Experiments 1–3, we found converging evidence suggesting that an item (e.g., animal picture) that has been associated with a particular level of conflict (MC or MI) continues to trigger a corresponding control setting (relaxing or heightening) regardless of whether WM resources are occupied by maintenance or updating of information. This suggests item-specific adjustments in attention, an indicator of reactive control, may be automatic with respect to the ability to operate in parallel with concurrent WM tasks.

In Experiment 4, we harness prior findings demonstrating "transfer" of reactive item-specific control to examine whether automaticity also transfers to novel items. That is, we examine whether the apparent automaticity of item-specific control even in the face of high updating demands (Experiment 3) extends beyond stimuli for which one has had prior direct experience acquiring and applying attentional control settings. Transfer is examined by investigating ISPC effects for novel, unbiased (PC-50) exemplars from trained ("inducer") categories (e.g., Bugg et al., 2011; Bugg & Dey, 2018; cf. Bugg & Hutchison, 2013). In the preceding experiments, the inducer items were the mostly congruent (PC-75; e.g., birds and cats) or mostly incongruent (PC-25; e.g., dogs and fish) animal pictures. In Experiment 4, the inducer items were again presented; however, in addition, transfer ("diagnostic") items were intermixed with inducer items. The diagnostic items comprised unique exemplars from the same animal categories as the inducer items (see Figure 5 for sample items) but diagnostic items were unbiased (PC-50) regardless of the animal category. Previous studies have shown that diagnostic items also show an ISPC effect such that a reduced Stroop effect is found for novel dogs and fish compared with novel birds and cats (although all of these items are 50% congruent), which is referred to as an ISPC *transfer effect*. This transfer effect has been observed reliably across many experiments and like the ISPC effect for inducer items, it presents as an asymmetrical interaction whereby the effect is mostly driven by differences on incongruent trials (slower for incongruent trials from MC condition than MI condition as in the preceding experiments within this series; see Bugg et al., 2011; Bugg & Dey, 2018; Gonthier et al., 2016). Critically, the transfer effect has proven to be highly robust to a variety of manipulations that were designed to disrupt the effect (e.g., familiarizing participants with all items prior to the Stroop task; individuation instructions; using diagnostic items that defy the PC of the inducer items; Bugg & Dey, 2018).

The critical question is whether retrieval and execution of item-specific control settings will continue to be automatic when a novel exemplar is presented that differs from the inducer items. In other words, is the automaticity of reactive control dependent on retrieval processes being able to make direct contact with prior perceptual experiences (i.e., prior presentations of stimuli [instances]; Logan, 1988) that supported the learning of item-control associations, or alternatively, will retrieval and execution of item-specific control for similar yet distinct items be vulnerable to a concurrent WM load?

As in Experiment 3, participants were asked to perform the picture-word Stroop task with a concurrent n-back task. We again used the *n*-back task to manipulate the WM load given prior evidence suggesting that the updating and not maintenance of WM contents interferes with one form of reactive control (sequential adjustments based on congruency; Soutschek et al., 2013). If the automatic retrieval of item-specific control settings depends on retrieval of learned item-control associations tied to direct prior experiences, we expect to see no WM load effect for the inducer items (replicating Experiment 3). However, in that case, load should affect performance on the diagnostic items such that the ISPC effect will be disrupted by high WM load compared with low WM load. On the other hand, if the automaticity of reactive

Figure 5
Examples of Stimuli Used for Inducer and Diagnostic Items in Experiment 4



Note. MC = mostly congruent; MI = mostly incongruent. Three of four exemplars were inducer items (PC-75/PC-25), and the remaining exemplar was a diagnostic item (PC-50). The pictures of birds/cats were MC and dogs/fish were MI for a half of participants (as described in this example); the other half of participants followed the opposite assignment.

control transfers to novel unbiased exemplars from the same categories as inducer items, then there should be no effect of WM load on either the inducer or diagnostic items.

Method

Participants

Forty-eight undergraduate students (*M* age = 19.44, *SD* = 1.41, 33 Female) from Washington University in St. Louis participated in the study to fulfill a credit as a partial requirement of a course. All participants were native English speakers with normal or corrected-to-normal vision.

Apparatus, Stimuli, and Procedure

Apparatus, stimuli, and procedure were identical to those of Experiment 3 except that diagnostic items (PC-50) were added in Experiment 4 (see Table 5 for the frequency of picture and word pairs). Consistent with Experiment 3, two animal categories were mostly congruent (e.g., birds/cats) and two animal categories were mostly incongruent (e.g., dogs/fish). For each animal category, three exemplars were PC biased (inducer), such that they were either mostly congruent (PC-75) or mostly incongruent (PC-25), and one exemplar was unbiased (diagnostic; PC-50). Each block included 144 inducer and 72 diagnostic items comprising 216 trials total. The inducer and diagnostic items when combined yielded 67% PC for MC items and 33% PC for MI items. The animal categories and PC mappings and exemplar type for the diagnostic item were counterbalanced across participants.

Design

A 3 (WM load: no WM load vs. low WM load (0-back) vs. high WM load (2-back)) × 2 (item type: inducer vs. diagnostic) × 2 (ISPC: MC vs. MI) × 2 (trial type: congruent vs. incongruent) within-subject design was used. The experimental design was identical to that of Experiment 3 except for the inclusion of diagnostic items. Because of the diagnostic items, participants completed 216 trials per each block instead of 192 trials (see Table 5).

Results

One participant was excluded from analysis due to excessive errors made on the memory probe task (above 50%). The other participant was excluded because the voice-key was unable to detect the participants' responses on most trials.

WM Task

The *n*-back memory probe accuracy was higher for the low WM load (0-back; 100%) compared with the high WM load (2-back; 87%) task, $t(45) = 12.25$, $p < .001$, d = 1.81. In addition, probe RTs were faster for the low WM load (525 ms) compared

Table 5
Frequency of Picture and Word Pairs in Each Block Used in Experiment 4

		Picture			
Item type	Word	Bird	Cat	Dog	Fish
Inducer	BIRD	*27*	3	9	9
	CAT	3	*27*	9	9
	DOG	3	3	*9*	9
	FISH	3	3	9	*9*
Diagnostic	BIRD	*9*	3	3	3
	CAT	3	*9*	3	3
	DOG	3	3	*9*	3
	FISH	3	3	3	*9*

Note. The italicized values indicate congruent pairs. Each frequency was derived from the repetition of three unique exemplars (i.e., for the BIRD–Bird inducer pair, each picture repeated nine times). The frequencies described here represent the one of two counterbalancing conditions where birds and cats are mostly congruent, and dogs and fish are mostly incongruent.

with the high WM load (616 ms), $t(45) = 6.28$, $p < .001$, Cohen's d = .93. Both patterns replicate Experiment 3.

Stroop Task

The same trimming procedures were applied as in the preceding experiments (less than 1% of trials were removed based on the RT trim). Again, only correct Stroop responses were included in the RT analysis. Consistent with Experiment 3, we excluded trials with incorrect WM probe responses.[16] We analyzed the inducer and diagnostic items separately as in prior studies (Bugg et al., 2011; Bugg & Dey, 2018).

Reaction Time

Inducer Items. Table 6 summarizes mean RT and error rate for each condition.

Manipulation Check: Item-Specific Learning Under No WM Load. A 2 (ISPC: MC vs. MI) × 2 (Trial Type: congruent vs. incongruent) repeated-measures ANOVA revealed that the congruent trials (652 ms) were faster than incongruent trials (758 ms), $F(1, 45) = 149.32$, $p < .001$, $\eta_p^2 = .77$, $BF_{01} = .00$. However, the ISPC × Trial Type interaction was not significant, $F(1, 45) = 2.65$, $p = .11$, $\eta_p^2 = .06$, $BF_{01} = 2.65$.

Effects of WM Load. A repeated-measures ANOVA with factors of WM load, ISPC, and trial type showed that the marginal RT was faster for congruent (767 ms) compared with incongruent (858 ms) trials, $F(1, 45) = 161.05$, $p < .001$, $\eta_p^2 = .78$, $BF_{01} = .00$. In addition, the main effect of WM load was significant, $F(1, 45) = 79.68$, $p < .001$, $\eta_p^2 = .64$, $BF_{01} = .00$, showing that marginal RT was slower under the high WM load (883 ms) compared with low WM load (742 ms).

The ISPC × Trial Type interaction was significant, $F(1, 45) = 11.60$, $p = .001$, $\eta_p^2 = .20$, $BF_{01} = 2.18$, reflecting a typical ISPC effect. The Stroop effect was 109 ms for MC items but it was reduced to 73 ms for MI items. As in Experiment 3, the three-way interaction revealed that the ISPC effect was not influenced by the concurrent WM load, $F(1, 45) = 1.82$, $p = .18$, $\eta_p^2 = .04$, $BF_{01} = 3.73$ (see Figure 6).[17] Regardless of the WM load, the Stroop effect was bigger for MC items compared with MI items showing a persistent ISPC effect.

Diagnostic Items. Table 7 summarizes mean RT and error rate for each condition. The same ANOVA analysis was conducted as for the inducer items.

