Canadian Journal of Experimental Psychology /
Revue canadienne de psychologie expérimentale

Proactive and Reactive Cognitive Control in the Absence of Learning and Memory Confounds: Evidence From a Cross-Modal Trial-Unique Stroop Task

Nicholaus P. Brosowsky

Online First Publication, April 11, 2024. https://dx.doi.org/10.1037/cep0000325

CITATION

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*. Advance online publication. https://dx.doi.org/10.1037/cep0000325

CANADIAN PSYCHOLOGICAL ASSOCIATION SOCIÉTÉ CANADIENNE DE PSYCHOLOGIE

Canadian Journal of Experimental Psychology /
Revue canadienne de psychologie expérimentale

ISSN: 1196-1961

https://doi.org/10.1037/cep0000325

Proactive and Reactive Cognitive Control in the Absence of Learning and Memory Confounds: Evidence From a Cross-Modal Trial-Unique Stroop Task

Nicholaus P. Brosowsky
Department of Psychology, University of Manitoba

Goal-directed behaviour is typically conceptualized as striking a balance between two antagonistic cognitive control states such as proactive and reactive control, as demonstrated by conflict phenomena such as the list-wide proportion congruency and congruency sequence effects. However, control-based explanations for these phenomena have come under criticism due to low-level associative regularities that are frequently confounded with conflict manipulations within these experimental designs. In the present study, a novel Stroop paradigm referred to as the "trial-unique Stroop task" was developed to examine whether these effects could be observed in the absence of low-level associative regularities. On each trial, participants typed a word they heard spoken aloud while ignoring a word visually displayed on the screen. Importantly, each word only appeared in a single trial throughout the experiment, and because stimuli and responses were never repeated, there were no low-level associative regularities across trials. Using this paradigm, we observed both congruency sequence (Experiment 1) and list-wide proportion congruency (Experiment 2) effects, providing the strongest evidence to date for control-based explanations of these phenomena. Split-half analyses revealed much higher reliability than traditional colour–word Stroop tasks for the congruency effect ($r_{SB} = .98$), the congruency sequence effect ($r_{SB} = .42$), and the list-wide proportion congruency effect ($r_{SB} = .85$). Moreover, the methodological advantages of the trial-unique Stroop task allow for the independent manipulation of task features related to control, learning, and memory processes. The promising results of this study support the application of the trial-unique Stroop task in this context and open new avenues for future research.

Public Significance Statement

Cognitive control is crucial in guiding behaviour amidst distractions and conflicting information. However, traditional methods struggle to isolate the effects of control from other learning processes because of confounding trial-to-trial regularities. This study introduces the "trial-unique Stroop task" as an effective tool for measuring cognitive control, free from learning and memory confounds, providing robust evidence and practical insights into how we effectively manage distractions and conflict.

Keywords: cognitive control, conflict adaptation, attention, proportion congruency effect, congruency sequence effect

Cognitive control enables flexible goal-directed behaviour by adjusting attention and actions to prioritize relevant over irrelevant information. Recent theoretical accounts conceptualize cognitive control as striking a balance between two antagonistic control states (Braem & Egner, 2018; Braver, 2012; Brosowsky & Crump, 2018; Brosowsky & Egner, 2021; Bugg & Crump, 2012; Diamond, 2013; Dreisbach, 2012; Goschke, 2013; Hommel, 2015). According to the Dual Mechanisms of Control Framework (Braver, 2012), for instance, there is a proactive mode of control, which refers to the top-down maintenance of goal-relevant information and adjustment of attention in preparation for a stimulus onset, and a reactive mode of control, which refers to the bottom-up reactivation of goal-relevant information and reactionary adjustment of attention following the onset of a stimulus. Proactive and reactive control are often viewed as antagonistic in the sense that engaging in reactive control can engender attentional flexibility, but at the cost of focused attentional stability provided by proactive control—and vice versa (see also Dreisbach, 2012). The desirability of biasing control is, therefore, dependent on the context and can be driven by both our goals and prior experiences (Bugg & Crump, 2012).

Adaptations in cognitive control are frequently studied in the context of interference paradigms like the Stroop task (Stroop, 1935; see Braem et al., 2019) where a colour word is presented in a coloured font (e.g., "RED" in blue font) and participants identify the colour of the font while ignoring the text. On some trials, the distractor is congruent with the target–response (e.g., "RED" in red

---

Nicholaus P. Brosowsky https://orcid.org/0000-0003-4573-5468

This research was supported by a Discovery Grant RGPIN-2022-04498 from the Natural Sciences and Engineering Research Council of Canada awarded to Nicholaus P. Brosowsky.

