Attention, Perception, & Psychophysics (2026) 88:135
https://doi.org/10.3758/s13414-026-03279-8

Memory-guided selective attention: Evidence for context-specific control using the trial-unique Stroop task

Stephanie L. Souliere[1] · **Nicholaus P. Brosowsky**[1]

Received: 15 October 2024 / Accepted: 12 February 2026
© The Psychonomic Society, Inc. 2026

Abstract

Stroop effects can be influenced by context-specific cues associated with varying levels of proportion congruency, even for items that occur with equal frequency in each context. This theoretically important result points to the involvement of a cue-driven memory process in the dynamic updating of selective attention. However, attempts to replicate this result have been challenging, potentially due to limitations or confounds in the experimental design, or difficulties in inducing context-specific learning due to item-specific cues. The present study aimed to reevaluate the memory-guided hypothesis using the trial-unique Stroop paradigm, where each trial consists of a unique stimulus and response that never reappear throughout the experiment, thereby preventing trial-to-trial associative regularities. One context was associated with high conflict (20% proportion congruency), and the other with low conflict (80% proportion congruency). Experiment 1 demonstrated that trials within the low-conflict context exhibited larger congruency effects compared with those in the high-conflict context, aligning with the memory-guided selective attention hypothesis that attentional priorities become associated with contextual cues and are automatically reinstated through memory retrieval. Experiment 2 replicated this finding and examined whether both auditory and visual context cues were necessary by assigning participants to combined auditory–visual, auditory-only, or visual-only conditions. The context-specific proportion congruency effect was successfully replicated in both the combined and visual-only conditions, but not in the auditory-only condition, suggesting that visual contextual cues may be more effective than auditory cues. These findings provide compelling evidence for memory-guided selective attention while revealing important boundary conditions regarding contextual cue effectiveness.

Keywords Cognitive control · Conflict adaptation · Attention · Context-specific proportion congruency · Context-dependent

In day-to-day life, our prior experiences can often serve as effective guides for how we should direct our attention (Hutchinson & Turk-Browne, 2012). This might occur in a very explicit way: You remember how your co-worker sitting next to you constantly taps their pen and you prepare in advance to ignore the distraction and focus your attention on the task at hand. Cognitive control typically refers to the set of cognitive processes that would enable this topdown regulation of selective attention, which is thought to be a slow, effortful, and resourcedemanding process. Another possibility, however, is that memory might guide selective attention in a more automatic, implicit, manner. Consider, for instance, going to the same office every day, at the same desk, with the same co-worker beside you who persistently taps their pen in a distracting way. When faced with repeated experiences in a predictable environment, a cuedriven memory-retrieval process might sidestep the demanding task of

Public significance How does memory guide selective attention? Previous demonstrations of context-specific control are often confounded by item-specific associative learning. We address this using a novel trial-unique Stroop paradigm that eliminates item repetition. Our results provide clear evidence that context alone automatically triggers the retrieval of attentional settings, ruling out item-specific associative-learning explanations. We replicated this finding and identified critical boundary conditions regarding cue modality. This study validates the memory-guided selective attention hypothesis, offering a clean methodological foundation for future research and addressing longstanding replication difficulties in the field.

✉ Nicholaus P. Brosowsky
nicholaus.brosowsky@umanitoba.ca

[1] Department of Psychology, University of Manitoba, 190 Dysart Road, Winnipeg, MB R3T 2N2, Canada

Published online: 27 May 2026

updating selective attention. Instead, the environmental context could serve as a cue, automatically reinstating a previously used control process. The concept, initially seeming counterintuitive, is termed 'automatic control' (Jacoby et al., 2003; see also, Bugg & Crump, 2012) or 'memory-guided selective attention' (Brosowsky & Crump, 2018; see also, Crump, 2016). The evidence for memory-guided selective attention, however, is mixed, and its existence remains a subject of debate and uncertainty (for recent reviews, see Bugg et al., 2020; Weidler et al., 2022).

Within this context, selective attention is typically measured using interference tasks, such as the Stroop (1935) task, where participants identify the font color of color words (e.g., naming the blue font color of the word 'RED'). Stimuli can be congruent, with word meaning matching the font color ('RED' in red font), or incongruent, with a mismatch between color and word meaning ('RED' in green font).

Conflict in the Stroop task arises when the irrelevant stimulus dimension (word meaning) activates a response that competes with the response required by the relevant dimension (ink color). This creates a congruency effect where responses are typically slower and less accurate on incongruent trials compared with congruent trials. The size of this congruency effect indicates both the relative strength of the irrelevant dimension's influence and the level of cognitive control being applied. When participants effectively engage control mechanisms to prioritize processing of the relevant dimension (ink color) over the irrelevant dimension (word meaning), congruency effects are reduced. This dynamic adjustment of attentional resources in response to conflict represents a core aspect of adaptive cognitive control.

Importantly, this adaptive control system responds flexibly to different temporal patterns of conflict in the environment (for reviews, see Braem et al., 2019; Bugg & Crump, 2012). When, for example, conflict frequency increases across an entire experiment or block (the listwide proportion congruency effect), congruency effects diminish (Brosowsky, 2024; Logan & Zbrodoff, 1979; Lowe & Mitterer, 1982). This is taken as evidence for proactive attentional adjustments in response to the global likelihood of conflict. Similarly, the congruency sequence effect, where the congruency effect is smaller following an incongruent trial compared with a congruent trial, demonstrates reactive control adjustments triggered by recently experienced conflict (Brosowsky, 2024; Gratton et al., 1992).

Building on these findings, researchers have investigated whether cognitive control can also adapt to spatial contexts associated with different conflict frequencies. Critical evidence for memory-guided selective attention comes from the use of the 'context-specific proportion congruency' (CPSC) paradigm (see Braem et al., 2019; Bugg & Crump, 2012). In the traditional CSPC task, Stroop stimuli appeared in distinct locations associated with differing conflict frequencies (75% vs. 25% congruent). Although participants could not predict the upcoming trial's context, smaller congruency effects in the high versus low-conflict contexts suggested rapid, automatic attentional modulation cued by the location the stimulus appeared (e.g., Crump et al., 2006).

However, these CSPC effects can be explained by multiple theoretical mechanisms that differ in both their scope and underlying psychological processes (for a review, see Braem et al., 2019). The context-PC learning hypothesis suggests that attentional settings are reinstated by contextual cues, based on a cue-driven memory retrieval process (e.g., Brosowsky & Crump, 2018; Crump, 2016). After repeated exposure, an associative link between the context and the attentional setting most often used in that context is formed. Upon subsequent exposure to the context, the associated attentional setting is automatically reinstated. This attentional account predicts that contextual effects should generalize to any stimuli appearing within the biased context, regardless of whether those specific items were part of the original proportion manipulation.

The conjunctive learning hypothesis (also referred to as context + item-pc learning; Bugg et al., 2020) also proposes an attentional mechanism, but one driven by item-specific associations. Rather than pure contextual associations, participants learn that specific items (e.g., the word 'red') appear as mostly congruent in one location and mostly incongruent in another, leading to item-specific attentional adjustments within each context. While still an attentional account, this mechanism predicts that proportion effects should be limited to the specific items that were manipulated—new items appearing in the same contexts should show no contextual effects unless they share item-level features with the manipulated stimuli.