Manipulation Check: Item-Specific Learning Under No WM Load. A 2 (ISPC: MC vs. MI) × 2 (Trial Type: congruent vs. incongruent) repeated-measures ANOVA revealed that the congruent trials (654 ms) were faster than incongruent trials (758 ms), $F(1, 45) = 119.42$, $p < .001$, $\eta_p^2 = .73$, $BF_{01} = .00$. However, the ISPC × Trial Type interaction was not significant, $F(1, 45) = 2.26$, $p = .14$, $\eta_p^2 = .05$, $BF_{01} = 2.44$.

We assumed that the nonsignificant ISPC effect observed for both inducer and diagnostic items in the initial (no WM load) block might be attributable to lower power relative to the preceding experiments since the number of observations per each condition was reduced (48 observations per each condition in Experiment 1–3 but 36 observations per each condition for inducer items and 18 observations per each condition for diagnostic items in Experiment 4) and the overall PC was attenuated (from 75% for MC and 25% for MI items in Experiment 1 – 3% to 67% for MC items 33% for MI items in Experiment 4). Therefore, we combined inducer and diagnostic items to further examine the ISPC effect during the initial no load block. A 2 (Item Type: inducer vs. diagnostic) × 2 (ISPC: MC vs. MI) × 2 (Trial Type: congruent vs. incongruent) repeated-measures ANOVA was conducted revealing a marginally significant ISPC × Trial Type interaction, $F(1, 45) = 3.86$, $p = .06$, $\eta_p^2 = .08$, $BF_{01} = 1.49$, with the Stroop effect being 114 ms for MC items compared with 96 ms for MI items. The Item Type × ISPC × Trial Type interaction was not significant, $F < 1$, $BF_{01} = 2.34$, indicating that the ISPC effect was not different between inducer and diagnostic items.

Effects of WM Load. The overall RT was faster for congruent (755 ms) compared with incongruent (863 ms) trials, $F(1, 45) = 108.12$, $p < .001$, $\eta_p^2 = .71$, $BF_{01} = .00$. In addition, the main effect of WM load was significant, $F(1, 45) = 83.44$, $p < .001$, $\eta_p^2 = .65$, $BF_{01} = .00$, demonstrating that the marginal RT was significantly slower in the high WM load block (881 ms) compared with the low WM load block (737 ms). There was not a significant ISPC transfer effect overall, $F(1, 45) = 1.15$, $p = .29$, $\eta_p^2 = .02$, $BF_{01} = 5.74$, but this was qualified by a significant three-way interaction, $F(1, 45) = 4.54$, $p = .04$, $\eta_p^2 = .09$, $BF_{01} = 1.71$, indicating that the ISPC transfer effect was modulated by concurrent WM load (see Figure 7).[18] To disentangle the three-way interaction, we examined the ISPC effect under low and high WM load separately using 2 (ISPC) × 2 (Trial Type) ANOVAs. For the low WM load, the Stroop effect was 118 ms for MC items and 77 ms for MI items resulting in a significant interaction, $F(1, 45) = 5.86$, $p = .02$, $\eta_p^2 = .12$, $BF_{01} = .47$, that replicated the classic ISPC transfer effect (e.g., Bugg et al., 2011; Bugg & Dey, 2018). However, for the high WM load, the Stroop effect was 113 ms for MC items and 125 ms for MI items. The interaction was not significant, $F < 1$, $BF_{01} = 4.29$, under high WM load, and was in fact in the reversed direction to the classic effect.

Z-Transformed RT

We also conducted repeated measures ANOVAs on *z*-transformed RT as a function of WM load, ISPC, and trial type on inducer and diagnostic items separately.

---

[16] The same analysis on trials with both correct and incorrect probe responses yielded the same results. Further details can be found in the online supplemental materials.

[17] A 3 (WM Load: no load, low load, high load) × 2 (ISPC: MC, MI) × 2 (Trial Type: congruent, incongruent) repeated-measures ANOVA on inducer RT, *z*-transformed RT, and error rate showed the lack of three-way interaction (see Table S14 in the online supplemental materials for the full reports), suggesting that the magnitude of the ISPC effect was not different between the WM load conditions, including the no load condition.

[18] A 3 (WM Load: no load, low load, high load) × 2 (ISPC: MC, MI) × 2 (Trial Type: congruent, incongruent) repeated-measures ANOVA on diagnostic RT and error rate showed that the three-way interaction was only significant for z-transformed RT, $F(1,45) = 3.97$, p = .02, $\eta_p^2 = 0.08$, $BF_{01} = 0.33$. The omnibus ANOVA results are available in the online supplemental materials (Table S15). To disentangle the three-way interaction for z-transformed RT, we conducted separate follow-up ANOVAs, which showed that the ISPC effect was not different between the no and low load, $F(1,45) = 3.54$, $p = .07$, $\eta_p^2 = 0.07$, $BF_{01} = 0.81$, or between the no and high load, $F < 1$, $BF_{01} = 3.29$. Rather, as reported in the manuscript the interaction was driven by a difference between the low and high load conditions.

Table 6
Mean Reaction Times (RTs) and Percent Errors of Inducer Items as a Function of the WM Load, ISPC, and Trial Type in Experiment 4

		No load		Low load		High load	
ISPC	Trial type	*M* (*SD*)	% err (*SD*)	*M* (*SD*)	% err (*SD*)	*M* (*SD*)	% err (*SD*)
Mostly congruent	Congruent	656 (82)	0.16 (0.53)	696 (104)	0.58 (1.46)	838 (164)	1.01 (1.46)
	Incongruent	771 (130)	1.47 (2.76)	796 (143)	2.22 (3.65)	955 (188)	5.58 (8.81)
	Stroop effect	115	1.31	100	1.64	117	4.57
Mostly incongruent	Congruent	648 (90)	0.61 (1.77)	698 (121)	0.62 (2.14)	835 (162)	0.70 (2.38)
	Incongruent	745 (107)	2.05 (2.77)	776 (129)	1.56 (1.90)	903 (160)	3.45 (4.97)
	Stroop effect	97	1.44	78	0.94	68	2.75
ISPC effect		18	−0.13	22	0.70	49	1.82

Note. ISPC = item-specific proportion congruence; WM = working memory.

Inducer Items. The main effect of trial type, $F(1, 45) = 236.02$, $p < .001$, $\eta_p^2 = .84$, $BF_{01} = .00$, was highly significant while the main effect of WM load was not, $F(1, 45) = 1.59$, $p = 21$, $\eta_p^2 = .03$, $BF_{01} = 8.26$. The item-specific PC by trial type interaction remained significant, $F(1, 45) = 12.63$, $p < .001$, $\eta_p^2 = .22$, $BF_{01} = .04$, featuring the same pattern as reported in the raw RT analysis. Finally, the three-way interaction remained nonsignificant, $F(1, 45) = 1.10$, $p = .30$, $\eta_p^2 = .02$, $BF_{01} = 3.34$.

Diagnostic Items. The main effect of trial type, $F(1, 45) = 249.94$, $p < .001$, $\eta_p{}^2 = .85$, $BF_{01} = .00$, but not the effect of WM load, $F < 1$, $BF_{01} = 8.80$, was significant (as expected given the transformation). The two-way interaction of ISPC × Trial Type, $F(1, 45) = 5.12$, $p = .03$, $\eta_p^2 = .10$, $BF_{01} = .70$, was significant. The three-way interaction was highly significant, $F(1, 45) = 7.11$, $p = .01$, $\eta_p^2 = .14$, $BF_{01} = .69$, confirming the notion that the ISPC transfer effect was significantly attenuated under the concurrent high WM load.

Error Rate

Inducer Items.

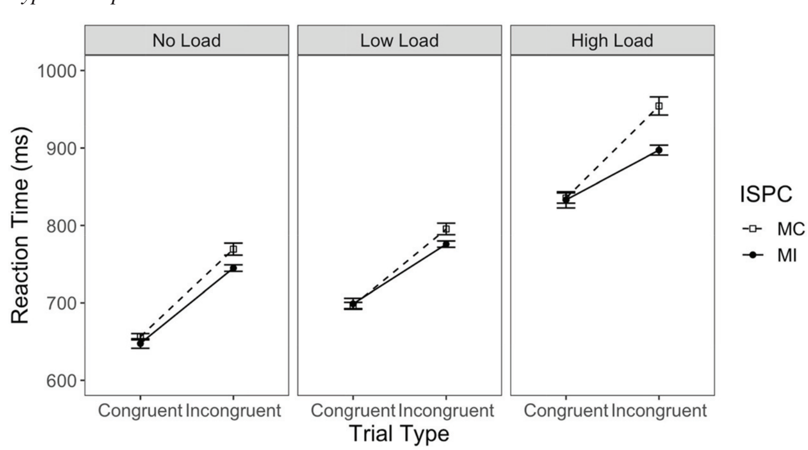
Manipulation Check: Item-Specific Learning Under No WM Load. A repeated-measures ANOVA with ISPC (MC vs. MI) and trial type (congruent vs. incongruent) revealed that the participants made more errors on incongruent (1.76%) compared with congruent (.39%) trials, $F(1, 45) = 20.53$, $p < .001$, $\eta_p^2 = .31$, $BF_{01} = .00$. However, the ISPC × Trial Type interaction was not significant, $F < 1$, $BF_{01} = 4.21$.