All data, materials, and code are available at https://osf.io/6jwe2.

Correspondence concerning this article should be addressed to Nicholaus P. Brosowsky, Department of Psychology, University of Manitoba, 190 Dysart Road, Winnipeg, MB R3T 2N2, Canada. Email: nicholaus.brosowsky@umanitoba.ca

font) and, on others, incongruent (e.g., "RED" in yellow font). Control is thought to be recruited in response to the level of perceived interference (e.g., Botvinick et al., 2001) and quantified by the size of the congruency effect—the difference in performance between incongruent and congruent trials. Modulations of this congruency effect are then taken as evidence of changes in control, such that smaller congruency effects reflect an increase in control and a shift of attention away from the distractor dimension (Braem et al., 2019; Bugg & Crump, 2012).

One phenomenon important for demonstrating shifts toward proactive control is the list-wide proportion congruency effect (Logan & Zbrodoff, 1979; Lowe & Mitterer, 1982). In this task, the frequency of conflict—the proportion of congruent versus incongruent trials—is manipulated globally across a block of trials. Congruency effects tend to be smaller in high-conflict lists as compared to low-conflict lists, which is explained in terms of proactive control: When the global level of conflict is high, participants prepare for conflict, proactively adjusting attention away from the distractor and reducing the congruency effect (Lindsay & Jacoby, 1994; Logan & Zbrodoff, 1979; Lowe & Mitterer, 1982; see also Bugg et al., 2008).

The congruency sequence effect, in contrast, has been important for demonstrating reactive control (Gratton et al., 1992; for a review, see Egner, 2007). The congruency sequence effect refers to the observation that congruency effects are smaller if the previous trial was incongruent rather than congruent. The congruency sequence effect is typically explained in terms of a reactive, ballistic shift of control in response to recently experienced conflict, which persists to affect performance on the following trial (Botvinick et al., 2001; but see Brosowsky & Crump, 2018).

Cognitive control explanations of these phenomena, however, have come under heavy scrutiny over the last 2 decades (for reviews, see Algom & Chajut, 2019; Algom et al., 2022; Braem et al., 2019; Schmidt, 2019). Numerous studies have now highlighted problematic experimental design characteristics that confound conflict manipulations with so-called "low-level" learning and memory manipulations. For example, in the traditional Stroop paradigm, when there are four words and colours presented with 50% congruency, individual congruent items appear more frequently than incongruent items. To the extent that a distractor is paired more frequently with one target over another, it can become predictive of a particular response or target feature (i.e., contingency learning, Schmidt & Besner, 2008; target–distractor correlations, Dishon-Berkovits & Algom, 2000; Melara & Algom, 2003). Participants learn these regularities, as evidenced by faster responding to frequency-biased stimuli and, because frequency biases are often confounded with congruency manipulations, offers an alternative explanation for a variety Stroop phenomena without appealing to control processes, including list-wide proportion and congruency sequence effects (see Algom et al., 2022; Schmidt, 2019).

In the list-wide proportion congruency paradigm, the frequency of conflict is manipulated by increasing or decreasing the proportion of congruent trials across a block of trials. However, due to limited stimulus set sizes, and an imbalance in the number of unique congruent versus incongruent stimuli, manipulating the relative frequency of incongruent to congruent stimuli also tends to alter stimulus–response contingency biases and target–distractor correlations in a manner that is confounded with the proportion congruency manipulation (for a review, see Spinelli & Lupker, 2021). Thus, traditional list-wide proportion congruency effects can be explained without invoking proactive control constructs.

Likewise, low-level associative regularities have also been shown to influence the congruency sequence effect. For instance, the magnitude and presence of the congruency sequence effect is sensitive to whether the response repeats from trial-to-trial (Mayr et al., 2003), whether target and/or distractor features repeat from trial-to-trial (Schmidt & Weissman, 2014), and trial-to-trial changes in contingency biases (Schmidt & De Houwer, 2011). These repetitions can also be confounded with trial-to-trial changes in congruency in traditional designs providing alternative explanations that do not require additional theoretical constructs like the conflict monitor (Botvinick et al., 2001) or reactive control (Braver, 2012).