In contrast, the compound-cue contingency hypothesis proposes a fundamentally different, nonattentional mechanism based on stimulus–response contingency learning that is bound by context (Schmidt & Lemercier, 2019). This account recognizes that when proportion congruency is manipulated, the frequency of individual congruent and incongruent items is also necessarily manipulated, creating biases in stimulus–response contingencies. For example, in a high-conflict context where most trials are incongruent, the word 'red' might frequently appear in green ink, creating a strong contingency between seeing 'red' and responding 'green' in that location. Conversely, in a low-conflict context, 'red' might frequently appear in red ink, creating a strong contingency between seeing 'red' and responding 'red.' The 'compound-cue' aspect refers to the context-dependent nature of these contingencies. The same stimulus can have different learned response associations depending on the spatial context. As a nonattentional account, this mechanism predicts that contextual effects should be restricted to the specific stimulus–response combinations that were

learned, with no generalization to new items within the same contexts.

The key distinction between these three accounts concerns whether contextual effects generalize beyond the specific items used in the proportion manipulation. The context-PC learning hypothesis predicts full generalization across all items within a context, while both the conjunctive learning and compound-cue contingency hypotheses predict no generalization to new items that do not share features with the manipulated stimuli.

One important approach for testing generalization is the 'transfer' design developed by Crump and Milliken (2009). This paradigm addresses a fundamental limitation of traditional CSPC studies: When all items are frequency-biased, it becomes impossible to determine whether contextual effects reflect broad attentional adjustments or item-specific learning (Braem et al., 2019). In the transfer design, Stroop items were divided into an 'inducer' set, biased at 75% congruency in one location and 25% in another, and a 'diagnostic' set, maintaining 50% congruency across both locations. The critical diagnostic set assessed whether the effects of contextual cues would generalize beyond the specific items used in the proportion manipulation. If context-PC learning drives CSPC effects, diagnostic items should show smaller congruency effects in high versus low-conflict contexts. If conjunctive learning or compound-cue contingency drives the effects, diagnostic items should show no contextual differences.

Crump and Milliken's (2009) results supported the context-PC learning account, showing smaller congruency effects for diagnostic items in high versus low-conflict contexts. This suggested that contextual cues had triggered the retrieval of attentional priorities that applied to all items within the context, not just those with specific learned associations (Braem et al., 2019; Bugg et al., 2020).

However, replicating this critical transfer effect has proven surprisingly difficult. Despite the theoretical importance of these findings for distinguishing memory-guided attention from alternative learning processes (Braem et al., 2019; Bugg et al., 2020), the small literature shows a mix of both positive and negative outcomes (see Table 1). Given these divergent results, multiple explanations have emerged to address the inconsistencies.

First, the inconsistencies may arise because observed CSPC effects represent Type 1 errors resulting from methodological limitations, task variability, and publication bias (e.g., see Hutcheon & Spieler, 2017). The CSPC effect can only be observed through a complex interaction and typically manifests as a small effect size (e.g., 16-ms difference-of-difference score; Crump et al., 2006). Initial demonstrations used small samples that were likely underpowered to reliably detect such effects (e.g., 16 participants, Crump et al., 2006; 17 participants, Crump & Milliken, 2009). This becomes evident when considering the sample sizes needed for adequate detection. Crump et al. (2017) estimated that 100 participants may be necessary to detect 20-ms CSPC effects

Table 1 Summary of context-specific proportion congruency studies using inducer and diagnostic transfer items

Study	Exp	Task	Inducer PC (%)	Observed effect	
				Inducer	Transfer
Crump and Milliken (2009)	1	Prime-Probe Stroop	100/0	–	Y[a]
	2	Prime-Probe Stroop	92/8	Y[a]	Y
Hutcheon and Spieler (2017)	1	Prime-Probe Stroop	92/8	N	N
	2	Prime-Probe Stroop	83/17	N	N
Crump et al. (2017)	1	Prime-Probe Stroop	100/0	–	Y
	2	Prime-Probe Stroop	92/8	N	N
	3	Letter Flanker	100/0	–	Y
	4	Letter Flanker	92/8	N	Y
Bugg et al. (2020)	1	Picture-Word Stroop	75/25	N	N
Gonthier and Blaye (2021)	2	Modified Arrow Flanker	100/0	–	Y
Weidler et al. (2022)	3A	Arrow Flanker	100/0	–	Y
	3B	Arrow Flanker	92/8	Y	Y
Hutcheon (2022)	1	Stroop w/Face Cues	92/8	Y	N
	2	Stroop w/Face Cues	92/8	Y	N
	3	Stroop w/Face Cues	92/8	Y	N
	4	Stroop w/Face Cues	92/8	Y	N

PC proportion congruency within each context; *observed effect* indicates whether a statistically significant ($p < .05$) context by proportion congruency interaction for reaction times was observed within the frequency biased inducer item set or the 50% congruent transfer item set

[a]effect was only significant in the second half of trials

reliably, but our calculations suggest the requirements may be even higher. Using the *Superpower* R package (Lakens & Caldwell, 2021) and assuming a standard deviation of 125 and correlation of 0.8, we estimated 250 participants would be necessary to detect a 20-ms CSPC effect. The problem extends beyond sample size to replication practices. Direct replications have been rare (e.g., Crump et al., 2017), with researchers more often conducting 'conceptual' replications that use different tasks (e.g., Stroop, prime–probe Stroop, arrow flanker, letter flanker) and/or modify the task in various ways (e.g., see Table 1).

This reliance on conceptual rather than direct replications is particularly problematic because it can systematically distort the literature (Pashler & Harris, 2012). When conceptual replications fail, researchers can attribute the failure to methodological differences rather than questioning the original effect, and these failures may not get published. Conversely, successful conceptual replications, which can occur by chance alone in small, underpowered studies, are viewed as novel extensions worthy of publication. This dynamic is even evident in the small CSPC literature, where Crump et al., (2008, p. 26) mention in a footnote that they adopted the prime–probe design only after failing to observe CSPC effects using the traditional Stroop design in an unpublished pilot study. Thus, the combination of low-powered designs, methodological variability across studies, and publication bias may have created a literature that overrepresents positive findings for an effect that is not genuinely replicable (see Rosenthal, 1979).

Second, assuming the effect is real, the inconsistencies might stem from inherent measurement challenges (Crump et al., 2017). CSPC effects are typically observed in reaction times rather than error rates and manifest as small interaction effects—difference-of-difference scores—ranging from 10 to 40 ms (Crump et al., 2017; Weidler et al., 2022). This makes the effect inherently difficult to measure reliably (see Hedge et al., 2018, for a review of the reliability of difference scores). The transfer design presents additional challenges as it requires dividing trials into inducer and diagnostic sets, effectively removing at least half the trials from analysis and further reducing statistical power by increasing variability.

Consequently, even if the CSPC effect exists, these methodological limitations may render it exceptionally difficult to measure consistently across studies.

From a theoretical perspective, numerous studies have investigated the factors necessary for context-specific learning. These factors change unsystematically from study to study, which may help explain why context effects are not consistently observed.

First, the nature of the context cue itself appears to be an important determinant of contextual control. Context cues that are naturally processed with the stimulus, such as location (e.g., Crump et al., 2017), irrelevant stimulus color (Lehle & Hübner, 2008), and font type (Bugg et al., 2008), have been shown to produce CSPC effects, while other cues such as irrelevant shapes (Crump et al., 2008) and faces (Hutcheon, 2022) have not (see also Gottschalk & Fischer, 2016). Additionally, the task relevance of the context cue may be critical, particularly when the cue is not embedded in the stimulus (Brosowsky & Crump, 2021; Bugg et al., 2020; Crump et al., 2008). Brosowsky and Crump (2021) presented context cues associated with 50% congruency that shared one feature with a 100% congruent context cue and one feature with a 0% congruent context cue. They manipulated the task-relevance of these features across blocks by instructing participants to count one of the two shared features. When counting the 100% congruent cue feature, they found a larger congruency effect for the 50% trials compared with when participants counted the 0% congruent cue feature. By changing which contextual feature was task relevant, they were able to manipulate the direction of the CSPC transfer effect across blocks, suggesting that task relevance is necessary to observe CSPC effects.