Effects of WM Load. A repeated-measures ANOVA as a function of WM load (low load [0-back] vs. high load [2-back]), ISPC (MC vs. MI), and trial type (congruent vs. incongruent) was conducted on the error rate (%). The error rate was higher for incongruent (3.20%) compared with congruent (.73%) trials, $F(1, 45) = 21.93$, $p < .001$, $\eta_p^2 = .33$, $BF_{01} = .00$. It was also influenced by the concurrent WM load, $F(1, 45) = 19.18$, $p < .001$, $\eta_p^2 = .30$, $BF_{01} = .03$, with a higher error rate under high load (2.68%) than low load (1.24%). Again, the Stroop effect for MC items (3.67%) was greater than that of MI items (1.29%), $F(1, 45) = 3.13$, $p = .08$, $\eta_p^2 = .07$, $BF_{01} = 2.03$. Most importantly, the three-way interaction was not significant, $F < 1$, $BF_{01} = 4.46$, indicating that the item-specific PC effect was equivalent regardless of the concurrent WM load.

Diagnostic Items.

Manipulation Check: Item-Specific Learning Under No WM Load. A repeated-measures ANOVA with ISPC (MC vs. MI) and trial type (congruent vs. incongruent) revealed that the

Figure 6
Reaction Time for Inducer Items as a Function of WM Load, ISPC, and Trial Type in Experiment 4



Note. ISPC = item-specific proportion congruence; MC = mostly congruent; MI = mostly incongruent; WM = working memory. Error bars depict one within-subject standard error.

participants made more errors on incongruent (2.64%) compared with congruent (.31%) trials, $F(1, 45) = 23.15$, $p < .001$, $\eta_p^2 = .34$, $BF_{01} = .00$. However, the ISPC × Trial Type interaction was not significant, $F < 1$, $BF_{01} = 4.46$.

We again combined inducer and diagnostic items to examine the pooled ISPC effect in the first (no WM load) block with more observations. A repeated measures ANOVA with item type (inducer vs. diagnostic), ISPC (MC vs. MI) and trial type (congruent vs. incongruent) was conducted revealing that error rate was higher in incongruent trials (2.20%) compared with congruent (.35%) trials, $F(1, 45) = 94.75$, $p < .001$, $\eta_p^2 = .68$, $BF_{01} = .00$. The ISPC × Trial Type interaction was not significant, $F < 1$, $BF_{01} = 6.77$. The three-way interaction was not significant either, $F < 1$, $BF_{01} = 4.31$.

Effects of WM Load. The marginal error rate was greater for incongruent (3.18%) compared with congruent (.42%) trials, $F(1, 45) = 28.75$, $p < .001$, $\eta_p^2 = .39$, $BF_{01} = .00$. The error rate was influenced by the WM load, $F(1, 45) = 12.64$, $p < .001$, $\eta_p^2 = .22$, $BF_{01} = .06$, such that the error rate under the high WM load (2.47%) was higher than that of low WM load (1.13%). Consistent with the RT results, the ISPC effect was not observed in error rate, $F(1, 45) = .73$, $p = .40$, $\eta_p^2 = .02$, $BF_{01} = 5.25$. The three-way interaction was also not significant, $F < 1$, $BF_{01} = 4.76$.

Discussion

As in Experiment 3, the ISPC effect was observed regardless of whether the WM load was low or high for the inducer items. The results also mirrored those of Experiment 3 with respect to the tendency for a numerically larger ISPC effect under the high WM load (49 ms) compared with the low WM load (22 ms) condition. A difference from Experiment 3 was that the ISPC effect was not statistically reliable in the initial no load block although it followed the same pattern as the preceding experiments. Given that the results for the inducer items in Experiment 4 replicated those of Experiment 3, it does not seem this difference is consequential.

The novel finding in Experiment 4 was that the ISPC transfer effect for the diagnostic items was abolished under high WM load (−12 ms), whereas it was highly robust under the low WM load (41 ms). The significant disruption of the ISPC transfer effect under high load is striking considering past studies that have demonstrated the robustness of this effect and its apparent immunity to multiple manipulations designed to disrupt it (for several examples, see Bugg & Dey, 2018).

The nonsignificant ISPC effect for inducer and diagnostic items in the initial no load block indicates that learning of the item-control associations was weaker in Experiment 4 relative to Experiments 1 – 3. One possibility is that participants were still learning these associations by the end of the no load block. To test this possibility, we included half (first vs. second) as a factor in an additional analysis (see Footnote 19). The ISPC effect did not differ between the first and second half of the no load block.[19] This does not support the idea that participants were still learning the associations; we thus believe that the weaker ISPC effect in Experiment 4 compared with the earlier experiments primarily occurred because of the reduced overall proportion congruency bias (MC: PC 64%, MI: PC 33%) caused by the addition of the diagnostic items. Still, one might question whether the lack of an ISPC effect on diagnostic items under the high load was caused by insufficient initial learning of the item-control associations.

To address this question, we tested whether the order of the WM load manipulation (low load first vs. high load first) influenced the degree to which the ISPC effect for diagnostic items was affected by the high WM load. The idea was that participants who completed the low load first may have had more opportunity for additional learning of item-control associations, which may have made them less vulnerable to the high WM load in the subsequent block (compared with participants that completed the high load first). The results showed that the degree to which the diagnostic ISPC effect was disrupted by concurrent high WM load was not influenced by whether participants started with low or high WM load (see Footnote 19). Taken together, the findings from these analyses suggest that insufficient learning might not be the main cause of the selective impairment of the diagnostic ISPC effect under high WM load.

The primary take-home message of Experiment 4 is as follows: whereas reactive control may be automatic with respect to the retrieval of control settings for (inducer) items that supported learning of the item-control associations, this automaticity does not appear to transfer to novel, unbiased exemplars. We will reserve discussion of this interesting asymmetry for the General Discussion. First, we report an analysis of the combined data for inducer items across Experiments 1–4.

Combined Data Analysis

Across four experiments, we found that various WM load manipulations (storage and updating) did not affect reactive control, as evidenced by the persistence of the ISPC effect in the low and high load conditions following an initial learning block. Although our power analysis (see Experiment 1) indicated that our experiments were well-powered and in a number of experiments the ISPC effect was actually larger in the high WM load condition, we nonetheless thought it would be valuable to provide a more highly powered test of the effect of concurrent WM load on item-specific control by combining the data from the Stroop task in the load blocks in Experiments 1 through 4 (only inducer trials, which were included in all experiments, were included in this analysis). A mixed-design repeated measures ANOVA with a between-subjects factor of unique experiment number (4 levels) and within subjects factors of WM load (low WM load vs. high WM load), ISPC (MC vs. MI) and trial type (congruent vs. incongruent) was conducted.

[19] A 2 (Half: first, second) × 2 (ISPC: MC, MI) × 2 (Trial Type: congruent, incongruent) repeated-measures ANOVA on diagnostic RT in the no load block showed a lack of three-way interaction, $F < 1$, $BF_{01} = 5.00$, suggesting that the ISPC effect was not different between the first and second half of the block (25 ms for the first half and 15 ms for the second half). The mixed measures ANOVA as a function of the WM block order (low load first vs. high load first), WM load, ISPC, and trial type confirmed that the influence of the WM load on the ISPC effect was not modulated by the order, $F(1,44) < 1$, $BF_{01} = 4.26$. Under high WM load, the diagnostic ISPC effect was −6 ms ($MC_{Stroop} = 128$ ms, $MI_{Stroop} = 135$ ms) for those participants who started with the high WM load, immediately followed by the initial no load block, and it was −17 ms ($MC_{Stroop} = 95$ ms, $MI_{Stroop} = 122$ ms) for those who started with the low WM load. For completeness, we note that under low WM load, the diagnostic ISPC effect was 40 ms ($MC_{Stroop} = 101$ ms, $MI_{Stroop} = 61$ ms) for those who started with the low WM load and it was 40 ms ($MC_{Stroop} = 138$ ms, $MI_{Stroop} = 98$ ms) for those who started with the high WM load. We thank an anonymous reviewer for suggesting these analyses.

Table 7
Mean Reaction Times (RTs) and Percent Errors for Diagnostic Items as a Function of the WM Load, ISPC, and Trial Type in Experiment 4

		No load		Low load		High load	
ISPC	Trial type	*M* (*SD*)	% err (*SD*)	*M* (*SD*)	% err (*SD*)	*M* (*SD*)	% err (*SD*)
Mostly congruent	Congruent	657 (90)	0.38 (1.45)	685 (112)	0.50 (1.63)	828 (134)	2.26 (1.25)
	Incongruent	771 (135)	2.85 (3.90)	803 (145)	1.81 (3.11)	941 (180)	5.09 (6.49)
	Stroop effect	114	2.47	118	1.31	113	2.83
Mostly incongruent	Congruent	650 (86)	0.25 (1.18)	691 (108)	0.50 (2.05)	815 (147)	0.42 (1.61)
	Incongruent	746 (115)	2.44 (4.40)	768 (142)	1.71 (3.70)	940 (211)	4.12 (6.79)
	Stroop effect	96	2.19	77	1.21	125	3.7
ISPC transfer effect		18	0.28	41	0.10	−12	0.87

Note. ISPC = item-specific proportion congruence; WM = working memory.