A considerable amount of research has been dedicated to disentangling cognitive control from low-level associative mechanisms by creating so-called "confound minimized" designs (for a review, see Braem et al., 2019). For the congruency sequence effect, trial lists have been designed to control for the various trial-to-trial feature repetitions by excluding problematic transitions (Kim & Cho, 2014; Schmidt & Weissman, 2014; Weissman et al., 2014), and analysis strategies have been developed to try to control for repetitions post hoc (Notebaert & Verguts, 2007). For the list-wide proportion congruent design, stimuli are typically separated into a "biased" set, which includes the proportion manipulation with contingency biases and an "unbiased" transfer set, which does not. Evidence for proactive control comes from whether the congruency effect is manipulated in the transfer set (Bugg, 2014; Hutchison, 2011; Tang et al., 2023).

Despite recent successes using confound-minimized designs, however, some remain sceptical, arguing that low-level associative mechanisms still provide a more general and parsimonious account (Algom et al., 2022). It certainly can be difficult to completely rule out all potential confounds and creating idiosyncratic trial lists that control for one confound have been known to inadvertently introduce another (see Tomat et al., 2021). Additionally, many interference paradigms produce relatively small effects that are often diminished further in confound-minimized variants (e.g., Bugg, 2014). Small effect sizes make it difficult to reliably observe the intended effect (Weissman et al., 2014) and limit their usefulness because of the difficulty in reliably measuring experimental manipulations of those effects (Brosowsky & Crump, 2021; Bugg et al., 2020; Crump et al., 2017). Some researchers have also implied that the reduction of effect sizes is further evidence that cognitive control is a poor explanation of these phenomena (e.g., Schmidt, 2019).

Finally, deconfounding low-level learning and memory effects does not eliminate their presence. These designs still include regularities that can be learned and, although we expect they are no longer confounded with the congruency manipulation, it is unclear to what extent that adaptations in control may depend on the presence (or absence) of such learning signals. For instance, in the transfer variant of the list-wide proportion congruency paradigm, the presence of contingency biases and predictability of the stimuli in the biased item set may result in a shift of attention between dimensions (e.g., Le Pelley et al., 2013). Alternatively, as others have suggested, the presence of such learning signals may dampen the use of control processing because control is only engaged as a last resort (Bugg et al., 2008).

In the present study, a novel task has been designed to circumvent these issues by using trial-unique stimuli and responses. The trial-unique Stroop task is a cross-modal paradigm, where a word is presented on the screen and another is heard spoken aloud. Participants are tasked with typing the word they hear and ignoring the word they see. On congruent trials, the visually presented word was identical to the auditory word, and on incongruent trials, the words differed. Importantly, every word is only seen and/or heard in a single trial across the experiment, and because stimuli and responses are never repeated, there are no low-level associative regularities across trials. More specifically, because stimuli and responses are never repeated, there are no trial-to-trial regularities for participants to learn and, as such, there are no stimulus–response contingencies, target–distractor or distractor–response correlations, and no trial-to-trial repetitions of features and/or responses.

The trial-unique Stroop task incorporates several critical features necessary to differentiate it from non-Stroop selective attention tasks. For example, it involves a logical–semantic relationship between dimensions and requires a semantic analysis of both the targets and distractors (Algom et al., 2022). Similarly, much like traditional Stroop tasks, the current task entails interference at multiple levels, including (a) the task selection level, due to conflict between the visual and auditory processing channels, (b) the stimulus selection level, due to perceptual and semantic conflict between the relevant and irrelevant dimensions, and (c) the level of response selection between two typed responses activated by the visual and auditory word features (e.g., Viviani et al., 2022). As such, the task would also meet the definition of a "type-eight ensemble" according to Kornblum et al.'s (1990) classification. Finally, previous research has consistently shown that interference effects are larger for vocal responses than for manual responses (e.g., Sharma & McKenna, 1998; White, 1969). This finding suggests that managing interference from competing speech codes is more challenging than manual keypresses (Augustinova et al., 2019; see also Roelofs, 2005). In the present study, typed responses were chosen to replicate the speech production effect and maintain consistency with traditional Stroop tasks.