Second, prior experience with the context–conflict associations may be important for contextual control to develop (Brosowsky & Crump, 2016; Lehle & Hübner, 2008). Many studies tend to use blocked practice where conflicting stimuli are consistently presented in their respective contexts with the implicit understanding that participants may require training to learn the associations (e.g., Bugg et al., 2020; Hutcheon, 2022; see also Brosowsky & Crump, 2016). Crump and Milliken (2009) only observed the CSPC effect in the second half of the experiment, suggesting that learning the context–attention associations requires experience. However, this learning effect was not replicated in Crump et al. (2016). Lehle and Hübner (2008) also provide some evidence for this claim, failing to observe the CSPC effect without practice but successfully observing it after including a blocked practice phase.

Finally, associative competition among stimulus features may be another important determinant of contextual control (Bugg et al., 2020). In any CSPC design, multiple features vary systematically with the proportions (e.g., location, word, color), meaning that associations could theoretically form between cognitive control and any one or combination of these features. There may be competition between which feature becomes associated with control adjustments. Inconsistencies in CSPC effects could arise because sometimes participants form item-level associations rather than context-level associations, which would eliminate the CSPC effect.

Bugg et al. (2020) tested this directly using a modified transfer design with a pictureword Stroop task where participants named pictures of animals while ignoring overlaid animal words. They manipulated both location (context feature) and animal category (item feature) as

predictors of proportion congruency. For example, cats and birds always appeared as 75% congruent in the upper location but 50% congruent in the lower location, while fish and dogs showed the reverse pattern. This created competing predictors: if diagnostic items associated with the context in which they were presented, they should show a pattern consistent with that context, but if they associated with their item category, they should show the opposite pattern. Their results showed a typical CSPC effect for inducer items but the opposite pattern for diagnostic items, suggesting that participants learned item-based rather than location-based associations. This demonstrates that when multiple features are available as potential cues, control may selectively associate with item-specific rather than context-specific features. Bugg et al. suggest that failures to reproduce CSPC effects may be better explained by this associative competition rather than by underpowered studies or weak manipulations. This implies that to consistently observe context-based learning, competing item-specific associations may need to be eliminated.

In summary, the CSPC literature presents a complex picture that may largely stem from inherent limitations of traditional designs that require item repetition. When individual stimuli must repeat across trials to create proportion congruency manipulations, several problematic confounds emerge: stimulus–response contingencies become biased, creating nonattentional learning opportunities that can masquerade as contextual control (e.g., Schmidt & Lemercier, 2019); competing item-specific associations can override context-specific learning (e.g., Bugg et al., 2020), leading to inconsistent effects; and the resulting confounds necessitate complex 'transfer' designs that reduce statistical power and introduce additional sources of variability. Furthermore, the small effect sizes typical of these complex interactions, combined with methodological variability across studies, may have created a literature where genuine replicable effects are difficult to distinguish from statistical artifacts. These design-related challenges suggest that many of the theoretical debates and replication difficulties in the CSPC literature may be methodological rather than theoretical in nature.

Given the theoretical importance of the CSPC effect, and the inconsistency of the literature, the focus of the present study was to further test the memory-guided selective attention hypothesis using a trial-unique Stroop task (Brosowsky, 2024). In this task, a word is presented visually, and another is spoken aloud. Participants type the word they hear while ignoring the word they see (see Fig. 1A). A trial is congruent when the visual and auditory words are identical, and incongruent when they differ. A critical feature of this task is that each word is only ever presented on a single trial; this lack of stimulus repetition effectively prevents low-level associative regularities and eliminates competing associative learning (Brosowsky, 2024; see also Braem et al., 2019).

We adapted the trial-unique Stroop task to incorporate contextual cues, creating an optimal test of the memory-guided selective attention hypothesis (see Fig. 1B). Our design addresses the key limitations identified in prior research. First, the trial-unique design ensures no competition from item-specific associative learning since items never repeat. Second, we ensured participants had a learning history by providing practice with contexts at 100%/0% congruency. Third, we created high-salience contexts by embedding redundant context features into both target and distractor dimensions. Additionally, this task offers several methodological advantages: (1) congruency effects in this task tend to be more pronounced than in other variants (Brosowsky, 2024); (2) the absence of repeating items eliminates the need for 'inducer' or 'diagnostic' sets, allowing for the inclusion of all trials in the analyses; and (3) higher reliability than previously used variants ensures enhanced measurement precision (Brosowsky, 2024; see also Hedge et al., 2018; Rouder & Haaf, 2019).

Critically, this design allows for a clear theoretical interpretation of the results. If CSPC effects emerge in our trial-unique task, this would provide strong evidence for the memoryguided selective attention hypothesis, demonstrating that contextual cues can indeed trigger retrieval of appropriate attentional control settings. Conversely, if CSPC effects fail to emerge despite optimal conditions for context-specific learning, this would suggest that previously observed CSPC effects may have been driven primarily by item repetition and item-specific associative learning rather than genuine contextual control.

We tested our hypothesis through two experiments. Experiment 1 embedded redundant contextual cues in both auditory and visual dimensions. Given concerns about replicability, Experiment 2 served as both a direct replication and extension, examining whether redundant bimodal context cues are necessary. Participants were assigned to one of three groups: direct replication of Experiment 1, auditory context cues only, or visual cues only. This design allowed us to determine both whether our original findings were replicable and which cues were essential for contextual effects.

Experiment 1

Method

Sample-size rationale

We preregistered our analysis and sampling plan (https://osf.io/xer52). In a prior study using the same task, we found a Cohen's d of ~ .65 for congruency sequence and list-wide proportion congruency effects (Brosowsky, 2024). We chose 80 participants based on practicalities like subject