The combined data showed that the main effects of WM load, $F(1, 186) = 173.21$, $p < .001$, $\eta_p{}^2 = .48$, $BF_{01} = .00$, and trial type on Stroop RT, $F(1, 186) = 603.98$, $p < .001$, $\eta_p^2 = .76$, $BF_{01} = .00$, were highly significant replicating the same pattern as reported in Experiments 1 – 4. The ISPC effect was significant, $F(1, 186) = 88.62$, $p < .001$, $\eta_p^2 = .32$, $BF_{01} = .00$. Lastly and most importantly, the three-way interaction remained not significant, $F(1, 186) = 2.14$, $p = .15$, $\eta_p^2 = .01$, $BF_{01} = 7.80$, and the BF value provided substantial evidence favoring the null (see Figure 8). The ISPC effect was 37 ms under the low WM load ($Stroop_{MC} = 109$ ms, $Stroop_{MI} = 72$ ms) and 49 ms under the high WM load ($Stroop_{MC} = 119$ ms, $Stroop_{MI} = 70$ ms).

The same mixed-design repeated-measures ANOVA with *z*-transformed RT revealed that the main effect of trial type, $F(1, 186) = 750.84$, $p < .001$, $\eta_p^2 = .00$, $BF_{01} = .00$, and ISPC, $F(1, 186) = 42.41$, $p < .001$, $\eta_p^2 = .19$, $BF_{01} = .00$, were significant. The main effect of WM load was not significant, $F < 1$, $BF_{01} = 17.53$. The ISPC effect was significant, $F(1, 186) = 96.66$, $p < .001$, $\eta_p^2 = .34$, $BF_{01} = .00$, replicating the pattern observed in raw RT data. Most importantly, the three-way interaction remained not significant, $F < 1$, $BF_{01} = 7.58$.

The combined error rate data showed that the main effect of WM load, $F(1, 186) = 9.14$, $p = .003$, $\eta_p^2 = .05$, $BF_{01} = .57$, ISPC, $F(1, 186) = 12.42$, $p < .001$, $\eta_p^2 = .06$, $BF_{01} = .02$, and trial type, $F(1, 186) = 73.84$, $p < .001$, $\eta_p^2 = .28$, $BF_{01} = .00$, were significant. The ISPC effect was also significant, $F(1, 186) = 11.78$, $p < .001$, $\eta_p^2 = .06$, $BF_{01} = .15$, showing that the Stroop effect for MC items (2.30%) was greater than that of MI items (1.24%). Finally, the three-way interaction remained not significant, $F < 1$, $BF_{01} = 7.13$, suggesting that the magnitude of the ISPC was not influenced by WM load.

The results of the combined analysis bolster our conclusion that the triggering of item-specific control for inducer items continues to operate efficiently even under WM load, which suggests that it may operate in an automatic manner. Also, the numerical trends from the combined analysis showed that the ISPC effect was even greater (though not significantly so) under the high WM load (49 ms) than low WM load (37 ms) condition. To further examine the robustness of this numerical trend, we tested it separately based on the broader category of WM tasks, again focusing the analysis on the inducer items.[20] We combined Experiments 1 and 2 where memory-items were maintained in WM (i.e., Sternberg-like memory task) and Experiments 3 and 4 where participants were constantly updating memory-items in WM (i.e., n-back memory task). The analysis on the combined data set from Experiments 1 and 2 showed that the three-way interaction was not significant, $F < 1$, $BF_{01} = 5.80$. In contrast, the analysis on the combined data set from Experiments 3 and 4 showed that the Load × ISPC × Trial Type interaction was significant, $F(1, 93) = 4.88$, $p = .03$, $\eta_p^2 = .05$, $BF_{01} = 2.30$, confirming the trend of a larger ISPC effect under high WM load compared with low WM load.

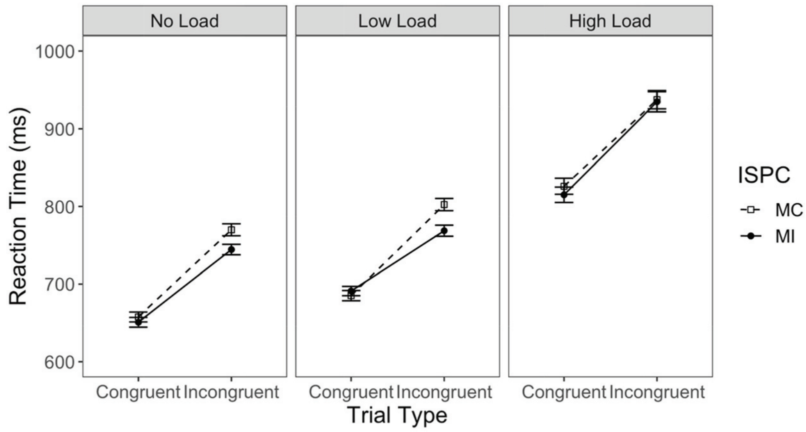
General Discussion

Traditional theories of cognitive control highlighted slow-acting, effortful, and strategic, mechanisms. Against this backdrop, the notion of "automatic control" as forwarded by Jacoby et al. (2003) as a potential explanation for the ISPC effect was clearly an oxymoronic concept (Bugg, 2015a). However, an emerging view reasons that cognitive control may also operate quickly and effortlessly with adjustments occurring outside of awareness, leaving open the possibility that cognitive control could, under certain conditions, be automatic. A key line of initial support for this view stemmed from studies demonstrating ISPC effects—the pattern whereby Stroop effects are reduced for mostly incongruent items compared with mostly congruent items—a key indicator of reactive control (Gonthier et al., 2016; see also evidence for CSPC effects, e.g., Crump & Milliken, 2009). The ISPC effect is a robust phenomenon that has been reproduced in many experiments and thus was ideally suited for investigating the present question of whether reactive control may be automatic with respect to the criterion of continuing to operate efficiently under load. Extending Spinelli et al. (2020), the present study systematically examined for the first time whether the triggering and execution of item-specific control in the overlapping sets design, a design that is confound-minimized relative to the two-item set design, is dependent on the amount of central cognitive resources as could be expected by traditional theories of control. To address this question, we exposed participants to the ISPC manipulation in an initial block so that they could learn the item-control associations. Then in subsequent blocks, we manipulated concurrent WM load to determine whether retrieval and execution of previously learned, item-specific control settings occurred efficiently even when WM load was high.

[20] We thank an anonymous reviewer for suggesting this analysis.

Figure 7

Reaction Time for Diagnostic Items as a Function of WM Load, ISPC, and Trial Type in Experiment 4

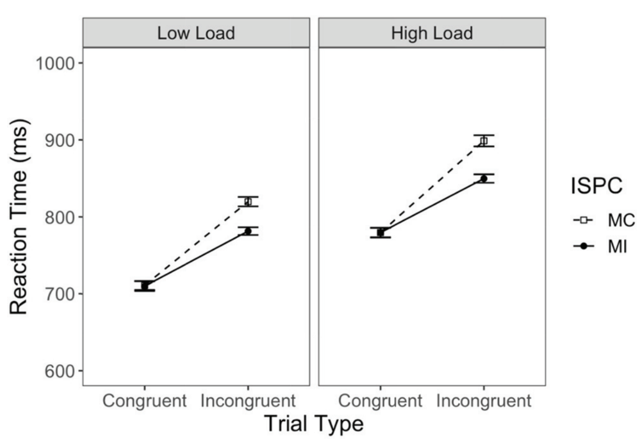


Note. ISPC = item-specific proportion congruence; MC = mostly congruent; MI = mostly incongruent; WM = working memory. Error bars depict one within-subject standard error.

In Experiments 1–3, we examined the effects of three different WM load manipulations on the ISPC effect while keeping other design and procedural features the same. In Experiment 1, we employed a verbal WM load requiring storage of digits (cf. Spinelli et al., 2020). In Experiment 2, we employed a visuospatial WM load requiring storage of spatial locations of visual objects. In Experiment 3, we switched gears and manipulated updating rather than maintenance demands, as inspired by previous findings showing other indices of control are detrimentally influenced by updating demands (proactive control; Kalanthroff et al., 2015; congruency sequence effects; Soutschek et al., 2013). Finally, in Experiment 4, we used the same updating manipulation but additionally tested diagnostic items to examine whether the automaticity of reactive control as evidenced in Experiments 1–3 for inducer (training) items transfers to diagnostic items that are novel, 50% congruent exemplars. The results of all four experiments are summarized in Figure 9.

Figure 8

Reaction Time of Combined Data From Inducer Trials as a Function of WM Load, ISPC, and Trial Type



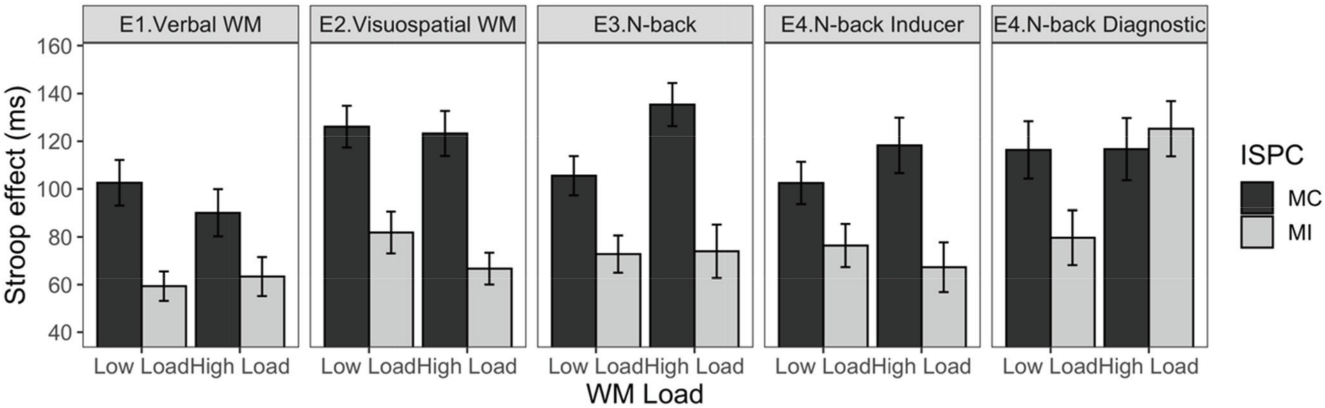
Note. ISPC = item-specific proportion congruence; MC = mostly congruent; MI = mostly incongruent; WM = working memory. Error bars depict one within-subject standard error.