The trial-unique Stroop task, therefore, provides a strong experimental test of control-based explanations of Stroop phenomena by effectively preventing trial-to-trial associative regularities and maintaining critical task features that distinguish Stroop from non-Stroop stimuli. The aim of the present study was to conceptually replicate and extend the work by Spinelli et al. (2019) using the cross-modal trial-unique Stroop paradigm. In Experiment 1, a 50% congruent paradigm was used to determine whether congruency sequence effects could be observed in the trial-unique Stroop task, where there were no trial-to-trial stimulus and/or response repetitions (preregistration is available at https://osf.io/hw7av; Brosowsky, 2023). If low-level associative mechanisms are sufficient for explaining congruency sequence effects, there should be no trial-to-trial modulations in the congruency effect within this design. In Experiment 2, the proportion congruency was manipulated across blocks of trials (80% congruent vs. 20% congruent) to examine whether list-wide proportion congruency effects would be observed when there are no low-level learning biases present (preregistration is available at https://osf.io/j9n56; Brosowsky, 2023). Again, if contingency and/or correlation biases are necessary for producing list-wide effects, there ought to be no differences in the size of the congruency effect across blocks.

Experiment 1: Congruency Sequence Effects

Method

Sample Size Rationale

Sample sizes were determined by estimating the power to detect a range of effects using a Monte Carlo simulation approach (Brosowsky & Egner, 2021; Brosowsky et al., 2021; Crump et al., 2017). The ex-Gaussian distribution parameters ($\mu = 661$ ms, $\sigma = 103$ ms, and $\tau = 117$ ms) and variance–covariance structure representative of participant responses were taken from a prior typing Stroop task (Crump et al., 2017; Experiment 2b). For every simulated individual, ex-Gaussian parameters were drawn from a multivariate normal distribution, ensuring that no parameter was below the least observed participant parameters in Crump et al. (2017). Four hundred eighty response times were then sampled from the ex-Gaussian distribution for each. The congruency sequence effect was of primary interest. Trials were categorized into four within-subject conditions (trial $N - 1$ congruency and trial N congruency), and congruency sequence effects were generated by increasing the congruency effect in the trial $N - 1$ congruent condition by a set effect size (e.g., 25 ms). For each combination of effect size and sample size, 1,000 simulations were run, analyzing the resultant data with a linear mixed-effect model, considering $N - 1$ congruency and congruency as fixed effects and individual as a random effect.

From these simulations, it was estimated that a minimum of 50 participants was necessary to detect congruency sequence effects of 20 ms with 90% power (see also Brosowsky & Crump, 2021). All code is available online at https://osf.io/6jwe2, and the preregistration is available at https://osf.io/hw7av (Brosowsky, 2023). The goal was to gather between 50 and 100 participants over 2 weeks, concluding either when 100 individuals finished the experiment or at the end of 2 weeks, depending on which occurred earlier. After the 2-week duration, 55 participants completed the study.

Participants

Participants were 55 University of Manitoba undergraduate students who completed the experiment online for course credit. Participant ages ranged from 18 to 21 years. When asked to indicate their gender using a free response questionnaire, 31 participants identified as a woman or female, 22 as a man or male, and two chose not to respond. Additionally, 50 participants indicated that they were right-handed, four left-handed, and one ambidextrous. The experiment was approved by the University of Manitoba Research Ethics Board.

Stimuli

All experiments were programmed in JavaScript. Audiovisual stimuli were presented using the Tone.js JavaScript library (https://tonejs.github.io/; Mann, 2015; see Brosowsky et al., 2023; Reimers & Stewart, 2016). Tone.js delivers audio stimuli by "looking ahead" and scheduling the auditory presentation using the highly accurate audio clock. Visual stimuli were scheduled using this same look ahead method, synced with the audio clock, and presented using the requestAnimationFrame method to provide accurate audiovisual synchronization.

Word and auditory stimuli were selected from the Auditory English Lexicon Project database (Goh et al., 2020) using their online word generation site (https://inetapps.nus.edu.sg/aelp/generate). Words were restricted to four-, five-, or six-letter words (full selection criteria are available on the additional online material at https://osf.io/6jwe2). The stimulus selection process produced a total of 2,218 unique words. The full stimulus set is available online.

For each participant, 240 words were randomly sampled without replacement for the congruent trials (80 four-letter words, 80 five-letter words, and 80 six-letter words) and 480 words were randomly sampled without replacement for the incongruent trials (160 four-letter words, 160 five-letter words, and 160 six-letter words). On congruent trials, the same word was presented visually and spoken aloud. For each incongruent pair, the two words selected were required to have the same number of letters, different first letters, and a Levenshtein's edit distance of less than three (Levenshtein, 1966). If a pair of words were randomly selected that did not meet these criteria, words were redrawn until a pair was selected that did meet the criteria. No word appeared in more than one trial throughout the experiment.