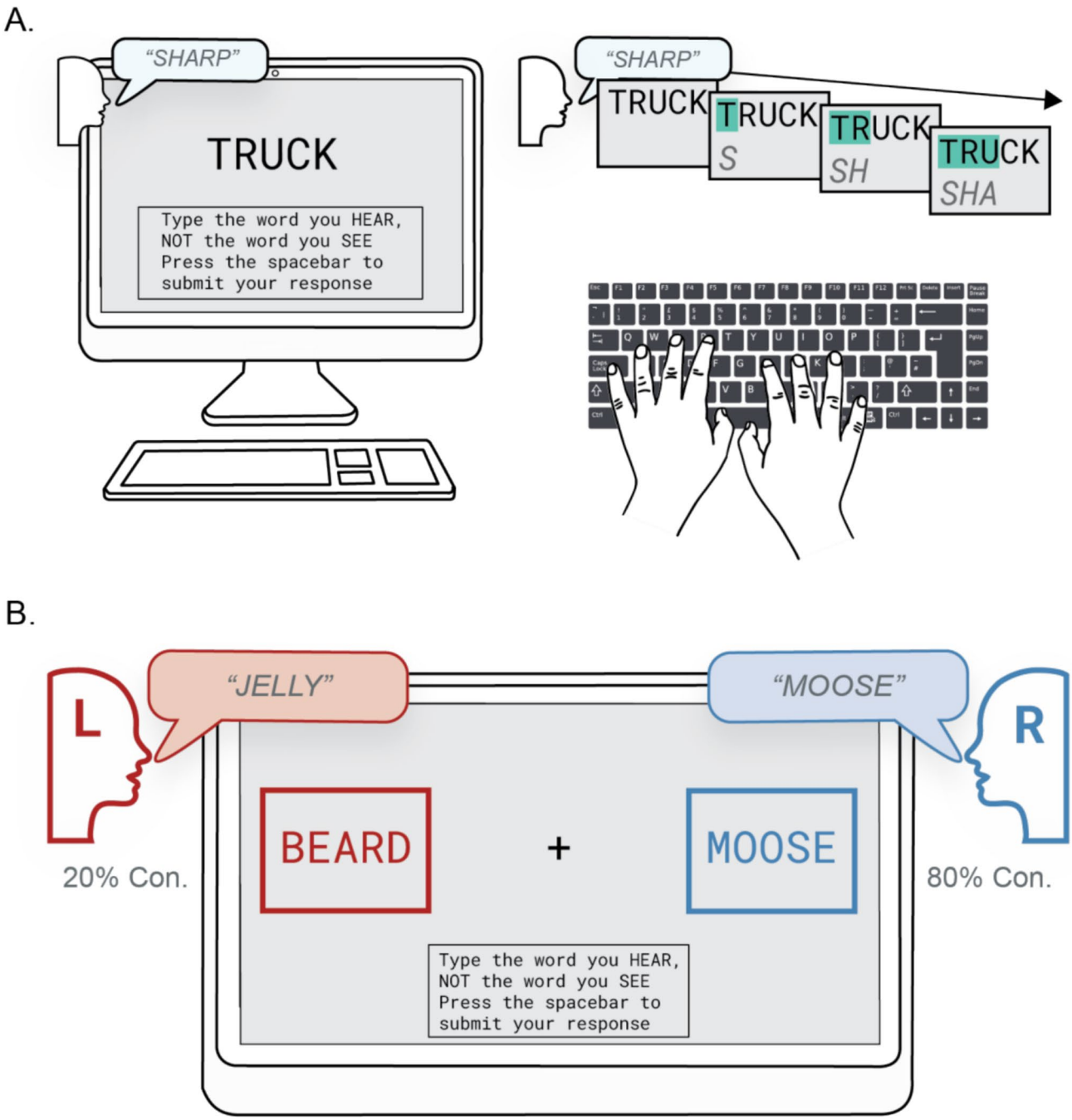

Fig. 1 Illustration of the trial-unique Stroop task. **A.** An illustration of the trial-unique Stroop task (Brosowsky, 2024). Participants heard a word while seeing another on-screen. They typed the spoken word while ignoring the visual word. As they typed, each letter's background shifted from grey to green, with no feedback on typed letters. The displayed 'SHA' is for illustration; it was not shown to participants. **B.** An illustration of the context-specific task. On each trial, the visual word was displayed on either the left or right side of the screen in red or blue font, while the spoken word played through either the left or right headphone in one of two distinct voices. One set of context features (e.g., left/red/voice 1) was associated with high conflict (20% congruent) and the other (e.g., right/blue/voice 2) low conflict (80% congruent). (Color figure online)

pool availability and data collection time, as well as a power analysis.

With 75 participants, we could detect a Cohen's *d* of .35 with 80% power, about 54% of the observed effects. We aimed for 80 participants, anticipating some removals for low accuracy. This aligns with past studies where we estimated power for interaction effects using Monte Carlo

simulations (e.g., Brosowsky & Egner, 2021; Brosowsky et al., 2021; Crump et al., 2017).

Participants

Participants were 80 University of Manitoba undergraduate students who completed the experiment online for course credit. Participant ages ranged from 18 to 45 ($M = 20$, $SD = 4$). When asked to indicate their gender with a free response, 45 identified as a woman or female; 32 as a man or male; one as nonbinary; and two chose not to respond.

Stimuli

The experiment was programmed in JavaScript using the Tone.js JavaScript library[1]

(Mann, 2015; see also Brosowsky, 2024; Brosowsky et al., 2020; Reimers & Stewart, 2016).

Words were restricted to four-letter, five-letter, or six-letter words selected from The Auditory English Lexicon Project database (Goh et al., 2020) in the same manner as Brosowsky (2024). For each participant, 240 words (80 each of four, five, and six letters) were chosen for congruent trials and 480 words (160 each of four, five, and six letters) for incongruent trials without repetition. Incongruent pairs had words of equal length, distinct initial letters, and a Levenshtein edit distance below three (Levenshtein, 1966). Pairs not fitting these criteria were redrawn. Each word was used only once in the experiment. Context features were (1) visual spatial location (left/ right); (2) color (red/ blue); (3) auditory location (left/ right); and (4) speaker voice ('American Male, M2'/ 'American Female, F2'; see Goh et al., 2020). Features were randomly assigned to either 80% or 20% congruency with the restriction that the left/right visual and auditory locations were always consistent with one another (i.e., left auditory with left visual).

Task and procedure

Participants first completed an audio test to confirm that they could hear and respond to the words. They completed a headphone check (Woods et al., 2017), then an audio test to adjust the volume to comfortable level. Next, participants took a typing test, entering four-letter words spoken aloud. On initiation, four *X*s appeared on-screen, changing to the typed letter as feedback. After typing four letters, they pressed the spacebar to submit. To advance, participants needed 7/10 correct trials. Failure to meet this threshold required retaking the test.

After successfully completing the audio test phase, participants moved on to the experimental phase. Every trial began with a fixation cross for 500 ms, followed by a blank screen that randomly varied in duration between 250 and 750 ms, followed by the target stimulus.

On every trial a word was presented on the screen, and another was heard spoken aloud. Participants were instructed to type the words they heard and ignore the words that were presented on the screen (a task reminder was presented below the stimulus throughout the experiment; see Fig. 1). The background of each letter changed from grey to green as participants typed their response. However, no visual feedback was provided as to which letters were typed. When the participant had typed the minimum number of letters for that trial (four, five, or six), they could submit their response by pressing the spacebar, which initiated the next trial. Participants completed two phases, each beginning with a practice block of 60 trials of 100%/0% congruency, followed by 180 trials of 80%/20% congruency. The practice blocks alternated between subblocks of 10 trials of 100% congruent and 0% congruent trials. The experimental blocks presented trials in a randomized order.

Results

Data analysis

Following our preregistered data cleaning procedures (https://osf.io/xer52), participants with accuracy below 60% were excluded, eliminating six participants. For reaction time analyses, trials outside 300 ms to 3 s were discarded (1.26% of observations). Error trials were also excluded before using the van Selst and Jolicoeur nonrecursive outlier removal (Van Selst & Jolicoeur, 1994), which further removed 2.95% of observations. All data and analysis code are available online (https://osf.io/c5m4n/).