There were two major findings that emerged from the present investigation. First, although the load manipulation was effective as indicated by slower and/or more errant performance when responding to the probes in the high load condition than the low load condition, we repeatedly observed across experiments that the magnitude of the ISPC effect (on inducer items) was not influenced by concurrent WM load. The ISPC effect was consistently robust under low and high load conditions. These findings support the concept of "automatic control" (Jacoby et al., 2003) and the view that reactive control continues to operate efficiently when a high load is imposed.

The current findings are consistent with and extend Spinelli et al. (2020). Several extensions are notable. First, our findings were observed using a confound-minimized ISPC design, and not one in which ISPC is perfectly confounded with contingency (thereby making it challenging to determine the contribution of control and contingency to observed effects). Second, the design of the present study enabled us to more precisely conclude that the triggering and execution of item-specific control (as opposed to learning of the item-control associations) is not dependent on WM maintenance or updating resources. In Spinelli et al. the WM load was imposed from the beginning of the ISPC paradigm such that any effect of load could have been attributable to the dependence of learning on WM resources rather than item-specific control per se. Third, the results also significantly extended their findings by showing that item-specific control remained robust regardless of the type of concurrent WM task. Unlike Spinelli et al., who used a verbal WM load across their experiments, we used verbal WM (Experiment 1), visuospatial WM (Experiment 2), and *n*-back updating (Experiment 3 and 4) loads, and all revealed the automatic

Figure 9

Stroop Effect for MC and MI Items as a Function of the WM Load in Experiments 1– 4



Note. ISPC = item-specific proportion congruence; MC = mostly congruent; MI = mostly incongruent; WM = working memory. The error bars depict one within-subject standard error.

nature of item-specific control (no effect of load on the ISPC effect) for inducer items, a pattern confirmed by our cross-experimental analysis.

The second major finding was that the ISPC effect for diagnostic[21] items (i.e., ISPC transfer effect) was abolished under high WM load but not low WM load. This is striking considering both the robustness of the ISPC effect for inducer items to the WM load manipulations in all of the present experiments, and the robustness of the ISPC effect and ISPC transfer effect to various manipulations designed to disrupt the effects in prior experiments (Bugg & Dey, 2018). An intriguing theoretical question is why performance on the diagnostic but not inducer items was susceptible to the load manipulation, that is, why the ISPC transfer effect does not appear to be governed by the same automatic processes as the ISPC effect. Here, we propose two potential explanations which we refer to as the *strength of learning hypothesis* and *abstraction hypothesis*. The strength of learning hypothesis posits that diagnostic items failed to trigger their corresponding control settings under high load because participants had less experience with diagnostic items (one presentation of a diagnostic item for every two presentations of an inducer item) and therefore less experience retrieving the requisite item-control associations in response to these items. This hypothesis fits with the theoretical possibility that reactive control, like other skills, may develop with experience (practice), with increases in skill learning corresponding to increases in automaticity (Shiffrin & Schneider, 1977), which could conversely explain why the ISPC effect for inducer items was highly robust to the load manipulation. However, the results of the additional analyses suggested by the reviewer (see Footnote 19) did not support the view that insufficient (weaker) learning was the primary cause of the disruption to the ISPC effect for diagnostic items under high load.

The abstraction hypothesis posits that the retrieval and execution of item-specific control settings is automatic only for items that supported learning of the item-control associations and not for similar but distinct items. The inducer items were mostly congruent or mostly incongruent. Experiences with these items thus afforded participants the opportunity to learn which animals were associated with a low probability of conflict and which were associated with a high probability of conflict, and importantly, the opportunity to adjust control accordingly directly in response to these items (i.e., relaxing or heightening control, respectively). Thus, these items enabled participants to directly learn item-control associations. In contrast, diagnostic items were 50% congruent regardless of the animal category—and therefore experiences with these items did not directly contribute to learning differences between mostly congruent and mostly incongruent items. Rather, participants had to generalize the item-control associations learned from the inducer items to the diagnostic items. Possibly this requires a degree of abstraction that is not supported by automatic control processes. In other words, an extra step may be needed to retrieve the indirectly associated control settings in the case of diagnostic items (e.g., a matching of the exemplar with the inducer exemplars; an overcoming of the initial tendency to retrieve a control setting representing a 50% item as opposed to a biased (67% or 33%) category), and this extra step may require WM resources.

Another interesting possibility is that the differential effects of the WM load manipulation on the ISPC effect for inducer and diagnostic items may indicate that representations of inducer and diagnostic items and their associated control settings were retrieved from different sources, long-term and WM, respectively. This would fit with findings from the visual search literature showing that targets that reappear across trials are initially held in WM but subsequently search is controlled by representations in long-term memory (Woodman et al., 2013). A *hybrid learning/abstraction hypothesis* could capture this distinction by positing that the relatively large amount of experience repeatedly retrieving and executing control settings for the inducer items led to the requisite representations (item-control associations) being stored in long-term memory whereas the representations for the diagnostic items were stored in WM. On this hybrid view, the diagnostic items were vulnerable to interference from a high WM load not solely because they required abstraction but also because representations

[21] Recall that diagnostic items refer to novel exemplars (i.e., distinct from inducer items) from MC or MI animal categories for which there was an equal proportion of congruent and incongruent trials.

of these items (including their item-control associations) required WM (similar to how representations of task sets have been proposed to be held in WM; see, e.g., Monsell, 2003; Oberauer et al., 2013).

Future research is needed to test these hypotheses. One potential strategy for doing so is inspired by several recent findings demonstrating a single prior experience with a control setting is sufficient to create an item-control association that can be retrieved after at least several intervening trials. Brosowsky and Crump (2018) showed that a control setting associated with a unique object (prime object) during a flanker task was later retrieved when the same object reappeared (probe object) even after a few hundred trials. More recently, Whitehead et al. (2020) extended this finding to a task-switching paradigm. Participants were presented with a picture of an object followed by either a task switch or a repetition of the same task. When the picture of the object was presented a second time two to seven trials later, switch-cost was reduced if the object was previously associated with a switch compared with when it was associated with a repeat of the task. Considered in conjunction with the present findings, an interesting and informative avenue for future research would be to examine the effects of WM load on item-specific control in such "one-shot" learning paradigms. The strength of learning hypothesis would predict that item-specific control that is based on a single experience with an inducer item should not be automatic and rather would be vulnerable to disruption when under load. However, if that single experience is stored in long-term memory, as certainly seems to be the case for effects observed by Brosowsky and Crump, then even control based on a single experience could be automatic according to the hybrid hypothesis. Investigations such as this would allow researchers to further develop theory regarding the automaticity of reactive control and define other potential boundary conditions for the automaticity of reactive control.

Different Reactive Control Mechanisms, Different Operating Characteristics?

The findings of Experiments 3 and 4 are also interesting in conjunction with earlier findings from Soutschek et al. (2013) in that they raise the possibility of another type of boundary condition for automatic control pertaining to the type of reactive control mechanism. In Soutschek et al. participants performed a color-word Stroop task with a concurrent n-back task (0-, 1-, and 2-back task). The congruency sequence effect was evident under 0-back task load, such that RT was faciliated following the previous incongruent trial. However, the effect disappeared under concurrent 1-back and 2-back task load. In Experiments 3 and 4 of the present study, we found that the ISPC effect was evident under 0-back task load *and* under concurrent 2-back task load (we did not test a 1-back task load as that would have greatly reduced observations in the other conditions). According to the dual mechanisms of control account (Braver et al., 2007), both congruency sequence effects and ISPC effects are indices of reactive control and therefore these contrasting patterns raise the question of whether different types of reactive control may have differing operating characteristics (see also Aschenbrenner & Balota, 2019, for evidence that the two effects are additive suggesting they represent independent processes), including the degree to which they operate automatically (i.e., efficiently under a concurrent load).