Task and Procedure

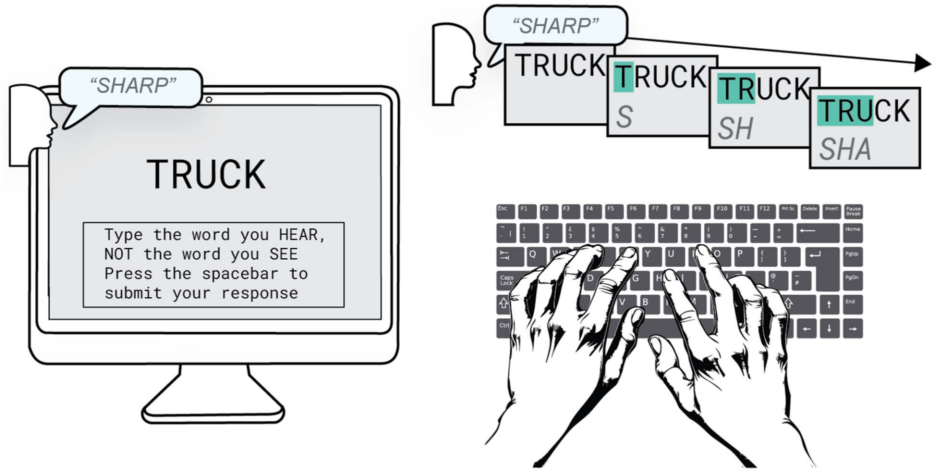
Participants first completed an audio test to confirm that they could hear and respond to the words. In the first part of the audio test, they heard the word "test" spoken aloud and were instructed to adjust their audio until it was at a comfortable level. Next, they typed four-letter words that had been spoken aloud. When each trial began, four Xs appeared on the screen. As they responded, the four Xs changed to the letter they had typed, giving them a visual indication of their responses. After typing four letters, they pressed the spacebar to submit their responses. To progress to the experiment, participants had to complete 10 test trials and type 7/10 accurately. They had to repeat the audio typing test if they did not meet this accuracy threshold.

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 word they heard and ignore the word that was presented on the screen (a task reminder was presented below the stimulus throughout the experiment; see Figure 1). During each trial, the background colour of each letter changed from grey to green as participants typed their responses, indicating how many letters they had typed. However, no visual feedback was provided to indicate which letters were typed correctly or incorrectly. Participants were not able to use the "backspace" or "delete" keys to undo their keypresses. Instead, they were instructed to continue typing the word if they made an error, and their accuracy would be evaluated based on how closely their response matched the target word. Once participants had typed the minimum number of letters required for the trial, they could submit their response by pressing the spacebar, which triggered the start of the next trial. Participants completed a total of 481 trials. Trial lists were constructed using a custom script that created lists with an equal number of congruency transitions (one additional trial is added to the beginning to give an equal number of congruency transitions).

Following the experimental phase, participants completed the Mindful Attention and Awareness Scale (Brown & Ryan, 2003) and the Deliberate and Spontaneous Mind Wandering scales (Carriere et al., 2013) followed by a demographics questionnaire

Figure 1

An Illustration of the Trial-Unique Stroop Task



Note. In each trial, a word was displayed on the screen and spoken aloud. A participant's task was to type the word they heard while ignoring the word they saw on the screen. As participants typed their responses, the background colour of each letter changed from grey to green, allowing them to keep track of how many letters they had typed, though no visual feedback indicated which letters they had typed. The typed letters "SHA" are shown in the figure for illustration only and were not shown to participants. Participants pressed the spacebar to submit their responses after typing the correct number of letters. See the online article for the color version of this figure.

and a debriefing. The questionnaires were included for exploratory purposes, and the data are available online, though not reported here (https://osf.io/6jwe2).

Data Analysis

All data, analysis, and article preparation code can be found at https://osf.io/6jwe2. Participants with less than 60% accuracy were excluded from all analyses. This removed three participants. Prior to all analyses, the first trial for each block was removed, and any trials with a response time greater than 3 s or less than 300 ms were removed (removing 0.96% of observations). In addition, for response time analyses, all error trials were removed before applying the Van Selst and Jolicoeur nonrecursive outlier removal (Van Selst & Jolicoeur, 1994), removing an additional 3.2% of observations.