The primary metric was the balanced integration score (BIS; Liesefeld & Janczyk, 2019; see also Brosowsky, 2024), a measure that combines speed and accuracy. The balanced integration score is calculated by standardizing the mean reaction time ($\bar{z}_{\overline{RT}}$) and accuracy (z_{ACC}) and then subtracting one standardized score from the other. As in Liesefeld and Janczyk (2019), We subtracted the proportion accuracy from the mean reaction time such that it was comparable in direction to Stroop reaction time effects (i.e., higher BIS indicates poorer performance).

$$BIS_{ij} = z_{\overline{RT}_{ij}} - z_{ACC_{ij}}$$

[1] https://tonejs.github.io/; Tone.js delivers audio stimuli by 'looking ahead' and scheduling them based on the precise audio clock. Using this same method, visual stimuli are scheduled in sync with the audio clock and displayed through the requestAnimationFrame method, ensuring tight audio-visual synchronization.

$$withz_{xij} = \frac{x_{ij} - \bar{x}}{SD_x}$$

All reaction times refer to response initiation, while accuracy is determined by the initial letter typed (Brosowsky, 2024; Crump et al., 2017; Logan & Zbrodoff, 1998). See Table 2 for summary statistics.

Our choice of BIS as the primary measure was preregistered and aligned with Brosowsky's (2024) original trial-unique paradigm design. Beyond replication purposes, the BIS offers methodological advantages for our analysis. It effectively addresses potential confounds from individual differences in speed–accuracy trade-offs, which could otherwise manifest as reaction-time differences even without significant group-level error rate variations. This concern is particularly relevant in Stroop designs, where asymmetrical patterns typically emerge with larger effects on reaction times than error rates. By integrating both reaction times and error rates at the individual level, the BIS provides a more comprehensive performance metric. Nevertheless, we also present separate analyses of reaction times and error rates to provide a complete picture of performance patterns.

Context-specific proportion congruency effects

Balanced integration scores Analyzing the balanced integration scores using a 2 × 2 repeated-measures analysis of variance (ANOVA) with PC context (80% PC, 20% PC) and congruency (congruent, incongruent) as the factors resulted in a significant main effect of congruency, $F(1,72) = 58.18$, $MSE = 2.71$, $p < .001$, $\hat{\eta}_p^2 = .447$, 90% CI [0.3, 0.55], no significant main effect of PC context, $F(1,72) = 1.04$, $MSE = 0.19$, $p = .310$, $\hat{\eta}_p^2 = .014$, 90% CI [0, 0.09], and a significant interaction between PC context and congruency, $F(1,72) = 7.74$, $MSE = 0.12$, $p = .007$, $\hat{\eta}_p^2 = .097$, 90% CI [0.02, 0.21] (see Fig. 2).

Reaction times Analyzing reaction times, we used a generalized linear mixed-effects model with an inverse gaussian distribution and identity link function (see Lo & Andrews,

Table 2 Experiment 1 results

		Congruency			
	PC	Inc	Con	Congruency effect	CSPC
BIS	MI	0.68 (0.18)	−0.67 (0.1)	1.36 (0.18)	0.23* (0.08)
	MC	0.85 (0.23)	−0.73 (0.1)	1.59 (0.22)	
RT (ms)	MI	1,435.6 (10.38)	1,217.08 (10.76)	218.52 (4.8)	46.2* (7.41)
	MC	1,441.88 (11.02)	1,177.16 (10.49)	264.72 (5.66)	
ER (%)	MI	2.77 (0.31)	0.9 (0.18)	1.87 (0.28)	0.3 (0.38)
	MC	3.34 (0.45)	1.18 (0.15)	2.16 (0.39)	

*$p < .05$; *BIS* balanced integration score; *RT* reaction time (ms); *ER* error percentage; *CSPC* context-specific proportion congruency effect; RTs and ERs were estimated using generalized linear models; Standard errors are presented in brackets

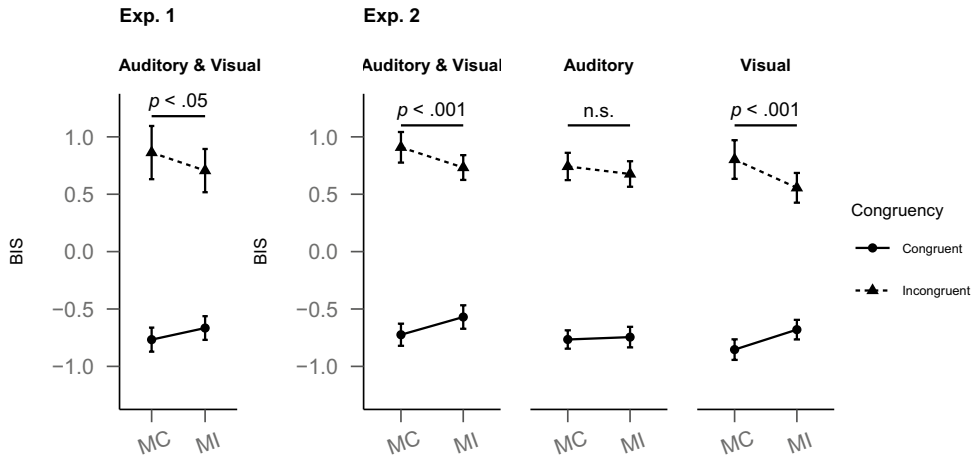

Fig. 2 Results from Experiments 1 and 2. This figure illustrates the balanced integration scores (BIS) across congruent and incongruent trials for the mostly congruent (MC) and mostly incongruent (MI) contexts for Experiments 1 and 2. Higher scores indicate better performance. Error bars represent standard error of the mean

2015), PC context and congruency as fixed effects (contrast coding: incongruent = −1 and congruent = +1; 80% PC = −1 and 20% PC = +1) and subjects as a random effect. Here, the main effect of congruency was significant, β = −123.2, 95% CI [−127.05, −119.35], $t = -62.76$, $p < .001$, the main effect of PC context was significant, β = −9.22, 95% CI [−12.99, −5.45], $t = -4.79$, $p < .001$, and the interaction between PC context and congruency was significant, β = −11.65, 95% CI [−15.37, −7.93], $t = -6.14$, $p < .001$.

Error rates Finally, analyzing error rates using a generalized linear mixed-effects model with a binomial distribution resulted in a significant main effect of congruency, β = −0.55, 95% CI [−0.66, −0.44], $t = -10.15$, $p < .001$, no significant main effect of PC context, β = 0.09, 95% CI [−0.01, 0.2], $t = 1.69$, $p = 0.091$, and no significant interaction between PC context and congruency, β = 0, 95% CI [−0.1, 0.11], $t = 0.02$, $p = 0.986$.

Experiment 2

Experiment 1 provided clear support for the memory-guided selective attention hypothesis. By using a trial-unique design, we observed a significant CSPC effect, characterized by larger congruency effects in the low-conflict context compared with the high-conflict context, without the influence of item repetition or item-specific contingency learning. These results suggest that contextual cues alone can trigger the automatic retrieval of attentional control settings.

However, Experiment 1 employed redundant contextual cues embedded in both the relevant (auditory location/voice) and irrelevant (visual location/color) dimensions. While this redundancy was designed to maximize cue salience, it leaves open the question of whether such redundancy is necessary to drive the effect, or if cues in a single modality are sufficient. Furthermore, given the replication challenges noted in the CSPC literature (e.g., Crump et al., 2017), it is critical to establish the reliability of the trial-unique paradigm.

Experiment 2 was designed to address these issues by serving as both a direct replication and a conceptual extension. We examined the necessity of redundant cues by assigning participants to one of three conditions: (1) a combined condition (auditory and visual) which served as a direct replication of Experiment 1; (2) an auditory-only condition; or (3) a visual-only condition. This design allowed us to test whether the CSPC effect relies on the integration of multisensory cues or if features within the target (auditory) or distractor (visual) dimensions are independently sufficient to cue memory-guided attention. Based on the robust effects in Experiment 1, we hypothesized that the combined condition would replicate the CSPC effect, while the single-modality conditions would reveal the relative efficacy of target-based versus distractor-based cues.