Essentially, the question is why the updating demands imposed by the *n*-back memory task interfered with the adjustments in control assessed by Soutschek et al. (2013) but not the adjustments in control assessed in the present study. The *n*-back task requires updating to-be-remembered information on a trial-by-trial basis, and in the case of the higher load conditions (1-back or 2-back) actively retaining information from at least one preceding trial. One possibility is that the n-back task interfered with the carry-over of congruency information or control from the preceding to the current trial, and thereby disrupted the congruency sequence effect (the intertrial relationship). In contrast, item-specific control in the present paradigm did not require retention of congruency information from the preceding trial. Participants flexibly retrieved the requisite control setting on each trial, but the ISPC effect was not dependent on carry over of control to the next trial. Another possibility invokes the hybrid hypothesis we discussed in the above section—sequential congruency effects may depend more on representations stored in WM compared with ISPC effects (at least for inducer items), which may depend more on representations stored in long-term memory.[22]

We can rule out the less interesting possibility that the congruency sequence effect was detrimentally affected by load (while the ISPC effect was not) because Soutschek et al. (2013) did not use a confound-minimized design. Their analyses controlled for feature (stimulus and response) repetitions, but the design did not control for contingency-learning contributions to the congruency sequence effect (Schmidt & Weissman, 2014) and prior research has shown that contingency-learning is detrimentally affected by WM load (Schmidt et al., 2010; Spinelli et al., 2020). Thus, one might suggest that the detrimental effects of load on the congruency sequence effect in Soutschek et al. reflect a contingency learning confound. Critically, however, if a contingency learning confound explained the Soutschek et al. (2013) findings, they should also have found that the load manipulation targeting WM maintenance in Experiment 1 of their study detrimentally affected the congruency sequence effect since this is the type of manipulation used by both Schmidt et al. (2010) and Spinelli et al. (2020). However, that manipulation did not disrupt the congruency sequence effect.

[22] To more directly compare the results of Soutschek et al. (2013) and ours, we calculated the congruency sequence effect by using the combined data sets from Experiment 3 and 4. As in Soutschek et al., trials with the feature repetition (e.g., a word/picture in previous trial repeated in the current trial) and previous incorrect responses were eliminated. Additionally, we aggregated the MC and MI items for this analysis. A 3 (WM Load: no vs. low vs. high load) × 2 (Prior Trial Type: congruent vs. incongruent) × 2 (Current Trial Type: congruent vs. incongruent) repeated-measures ANOVA on RT revealed a congruency sequence effect (i.e., Prior Trial Type × Current Trial Type interaction), $F(1,93) = 4.05$, $p = .05$, $\eta_p^2 = 0.04$, $BF_{01} = 7.00$, showing a smaller Stroop effect following a previous incongruent trial. When the congruency sequence effect was separately tested for each WM condition, it was marginally significant under no load, $F(1,93) = 3.75$, $p = .06$, $\eta_p^2 = .04$, $BF_{01} = 1.63$, but not significant under either low, $F(1,93) = 2.39$, $p = .13$, $\eta_p^2 = .03$, $BF_{01} = 3.60$, or high WM load, $F < 1$, $BF_{01} = 6.06$. Although these patterns generally mirror Soutschek et al., we did not find that the congruency sequence effect was significantly modulated by WM load, $F(2,186) = 0.27$, $p = .75$, $\eta_p^2 = .00$, $BF_{01} = 26.25$. However, the results must be interpretated with caution because our task was not designed to address this question (thus, 48% of trials were eliminated from the analysis based on the criteria above and cells were not equally balanced) and the ISPC manipulation might have hindered the pure sequential effects in this analysis.

Regardless of the precise explanation for the differing effects of WM load on the congruency sequence effect and the ISPC effect, the key point for present purposes is that not all forms of reactive control are necessarily automatic.

Limitations and Future Directions

In addition to pursuing the theoretical ideas discussed in the preceding sections, including studies that contrast the hypotheses we proposed, a number of other directions would paint an even richer picture of the efficiency of reactive control. The present findings in conjunction with several ERP and neuroimaging studies point to one such future direction. Briefly, prior studies have demonstrated neural underpinnings of ISPC effects early in visual processing (Blais et al., 2016; Shedden et al., 2013; cf. King et al., 2012). For example, using ERPs, Shedden et al. found that MC and MI items were differentiated quite early with differences observed in the N1 component (mean latency 150–175 ms after onset of stimulus) in the parieto-occipital area. This early segregation implies that early visual perception of an item may be the key signal that automatically triggers corresponding control settings (cf. Blais et al., 2016). This raises the interesting theoretical possibility that perceptual interference during the encoding of stimuli could prevent initiation of item-specific control.

Another future direction is to follow up on the overall trend of larger ISPC effects under higher WM load in the present study, which was a significant effect in the exploratory analysis targeting the combined data sets from Experiment 3 and 4. It is interesting to consider why this pattern emerged (see also Moss et al., 2020, for a similar pattern showing larger congruency sequence effects in error rate under high compared with low load). One possibility is that it may reflect a shift to more cost-saving processing (e.g., effort or other resources) under high load (see, e.g., Mäki-Marttunen et al., 2019, for supportive evidence in an AX-CPT task). This aligns with Spinelli and colleagues' (2020) view that reactive control may be "an even more convenient option than it is when [WM] resources are intact" (p. 19). However, it is important to clarify that we are not suggesting participants shift from purely proactive control in the low load (or even the no load condition) to reactive control in the high load condition–if participants engaged only proactive control in the low load (or no load) condition, we should not have observed ISPC effects. Rather, the suggestion is that a high load may have shifted participants' toward *even* greater use of reactive control. That is, in the low load condition participants may have occasionally tried to proactively filter the distracting words on some trials since they had available resources to do so. Since use of proactive control leads to weaker or absent ISPC effects, this may have produced nominally smaller ISPC effects in the low load condition.

Another possibility is that the tendency for the ISPC effect to be larger under high load may indicate that reinforcement learning was facilitated in that condition. This possibility is based on a consideration of two separate findings. First, the learned value of control model (Lieder et al., 2018), which is based on reinforcement learning principles, successfully captured the ISPC effect and other effects demonstrating that individuals learn to exert more control in response to certain features (i.e., features associated with MI items). Second, Collins et al. (2017) showed that the neural signatures of reinforcement learning were more potent under higher WM load. In their study, participants were shown unique objects one at a time and asked to press the correct key among three possible keys. The correct response for each object was not instructed to participants but feedback was given after each trial enabling participants to acquire the object-key mappings via trial and error (reinforcement learning processes). WM load was manipulated by varying the number of unique objects (six different objects under high load or one single object under low load) within a block. It was found that brain areas associated with reinforcement learning (e.g., striatum, left lateral prefrontal cortex, and parietal cortex) showed a stronger BOLD signal under the high load compared with low load condition. This suggests reinforcement learning signals were *more* robust when available central resources were scarce, and converges with a later finding showing reinforcement learning was enhanced in a high WM load condition (Collins et al., 2017). To the extent that item-specific control can be conceptualized as a process based on reinforcement learning (Lieder et al., 2018; but see Bejjani et al., 2020), it may be that the present trend of larger ISPC effects in the high load condition also reflects the interaction of reinforcement learning and WM.

The present study also had a few limitations that could be addressed in future studies. First, although we found participants were significantly less accurate and slower to respond to the WM task on high WM load compared with low WM load trials, it might be suggested that our participants represented higher WM capacity individuals that were, perhaps, immune to the effects of load on the ISPC effect. However, there is thus far little evidence suggesting that WM capacity modulates the ISPC effect (Hutchison, 2011; Spinelli et al., 2020) with the exception of Hutchison (2011) observing an ISPC effect in error rate selectively for the low WM capacity group (both low and high groups showed the effect in RT). Thus, if anything, we should have observed smaller ISPC effects in the present study if we had a nonrepresentative sample comprising only or mostly high WM capacity individuals. Still, future studies should examine the reproducibility of the present patterns in other populations. For example, it has been shown that older adults produce ISPC effects in the ISPC paradigm we employed that are similar in magnitude to younger adults' effects (i.e., Bugg, 2014). Thus, it would be interesting to examine whether a WM load manipulation disrupts ISPC effects for older adults. Such disruption may suggest that while older adults produce comparably sized ISPC effects to younger adults, the effect may not reflect automatic retrieval and execution of control settings.

A second limitation pertains to our decision to design our experiments to better isolate the retrieval and execution of item-specific control settings from the learning of item-control associations. Specifically, we had participants experience an initial block of ∼200 trials to learn the item-control associations before implementing the load manipulation. Consequently, this design did not enable us to examine whether such learning continues to occur efficiently under concurrent load and yet this learning is a critical component of reactive control. Thus, future studies should examine the degree to which load affects this component. Based on the findings of Spinelli et al. (2020), one might not expect load to affect learning; however, it is difficult to generalize their findings to the overlapping sets design used in the present study. Prior evidence from Jacoby et al. (2003) using the same design as Spinelli et al. found that learning of the item-response associations (i.e., contingency learning) occurs already within just 16 trials (Jacoby et al., 2003), thus not leaving much of a window for investigating effects of load on learning per se. An experiment that uses a similar design to the present one but breaks the initial block into several ISPC-matched mini blocks might enable examination of whether the time course of learning varies between conditions in which no load occurs in that first block relative to low or high WM load.