Method

Sample-size rationale

We preregistered our analysis and sampling plan (https://osf.io/a4we3) and estimated our power to detect the presence versus absence of the CSPC effect using the *SuperPower* R package (Lakens & Caldwell, 2021). With 100 participants per group, and assuming a standard deviation of 150 ms and within-subjects correlation of .9, we determined that we could reliably detect the presence versus absence of the CSPC effect between two groups with 93.5% power (Cohen's $f = .063$). We aimed to collect at least 115 per group to allow for some data removal due to poor performance (accuracy less than 60%, as in Experiment 1).

Participants

Participants were 348 University of Manitoba undergraduate students who completed the experiment online for course credit. The participant ages ranged from 18 to 66 ($M = 20$, $SD = 4.59$). When asked to indicate their gender with a free response, 210 identified as a woman or female; 127 as a man or male; eight as nonbinary; and four chose not to respond.

Stimuli

The stimuli were identical to Experiment 1. Context features were: (1) visual spatial location (left/ right); (2) color (red/ blue); (3) auditory location (left/ right); and (4) speaker voice ('American Male, M2'/ 'American Female, F2'; see Goh et al., 2020). Features were randomly assigned to either 80% or 20% congruency with the restriction that the left/ right visual and auditory locations were always consistent with one another (i.e., left auditory with left visual).

Participants were randomly assigned to one of three conditions that differed only in the combinations of contextual features presented. In the auditory and visual group, we presented stimuli with the same bimodal contextual cues as Experiment 1 (all four features). In the auditory-only group, words were always presented at the center of the screen in black font but could be presented in the left/right auditory location and M2/F2 voices (features 3 and 4). In the visual-only group, words were always presented in both left and right speakers using a single voice (randomly assigned as either M2 or F2) but could appear in the left/right visual spatial location and red/blue colored font (features 1 and 2).

Table 3 Experiment 2 results

Congruency					
Condition	PC	Inc	Con	CE	CSPC
BIS					
A & V	MI	0.73 (0.11)	−0.57 (0.10)	1.30* (0.07)	0.33* (0.08)
	MC	0.91 (0.13)	−0.72 (0.10)	1.63* (0.10)	
Auditory	MI	0.68 (0.11)	−0.74 (0.09)	1.42* (0.10)	0.09 (0.08)
	MC	0.74 (0.12)	−0.77 (0.08)	1.51* (0.11)	
Visual	MI	0.56 (0.13)	−0.68 (0.09)	1.24* (0.09)	0.42* (0.10)
	MC	0.80 (0.17)	−0.85 (0.09)	1.66* (0.14)	
RT (ms)					
A & V	MI	1394.25 (18.17)	1129.95 (18.32)	264* (3.49)	39* (5.36)
	MC	1403.41 (18.47)	1099.76 (18.13)	304* (4.07)	
Auditory	MI	1352.65 (18.17)	1081.17 (18.3)	271* (3.41)	8 (5.24)
	MC	1360.66 (18.45)	1081.22 (18.13)	279* (3.98)	
Visual	MI	1338.47 (18.08)	1093.94 (18.22)	245* (3.41)	33* (5.17)
	MC	1339.57 (18.36)	1061.93 (18.05)	278* (3.9)	
ER (%)					
A & V	MI	2.30 (0.21)	1.35 (0.19)	0.95* (0.74)	0.60 (2.17)
	MC	2.81 (0.32)	1.26 (0.13)	1.55* (0.47)	
Auditory	MI	2.61 (0.23)	1.41 (0.19)	1.20* (0.66)	0.23 (1.8)
	MC	2.76 (0.31)	1.34 (0.13)	1.43* (0.49)	
Visual	MI	2.39 (0.21)	1.43 (0.2)	0.96* (0.73)	1.09* (2.47)
	MC	3.34 (0.36)	1.29 (0.13)	2.05* (0.38)	

*$p < .05$; *A & V* auditory and visual condition; *BIS* balanced integration score; *RT* reaction time (ms); *ER* error percentage; *CE* congruency effect; *CSPC* context-specific proportion congruency effect; RTs and ERs are estimated using generalized linear models; standard errors are presented in brackets

Task and procedure

The task and procedure were identical to Experiment 1.

Results

Data analysis

Following our pre-registered analysis plan (https://osf.io/a4we3), participants with less than 60% accuracy were excluded from all analyses. This removed 23 participants. Prior to all analyses, trials with a response time greater than 3 s or less than 300 ms were removed (removing 1.7% of observations). In addition, for response time analyses, all error trials were removed before applying the van Selst and Jolicoer non-recursive outlier removal (Van Selst & Jolicoeur, 1994), removing 2.89% of observations. See Table 3 for a summary of the results.

Context-specific proportion congruency effects

Balanced integration scores Analyzing the balanced integration scores using a 3 × 2 × 2 mixed ANOVA with PC context (80% PC, 20% PC) and congruency (congruent, incongruent) as within-subjects variables and condition (auditory and visual, auditory only, or visual only) as the between-subjects factors. resulted in a significant three-way interaction, $F(2,322) = 4.75$, $MSE = 0.21$, $p = .009$, $\hat{\eta}_p^2 = .029$, 90% CI [0, 0.06]. We then analyzed each condition separately and found that the two-way interaction between congruency and PC Context was significant for the auditory and visual condition, $F(1,107) = 15.05$, $MSE = 0.18$, $p < .001$, $\hat{\eta}_p^2 = .123$, 90% CI [0.04, 0.22], and the visual-only condition, $F(1,108) = 18.98$, $MSE = 0.31$, $p < .001$, $\hat{\eta}_p^2 = .149$, 90% CI [0.06, 0.25], but was not significant for the auditory-only condition, $F(1,107) = 1.24$, $MSE = 0.15$, $p = .269$, $\hat{\eta}_p^2 = .011$, 90% CI [0, 0.07].

Reaction times Analyzing reaction times, we used a 3×2×2 generalized linear mixed-effects model with a Gamma distribution[2] and identity link function (see Lo & Andrews, 2015), PC context (80% PC, 20% PC), congruency (congruent, incongruent) and condition (auditory and visual, auditory only, or visual only) as fixed effects and subjects as a random effect. We used sum-to-zero contrast-coding for congruency and condition and treatment contrasts for condition with auditory and visual as the baseline. Overall, we found that the two-way interaction between congruency and PC Context was significant, $\beta = -9.84$, 95% CI [−12.46, −7.21], $t = -7.35$, $p < .001$, with smaller congruency effects in the mostly incongruent condition ($M = 260$ ms) as compared with the mostly congruent condition ($M = 287$ ms). However, this was qualified by a three-way interaction, indicating that the two-way interaction was significantly smaller in the auditory condition, compared with both the auditory and visual condition, $\beta = 7.85$, 95% CI [4.18, 11.52], $t = 4.19$, $p < .001$ and the visual-only condition, $\beta = 6.29$, 95% CI [2.68, 9.89], $t = 3.42$, $p < .001$. The two-way interactions, however, did not significantly differ between the auditory and visual and the visual-only conditions, $\beta = 1.56$, 95% CI [−2.09, 5.21], $t = 0.84$, $p = 0.401$.

We then analyzed each of the two-way interactions within each condition and found that the two-way interaction between congruency and PC context was significant for the auditory and visual condition, $\beta = -9.84$, 95% CI [−12.49, −7.18], $t = -7.26$, $p < .001$, and the visual-only condition, $\beta = -8.28$, 95% CI [−10.78, −5.77], $t = -6.48$, $p < .001$, but was not significant for the auditory-only condition, $\beta = -1.99$, 95% CI [−4.56, 0.58], $t = -1.52$, $p = 0.128$.

Error rates Finally, we analyzed error rates using a generalized linear mixed-effects model with a binomial distribution following the same procedure as above. Here, we again found a significant two-way interaction between congruency and PC Context, $\beta = -0.34$, 95% CI [−0.42, −0.26], $z = -8.16$, $p < .001$, indicating a smaller congruency effect in the mostly incongruent context ($M = 1.03\%$), as compared with the mostly congruent condition ($M = 1.67\%$). There were no significant three-way interactions. The auditory and visual condition did not significantly differ from the auditory-only, $\beta = 0.00$, 95% CI [−0.12, 0.11], $z = -0.04$, $p = .964$, or visual-only conditions, $\beta = -0.03$, 95% CI [−0.15, 0.08], $z = -0.59$, $p = .558$; and the visual-only condition did not significantly differ from the auditory-only condition, $\beta = 0.08$, 95% CI [−0.03, 0.19], $z = 1.47$, $p = .140$.

General discussion

The current study tested whether memory-guided selective attention drives context-specific proportion congruency effects using a novel, trial-unique Stroop paradigm (Brosowsky, 2024). By ensuring that stimuli were presented only once throughout the task, this design removed item repetition and eliminated competing associative learning explanations for any observed CSPC effects. The task incorporated salient contextual cues in both target and distractor dimensions and provided practice with these contexts. One context was associated with high conflict, characterized by a 20% proportion congruency, and the other with low conflict, featuring an 80% proportion congruency.

Experiment 1 provided clear evidence for the memory-guided selective attention hypothesis. Trials within the low-conflict context exhibited larger congruency effects compared with those within the high-conflict context. These results support the proposal that selective attentional priorities become associated with contextual cues and are automatically reinstated via memory retrieval (Brosowsky & Crump, 2018; Crump, 2016). While it is theoretically possible that these adjustments reflect voluntary, conscious strategy shifts, this interpretation is unlikely given the task structure and prior literature. First, contexts were interleaved in a randomized, unpredictable order, preventing participants from engaging in preparatory control of attention prior to trial onset. Second, whereas controlled attention is typically characterized as slow and effortful, the observed modulation occurred rapidly on a trial-by-trial basis. Finally, previous research has consistently demonstrated that participants remain unaware of context-specific proportion congruency manipulations, suggesting that these control adjustments operate in an implicit manner (e.g., Crump et al., 2006, 2008). Importantly, this effect emerged despite the absence of item repetition, demonstrating that CSPC effects can occur through genuine contextual control rather than item-specific associative learning alone.

Experiment 2 replicated this finding and revealed important boundary conditions. We tested whether both auditory and visual cues were necessary by assigning participants to one of three groups: auditory and visual context (replicating Experiment 1), auditory only context, or visual only context. We successfully replicated the CSPC effect in both the combined auditoryvisual condition and the visual-only

[2] We initially attempted to fit our model using *glmer* with the inverse Gaussian distribution but encountered convergence issues. Therefore, we used *glmmTMB*, which offers improved convergence properties for complex models while maintaining the same fundamental statistical approach. We used the Gamma distribution since *glmmTMB* does not support the inverse Gaussian distribution. The Gamma distribution with identity link is also recommended by Lo and Andrews (2015) for reaction time data.

condition, but not in the auditory-only condition. This pattern was unexpected and suggests that not all contextual cues are equally effective. We initially hypothesized that redundant context features would increase salience and strengthen context-PC associations. However, the visual cue alone was as effective as the combined cues, while the auditory cue alone was insufficient.

Two explanations could account for this asymmetry. First, the modality of the context may be important. Visual location cues (left/right position plus red/blue font) may provide more salient or effective learning signals than auditory location cues (left/right ear plus male/female voice). Second, the dimension in which the context is embedded may matter. Since the auditory dimension was the target and visual was the distractor, context-PC associations may depend on embedding cues in the distractor dimension rather than the target dimension.

The current data cannot distinguish between these possibilities. However, future studies could test both explanations using a fully crossed design. To properly invert the current task, participants would need to vocally respond to visual words while ignoring auditory words. This vocal response modality would maintain the appropriate stimulus–response compatibility, as the current task uses typed responses (compatible with visual processing) while the inverse would use vocal responses (compatible with auditory processing). This design consideration aligns with Kornblum's dimensional overlap model, which emphasizes that conflict emerges from overlapping stimulus and response dimensions (Kornblum et al., 1990). A fully crossed design comparing both task versions would determine whether the target/distractor dimension or the auditory/visual modality drives the observed asymmetry.

Beyond addressing modality effects, the trial-unique design opens several promising research directions. First, researchers could systematically examine the boundary conditions identified in prior CSPC work. While our task-irrelevant contexts produced robust effects, the role of task relevance in modulating effect size remains unclear. It appears that task relevance is not required but could be effective for enhancing or modulating the CSPC effect (e.g., Brosowsky & Crump, 2021). Second, although we included practice blocks, the necessity of prior experience for context-specific learning warrants direct testing through designs that manipulate practice systematically.

Perhaps most importantly, the trial-unique approach provides a clean foundation for investigating different learning mechanisms that have been difficult to disentangle in traditional paradigms. The current design eliminates trial-to-trial associative regularities such as contingency learning, target/distractor correlations, and item-specific associations. This creates an ideal baseline from which researchers could systematically reintroduce specific types of repetition to isolate their individual contributions. For example, selectively repeating certain stimulus features while keeping others unique could clarify how stimulus–response learning, contingency learning, and dimensional correlations interact with contextual control. Traditional paradigms struggle with this because repeating stimuli inevitably introduces multiple confounds simultaneously, making it nearly impossible to determine which learning mechanism drives observed effects. The trial-unique framework offers experimental control for parsing these competing explanations.

In conclusion, the results of our study provide compelling evidence that memory-guided selective attention drives context-specific proportion congruency effects. By using a trial-unique design that eliminates item repetition, we demonstrated that CSPC effects can emerge through genuine contextual control rather than low-level associative learning mechanisms such as contingency learning or item-specific associations. Importantly, we successfully replicated this finding across experiments, demonstrating the reliability of the effect when competing learning mechanisms are controlled. These findings contribute to addressing the longstanding replication difficulties in the CSPC literature, which may have stemmed from methodological limitations and design elements that favor competing learning mechanisms over contextual control. Our results suggest that when individuals encounter a context where a specific attentional strategy was previously effective, memory automatically reactivates that strategy. While this mechanism can be advantageous when contexts reliably predict optimal control settings, it may prove maladaptive when environmental demands have changed. Beyond these theoretical contributions, the trial-unique approach offers a methodologically robust foundation for future research, providing enhanced experimental control for investigating how different learning mechanisms interact with cognitive control.

Author's contributions Conceptualization: Stephanie Souliere and Nicholaus P. Brosowsky.

Formal analysis: Stephanie Souliere and Nicholaus P. Brosowsky.

Funding acquisition: Nicholaus P. Brosowsky.

Methodology: Stephanie Souliere and Nicholaus P. Brosowsky.

Supervision: Nicholaus P. Brosowsky.

Writing—original draft: Stephanie Souliere and Nicholaus P. Brosowsky.

Writing—review & editing: Stephanie Souliere and Nicholaus P. Brosowsky.