Lastly, we used a confound-minimized design (Braem et al., 2019) in the present study. Although this design is preferable to the confound-prone two-item set design (see Spinelli et al., 2020), this design does leave open the possibility that contingency learning (i.e., prediction of high contingency responses based on the word) could contribute to the ISPC effect because the words are 56% and 38% congruent for MC and MI items, respectively (though no study to date has established that this confound produces the ISPC effect in this design; Braem et al., 2019). As we have noted (Bugg & Dey, 2018) and demonstrated previously (e.g., Bugg et al., 2011; Experiment 2 vs. Experiment 3), the mere presence of the ISPC effect does not disambiguate the contingency and control accounts. What does do so, though, is the specific pattern that the ISPC effect (PC × Trial Type interaction) takes (see Bugg, 2015b for further discussion), and the ISPC pattern observed in the current experiments (and the prior experiments that used this design; Bugg et al., 2011; Bugg & Dey, 2018; Bugg & Hutchison, 2013) is not consistent with a contingency-learning mechanism. Specifically, contingency-learning should produce faster RTs for high-contingency compared with low-contingency cells—this prediction lies at the heart of the contingency-learning account of the ISPC effect (see Schmidt & Besner, 2008; for evidence supporting this prediction in the two-item set design). In the present design, this means that congruent trials for MC items should have been significantly faster than congruent trials for MI items. This prediction reflects that it is the congruent cells (MCC vs. MIC respectively) which differ most dramatically in contingency. This is because congruent word distractors (e.g., BIRD and CAT) appear much more often in congruent trials than in each possible incongruent trial for MC items (making the congruent response highly contingent) but for MI items the word distractors (e.g., FISH and DOG) are not more predictive of the congruent response.

To test this prediction in the present data, we examined whether the MCC cell yielded faster RTs than the MIC cell (i.e., we compared the congruent cells; see Figure S1, Table S16, and full report of analysis in the online supplemental materials). Inconsistent with a contingency learning account of the ISPC effect, the results showed that there was not a difference. In contrast, there was a significant difference between the MCI and MII cells (i.e., the incongruent cells) such that participants were faster for the MII cell which is consistent with the item-specific control account (Bugg, 2015b; Bugg et al., 2011; Bugg & Hutchison, 2013) and computational models of item-specific control (Blais et al., 2007; Verguts & Notebaert, 2008)) showing that item-specific control is conflict-dependent (i.e., triggered in the face of conflict on incongruent trials). The ISPC effect (i.e., interaction of PC and trial type), which considers movement in both congruent and incongruent trials, is thus clearly significant because of the effect on the incongruent trials and not the congruent trials. The absence of a difference between the MC and MI items on congruent trials (and the finding that for no-load trials, the difference between these trials is actually in the opposite direction than anticipated on a contingency-learning view[23]), a difference that is most clearly predicted by a word-based contingency-learning mechanism in the present design and a difference that *is* evident when the overlapping sets design is modified to bias participants to use contingency learning (by making the irrelevant word dimension the ISPC signal; see Bugg et al., 2011; Experiment 3), strongly suggests that contingency learning is not making a significant contribution to the ISPC effect in the present experiments.

Additionally, it is notable that the overlapping sets design used here where the words are 56% versus 38% congruent for the MC and MI items is precisely the design used by Bugg and Hutchison (2013; Experiment 2) for the inducer trials in their study. What differed was their approach to assessing transfer (i.e., the type of diagnostic item). To assess transfer, they presented *new* words in the trained colors in a final block of trials following learning of the ISPC effect (e.g., if the colors blue and red were MC and the colors white and green were MI, the inducer items comprised these trained colors and corresponding words [like the pictures and words in the present study]). The diagnostic items were also presented in blue, red, white, or green ink but the words differed from the inducer trials (e.g., YELLOW, PINK). The key finding was that participants were faster to name the color of these new words when they appeared in a MI ink color compared with an MC ink color. This demonstrated transfer of item-specific control from the inducer items to novel items comprising new words, which cannot be explained by any carry-over of contingency information from the word distractors on the inducer trials. Given we used exactly the same design for the inducer trials in the present study, this finding further supports an interpretation of the present ISPC effects based on item-specific control and not contingency learning. If the design encouraged contingency learning, Bugg and Hutchison should not have found transfer.

Finally, the contingency account also does not appear to be a plausible explanation of the current findings because in the two studies that have directly examined the effects of WM load on contingency learning (Schmidt et al., 2010; Spinelli et al., 2020), WM load significantly disrupted the contingency learning effect. This leads to the prediction that we should have found a significant disruption to the ISPC effect in the present study if the effect was driven by contingency learning—our results clearly did not support this prediction and, if anything, the trend was for a larger ISPC effect under higher WM load. One might then suggest that contingency learning could explain why the ISPC effect was disrupted under higher WM load for the diagnostic items but not the inducer items. However, the words were the same for inducer and diagnostic items (resulting in a similar predictiveness of the words when diagnostic trials are factored in 54% vs. 43% congruent for MC vs. MI), and these same contingencies were present under lower load where the ISPC effect *was* found for diagnostic items. Thus, a contingency account also does not appear to provide a coherent account of this pattern.

[23] In addition, we conducted the same analysis with trials under no load to rule out any contaminating influence of the concurrent WM load (Table S16 in the online supplemental materials). The difference between MCC and MIC cells in the no load block showed that RT was significantly slower for the MCC cell compared with the MIC cell (649 ms), which is opposite to what the contingency learning account predicts. The MCI cell (783 ms) was again significant slower than the MII cell (740 ms). The analysis is reported in the online supplemental materials.

In sum, although it is theoretically possible that the ISPC effect could in part reflect contingency learning in the confound-minimized design we employed, the evidence does not support this interpretation either in the present experiment or in past studies. Nonetheless, should better confound-minimized designs or confound-free designs (it remains to be determined whether this is possible given effects are rarely process pure) be developed and vetted that enable one to calculate an ISPC effect for inducer and diagnostic items, then it would be valuable to seek further converging evidence for the automaticity of item-specific control using such designs.[24]

Conclusion

In the present study, we investigated the automaticity of reactive control, namely the retrieval and execution of previous learned item-specific control settings. Across four experiments, we found consistent evidence that item-specific control continued to operate efficiently regardless of the concurrent WM load (low or high), type of WM contents (verbal or visuospatial), and the processing mode for WM contents (maintenance or updating). This pattern was observed for inducer items, which were directly associated with the differing, item-specific control settings (PC-25 or PC-75) via prior learning experiences. Quite interestingly, this pattern was not observed for diagnostic items—exemplars from the same category as the inducer items that were unbiased (PC-50) across animal categories. For the diagnostic items, the ISPC effect was observed in the low WM load condition but it was abolished under the high WM load condition, as shown in Experiment 4 where updating demands were manipulated. Together, these novel findings converge in suggesting that retrieval and execution of item-specific control settings is efficient regardless of the available central resources but only when the exact items are presented that supported the learning of the item-control associations. Our findings thus provide the first systematic evidence revealing the automatic nature of item-specific control including one boundary condition. Future research can inform the question of why inducer but not diagnostic items yield automatic control and why certain types of reactive control (i.e., ISPC effects) appear to be automatic while others (i.e., congruency-sequence effects) do not. In sum, the term "automatic control" (Jacoby et al., 2003) may not be entirely oxymoronic, and the present findings demonstrate that it is in fact possible to achieve high levels of cognitive control under conditions of high load, which may have practical implications for occupations such as air-traffic control.

[24] We thank an anonymous reviewer for encouraging us to address the role of contingency learning in the present designs.

References

Aschenbrenner, A. J., & Balota, D. A. (2019). Additive effects of item-specific and congruency sequence effects in the vocal Stroop task. *Frontiers in Psychology*, *10*, 860. https://doi.org/10.3389/fpsyg.2019.00860

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. *Unintended Thought*, *3*, 51–69.

Bargh, J. A. (1994). The four horsemen of automaticity: Intention, awareness, efficiency, and control as separate issues. In R. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 1–40). Erlbaum.

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*(4), 369–387. https://doi.org/10.1037/xhp0000720

Blais, C., Harris, M. B., Guerrero, J. V., & Bunge, S. A. (2012). Rethinking the role of automaticity in cognitive control. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *65*(2), 268–276. https://doi.org/10.1080/17470211003775234

Blais, C., Hubbard, E., & Mangun, G. R. (2016). ERP evidence for implicit priming of top–down control of attention. *Journal of Cognitive Neuroscience*, *28*(5), 763–772. https://doi.org/10.1162/jocn_a_00925

Blais, C., Robidoux, S., Risko, E. F., & Besner, D. (2007). Item-specific adaptation and the conflict-monitoring hypothesis: A computational model. *Psychological Review*, *114*(4), 1076–1086. https://doi.org/10.1037/0033-295X.114.4.1076

Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J. C., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring adaptive control in conflict tasks. *Trends in Cognitive Sciences*, *23*(9), 769–783. https://doi.org/10.1016/j.tics.2019.07.002

Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance* (pp. 713–737). MIT Press.

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 76–106). Oxford University Press.

Brosowsky, N. P., & Crump, M. J. C. (2018). Memory-guided selective attention: Single experiences with conflict have long-lasting effects on cognitive control. *Journal of Experimental Psychology: General*, *147*(8), 1134–1153. https://doi.org/10.1037/xge0000431

Bugg, J. M. (2012). Dissociating levels of cognitive control: The case of Stroop interference. *Current Directions in Psychological Science*, *21*(5), 302–309. https://doi.org/10.1177/0963721412453586

Bugg, J. M. (2014). Evidence for the sparing of reactive cognitive control with age. *Psychology and Aging*, *29*(1), 115–127. https://doi.org/10.1037/a0035270

Bugg, J. M. (2015a). "Automatic control" of Stroop interference: Evidence and implications for aging. In D. S. Lindsay, C. M. Kelley, A. P. Yonelinas, & H. L. Roediger, III, (Eds.), *Remembering: Attributions, processes, and control in human memory: Essays in honor of Larry Jacoby* (pp. 169–184). Psychology Press.