Nicholaus P. Brosowsky was supported by Natural Sciences and Engineering Research Council of Canada (NSERC).

Funding This research was supported by a Discovery grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) awarded to N. P. Brosowsky.

Data availability All data and materials are available on the Open Science Framework repository (https://osf.io/c5m4n/).

Code availability All relevant code is available on the Open Science Framework repository (https://osf.io/c5m4n/).

Declarations

Ethics approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This study was approved by the University of Manitoba Research Ethics Board (12/14/2023, HE2021-0208).

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent for publication Not applicable.

Open practices statement The data and analysis code are available online (https://osf.io/c5m4n/) and both Experiment 1 (https://osf.io/xer52) and Experiment 2 (https://osf.io/a4we3) were preregistered.

References

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.

Brosowsky, N. P. (2024). Proactive and reactive cognitive control in the absence of learning and memory confounds: Evidence from a cross-modal trial-unique Stroop task. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*. https://doi.org/10.1037/cep0000325

Brosowsky, N. P., & Crump, M. J. C. (2016). Context-specific attentional sampling: Intentional control as a pre-requisite for contextual control. *Consciousness and Cognition, 44*, 146–160. https://doi.org/10.1016/j.concog.2016.07.001

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*, 1134–1153. https://doi.org/10.1037/xge0000431

Brosowsky, N. P., & Crump, M. J. C. (2021). Contextual recruitment of selective attention can be updated via changes in task relevance. *Canadian Journal of Experimental Psychology/revue Canadienne De Psychologie Expérimentale, 75*(1), 19–34. https://doi.org/10.1037/cep0000221

Brosowsky, N. P., DeGutis, J., Esterman, M., Smilek, D., & Seli, P. (2020). Mind wandering, motivation, and task performance over time: Evidence that motivation insulates people from the negative effects of mind wandering. *Psychology of Consciousness: Theory, Research, and Practice, 10*(4), 475–486. https://doi.org/10.1037/cns0000263

Brosowsky, N. P., & Egner, T. (2021). Appealing to the cognitive miser: Using demand avoidance to modulate cognitive flexibility in cued and voluntary task switching. *Journal of Experimental Psychology: Human Perception and Performance, 47*(10), 1329.

Brosowsky, N. P., Murray, S., Schooler, J. W., & Seli, P. (2021). Attention need not always apply: Mind wandering impedes explicit but not implicit sequence learning. *Cognition, 209*, Article 104530. https://doi.org/10.1016/j.cognition.2020.104530

Bugg, J. M., & Crump, M. (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*. https://www.frontiersin.org/articles/https://doi.org/10.3389/fpsyg.2012.00367

Bugg, J. M., Jacoby, L. L., & Toth, J. P. (2008). Multiple levels of control in the Stroop task. *Memory & Cognition, 36*, 1484–1494. https://doi.org/10.3758/MC.36.8.1484

Bugg, J. M., Suh, J., Colvett, J. S., & Lehmann, S. G. (2020). What can be learned in a context-specific proportion congruence paradigm? Implications for reproducibility. *Journal of Experimental Psychology: Human Perception and Performance, 46*(9), 1029.

Crump, M. J. C. (2016). Learning to selectively attend from context-specific attentional histories: A demonstration and some constraints. *Canadian Journal of Experimental Psychology/revue Canadienne De Psychologie Expérimentale, 70*, 59–77. https://doi.org/10.1037/cep0000066

Crump, M. J. C., Brosowsky, N. P., & Milliken, B. (2017). Reproducing the location-based context-specific proportion congruent effect for frequency unbiased items: A reply to Hutcheon and Spieler (2016). *The Quarterly Journal of Experimental Psychology, 70*, 1792–1807. https://doi.org/10.1080/17470218.2016.1206130

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*, 316–321. https://doi.org/10.3758/BF03193850

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

Crump, M. J. C., Vaquero, J. M. M., & Milliken, B. (2008). Context-specific learning and control: The roles of awareness, task relevance, and relative salience. *Consciousness and Cognition, 17*, 22–36. https://doi.org/10.1016/j.concog.2007.01.004

Goh, W. D., Yap, M. J., & Chee, Q. W. (2020). The Auditory English Lexicon Project: A multitalker, multi-region psycholinguistic database of 10,170 spoken words and nonwords. *Behavior Research Methods, 52*(5), 2202–2231. https://doi.org/10.3758/s13428-020-01352-0

Gonthier, C., & Blaye, A. (2021). Preschoolers are capable of fine-grained implicit cognitive control: Evidence from development of the context-specific proportion congruency effect. *Journal of Experimental Child Psychology, 210*, Article 105211.

Gottschalk, C., & Fischer, R. (2016). Activation of context-specific attentional control sets by exogenous allocation of visual attention to the context? *Psychological Research*. https://doi.org/10.1007/s00426-016-0746-5

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*, 480–506. https://doi.org/10.1037/0096-3445.121.4.480

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods, 50*, 1166–1186.

Hutcheon, T. G. (2022). What is cued by faces in the face-based context-specific proportion congruent manipulation? *Attention, Perception, & Psychophysics, 84*(4), 1248–1263.

Hutcheon, T. G., & Spieler, D. H. (2017). Limits on the generalizability of context-driven control. *Quarterly Journal of Experimental Psychology, 70*(7), 1292–1304. https://doi.org/10.1080/17470218.2016.1182193

Hutchinson, J. B., & Turk-Browne, N. B. (2012). Memory-guided attention: Control fromhttps://doi.org/10.1016/j.tics.2012.10.003

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

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility-A model and taxonomy. *Psychological Review, 97*(2), 253–270. https://doi.org/10.1037/0033-295X.97.2.253

Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*. https://doi.org/10.1177/2515245920951503

Lehle, C., & Hübner, R. (2008). On-the-fly adaptation of selectivity in the flanker task. *Psychonomic Bulletin & Review, 15*, 814–818.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady, 10*(8), 707–710.

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speedaccuracy trade-offs(?). *Behavior Research Methods, 51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology, 6*, 1171.

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

Logan, G. D., & Zbrodoff, N. J. (1998). Stroop-type interference: Congruity effects in color naming with typewritten responses. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 978.

Lowe, D. G., & Mitterer, J. O. (1982). Selective and divided attention in a Stroop task. *Canadian Journal of Psychology/revue Canadienne De Psychologie, 36*(4), 684.

Mann, Y. (2015). Interactive music with Tone.js. In S. Goldszmid, N. Schnell, V. Saiz, & B. Matuszewski (Eds.), *Proceedings of the 1st annual Web Audio Conference*. IRCAM.

Pashler, H., & Harris, C. R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science, 7*(6), 531–536. https://doi.org/10.1177/1745691612463401

Reimers, S., & Stewart, N. (2016). Auditory presentation and synchronization in Adobe Flash and HTML5/JavaScript web experiments. *Behavior Research Methods, 48*(3), 897–908. https://doi.org/10.3758/s13428-016-0758-5

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638.

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin & Review, 26*(2), 452–467.

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

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

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology Section A, 47*(3), 631–650. https://doi.org/10.1080/14640749408401131

Weidler, B. J., Pratt, J., & Bugg, J. M. (2022). How is location defined? Implications for learning and transfer of location-specific control. *Journal of Experimental Psychology: Human Perception and Performance, 48*(4), 312.

Woods, K. J. P., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, & Psychophysics, 79*(7), 2064–2072. https://doi.org/10.3758/s13414-017-1361-2

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.