Bugg, J. M. (2015b). 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. https://doi.org/10.3758/s13414-014-0752-x

Bugg, J. M. (2017). Context, conflict, and control. In T. Egner (Ed.), *The Wiley handbook of cognitive control* (pp. 79–96). Wiley.

Bugg, J. M., & Chanani, S. (2011). List-wide control is not entirely elusive: Evidence from picture–word Stroop. *Psychonomic Bulletin & Review*, *18*(5), 930–936. https://doi.org/10.3758/s13423-011-0112-y

Bugg, J. M., & Crump, M. J. C. (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., & 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. https://doi.org/10.1037/xhp0000580

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*(2), 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*(3), 844–859. https://doi.org/10.1037/a0019957

Chiu, Y., Jiang, J., & Egner, T. (2017). The caudate nucleus mediates learning of stimulus-control state associations. *The Journal of Neuroscience*, *37*(4), 1028–1038. https://doi.org/10.1523/JNEUROSCI.0778-16.2017

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review*, *99*(1), 45–77. https://doi.org/10.1037/0033-295X.99.1.45

Collins, A. G. E., Albrecht, M. A., Waltz, J. A., Gold, J. M., & Frank, M. J. (2017). Interactions among working memory, reinforcement learning, and effort in value-based choice: A new paradigm. *Biological Psychiatry*, *82*(6), 431–439. https://doi.org/10.1016/j.biopsych.2017.05.017

Collins, A. G. E., Ciullo, B., Frank, M. J., & Badre, D. (2017). Working memory load strengthens reward prediction errors. *The Journal of Neuroscience*, *37*(16), 4332–4342. https://doi.org/10.1523/JNEUROSCI.2700-16.2017

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review*, *8*(2), 331–335. https://doi.org/10.3758/BF03196169

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, *123*(4), 354–373. https://doi.org/10.1037/0096-3445.123.4.354

Corballis, P. M., & Gratton, G. (2003). Independent control of processing strategies for different locations in the visual field. *Information Processing and Error Analysis: Retrospectives on a Career*, *64*(1), 191–209. https://doi.org/10.1016/S0301-0511(03)00109-1

Crump, M. J. C., Gong, Z., & Milliken, B. (2006). The context-specific proportion congruent Stroop effect: Location as a contextual cue. *Psychonomic Bulletin & Review*, *13*(2), 316–321. https://doi.org/10.3758/BF03193850

Crump, M. J. C., & Milliken, B. (2009). Short article: The flexibility of context-specific control: Evidence for context-driven generalization of item-specific control settings. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *62*(8), 1523–1532. https://doi.org/10.1080/17470210902752096

De Pisapia, N., & Braver, T. S. (2006). A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions. *Neurocomputing*, *69*(10-12), 1322–1326. https://doi.org/10.1016/j.neucom.2005.12.100

Diede, N. T., & Bugg, J. M. (2017). Cognitive effort is modulated outside of the explicit awareness of conflict frequency: Evidence from pupillometry. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(5), 824–835. https://doi.org/10.1037/xlm0000349

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 44, pp. 145–199). Elsevier Science.

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

Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive control in the Stroop task. *Memory & Cognition*, *44*(5), 778–788. https://doi.org/10.3758/s13421-016-0591-1

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*(4), 851–860. https://doi.org/10.1037/a0023437

Jacoby, L. L., Lindsay, D. S., & Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin & Review*, *10*(3), 638–644. https://doi.org/10.3758/BF03196526

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

Kalanthroff, E., Avnit, A., Henik, A., Davelaar, E. J., & Usher, M. (2015). Stroop proactive control and task conflict are modulated by concurrent working memory load. *Psychonomic Bulletin & Review*, *22*(3), 869–875. https://doi.org/10.3758/s13423-014-0735-x

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 336–358. https://doi.org/10.1037/0278-7393.26.2.336

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*(1), 47–70. https://doi.org/10.1037/0096-3445.132.1.47

Kim, S.-Y., Kim, M.-S., & Chun, M. M. (2005). Concurrent working memory load can reduce distraction. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(45), 16524–16529. https://doi.org/10.1073/pnas.0505454102

King, J. A., Korb, F. M., & Egner, T. (2012). Priming of control: Implicit contextual cuing of top-down attentional set. *The Journal of Neuroscience*, *32*(24), 8192–8200. https://doi.org/10.1523/JNEUROSCI.0934-12.2012

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

Lieder, F., Shenhav, A., Musslick, S., & Griffiths, T. L. (2018). Rational metareasoning and the plasticity of cognitive control. *PLoS Computational Biology*, *14*(4), e1006043 https://doi.org/10.1371/journal.pcbi.1006043

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492–527. https://doi.org/10.1037/0033-295X.95.4.492

Mäki-Marttunen, V., Hagen, T., & Espeseth, T. (2019). Task context load induces reactive cognitive control: An fMRI study on cortical and brain stem activity. *Cognitive, Affective & Behavioral Neuroscience*, *19*(4), 945–965. https://doi.org/10.3758/s13415-019-00691-6

Melnikoff, D. E., & Bargh, J. A. (2018). The mythical number two. *Trends in Cognitive Sciences*, *22*(4), 280–293. https://doi.org/10.1016/j.tics.2018.02.001

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*(2), 297–326. https://doi.org/10.1037/0033-2909.132.2.297

Moss, M. E., Kikumoto, A., & Mayr, U. (2020). Does conflict resolution rely on working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(12), 2410–2426. https://doi.org/10.1037/xlm0000801

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (pp. 1–18). Springer. https://doi.org/10.1007/978-1-4757-0629-1_1

O'Reilly, R. C., Braver, T. S., & Cohen, J. D. (1999). A biologically based computational model of working memory. In A. Miyake & P. Shah (Ed.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 375–411). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.014

Oberauer, K. (2019). Working memory and attention—A conceptual analysis and review. *Journal of Cognition*, *2*(1), 36. https://doi.org/10.5334/joc.58

Oberauer, K., Souza, A. S., Druey, M. D., & Gade, M. (2013). Analogous mechanisms of selection and updating in declarative and procedural working memory: Experiments and a computational model. *Cognitive Psychology*, *66*(2), 157–211. https://doi.org/10.1016/j.cogpsych.2012.11.001

Park, S., Kim, M.-S., & Chun, M. M. (2007). Concurrent working memory load can facilitate selective attention: Evidence for specialized load. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1062–1075. https://doi.org/10.1037/0096-1523.33.5.1062

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium*. Erlbaum.

Pratte, M. S., Rouder, J. N., Morey, R. D., & Feng, C. (2010). Exploring the differences in distributional properties between Stroop and Simon effects using delta plots. *Attention, Perception & Psychophysics*, *72*(7), 2013–2025. https://doi.org/10.3758/APP.72.7.2013

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General*, *123*(4), 374–393. https://doi.org/10.1037/0096-3445.123.4.374

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. https://doi.org/10.1037/0278-7393.34.3.514

Schmidt, J. R., Houwer, J. D., & Besner, D. (2010). Contingency learning and unlearning in the blink of an eye: A resource dependent process. *Consciousness and Cognition*, *19*(1), 235–250. https://doi.org/10.1016/j.concog.2009.12.016

Schmidt, J. R., & Weissman, D. H. (2014). Congruency sequence effects without feature integration or contingency learning confounds. *PLoS ONE*, *9*(7), e102337. https://doi.org/10.1371/journal.pone.0102337

Shedden, J. M., Milliken, B., Watter, S., & Monteiro, S. (2013). Event-related potentials as brain correlates of item specific proportion congruent effects. *Consciousness and Cognition*, *22*(4), 1442–1455. https://doi.org/10.1016/j.concog.2013.10.002

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127–190. https://doi.org/10.1037/0033-295X.84.2.127

Soutschek, A., Strobach, T., & Schubert, T. (2013). Working memory demands modulate cognitive control in the Stroop paradigm. *Psychological Research*, *77*(3), 333–347. https://doi.org/10.1007/s00426-012-0429-9

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*(11), 2007–2033. https://doi.org/10.1037/xlm0000934

Spinelli, G., & Lupker, S. J. (2020). 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*(3), 426–435. https://doi.org/10.3758/s13421-019-00980-y

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*(3736), 652–654. https://doi.org/10.1126/science.153.3736.652

Verguts, T., & Notebaert, W. (2008). Hebbian learning of cognitive control: Dealing with specific and nonspecific adaptation. *Psychological Review*, *115*(2), 518–525. https://doi.org/10.1037/0033-295X.115.2.518

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., . . . Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*(1), 58–76. https://doi.org/10.3758/s13423-017-1323-7

Whitehead, P. S., Pfeuffer, C. U., & Egner, T. (2020). Memories of control: One-shot episodic learning of item-specific stimulus-control associations. *Cognition*, *199*, Article 104220. https://doi.org/10.1016/j.cognition.2020.104220

Wood, J. N. (2011). When do spatial and visual working memory interact? *Attention, Perception & Psychophysics*, *73*(2), 420–439. https://doi.org/10.3758/s13414-010-0048-8

Woodman, G. F., Carlisle, N. B., & Reinhart, R. M. G. (2013). Where do we store the memory representations that guide attention? *Journal of Vision*, *13*(3), 1. https://doi.org/10.1167/13.3.1

Received September 7, 2020
Revision received December 26, 2020
Accepted January 30, 2021 ■