The main dependent measure of interest was the balanced integration score (Liesefeld et al., 2015), which integrates both speed and accuracy, and is relatively insensitive to speed–accuracy trade-offs (Liesefeld & Janczyk, 2019). This measure was chosen because of speed–accuracy trade-offs observed in pilot data—the reaction time (RT) and error analyses are also reported, however. The balanced integration score is calculated by standardizing the mean reaction time and accuracy and then subtracting one standardized score from the other. As in Liesefeld et al. (2015), the proportion accuracy was subtracted from the mean reaction time such that it was comparable in direction to Stroop reaction time effects (i.e., higher balanced integration scores [BIS] indicate poorer performance). As in previous typing Stroop tasks, for all analyses, reaction times refer to the response initiation time (Crump et al., 2017; Logan & Zbrodoff, 1998), and accuracy was determined by the first letter typed (Crump et al., 2017).

Note the deviation from the preregistered analysis plan. In the preregistration, it was indicated that full-word typing accuracy would serve as the primary accuracy measure, with first-letter accuracy as a secondary measure. However, upon further consideration, it became clear that the first-letter accuracy was a more appropriate metric. First-letter accuracy shows whether participants began the correct response, which seems more informative than identifying spelling errors—especially given that participants did not receive visual feedback while typing. This metric is also more inclusive (1.9–2.9% errors vs. 9.4–14.6% errors) and aligns with how typing responses have been coded in similar studies (e.g., Crump et al., 2017; Logan & Zbrodoff, 1998). Nevertheless, the full-word accuracy analysis is available online: Using the full-word accuracy, both the BIS and RT interactions proved significant ($p < .001$), whereas the error percentage interaction did not ($p > .05$).

Results

Congruency Sequence Effects

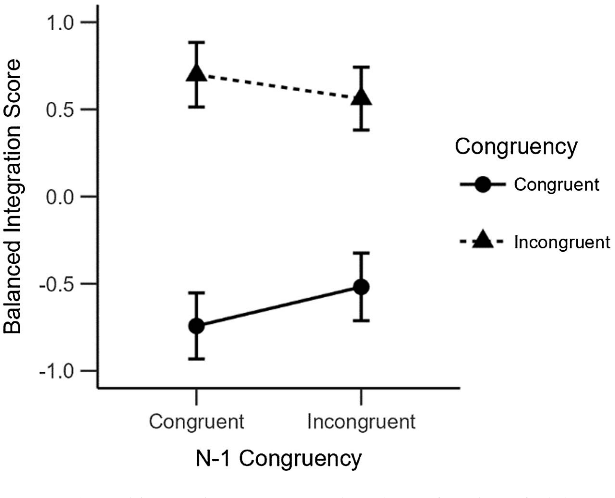
Balanced Integration Scores. BISs were analyzed using a 2 × 2 repeated measures analysis of variance with $N - 1$ congruency (congruent, incongruent) and congruency (congruent, incongruent) as the factors. The analysis resulted in a significant main effect of congruency, $F(1, 52) = 253.46$, $MSE = 0.33$, $p < .001$, $\eta_p^2 = .830$, 95% CI [0.75, 0.87], no significant main effect of $N - 1$ congruency, $F(1, 52) = 1.46$, $MSE = 0.07$, $p < .232$, $\eta_p^2 = .027$, 95% CI [0, 0.13], and a significant interaction between $N - 1$ congruency and congruency, $F(1, 52) = 28.25$, $MSE = 0.06$, $p < .001$, $\eta_p^2 = .352$, 95% CI [0.18, 0.49]. As evident in Figure 2 and Table 1, the congruency effect was larger following a congruent trial than an incongruent trial. Analyzing each of the $N - 1$ congruency conditions separately resulted in a significant interference effect for both the $N - 1$ congruent, $\Delta M = -1.44$, 95% CI [−1.97, −0.92], $t(103.96) = -5.44$, $p < .001$, and $N - 1$ incongruent, $\Delta M = -1.08$, 95% CI [−1.61, −0.55], $t(103.46) = -4.08$, $p < .001$, conditions.

Reaction Times. Reaction times and error rates were also analyzed separately. For the reaction times, a generalized linear mixed-effects model was used with an inverse Gaussian distribution and identity link function (see Lo & Andrews, 2015), $N - 1$ congruency and congruent conditions as fixed effects (dummy-coded; incongruent = 1 and congruent = 0; $N - 1$ incongruent = 0, $N - 1$ congruent = 1) and subjects as a random effect. Here, the main effect of congruency was significant ($\beta = 227.87$, 95% CI [220.4, 235.34], $t = 59.81$, $p < .001$); the main effect of $N - 1$ congruency was significant ($\beta = -42.5$, 95% CI [−48.24, −36.75], $t = -14.49$, $p < .001$); and the interaction between $N - 1$ congruency and congruency was significant ($\beta = 47.61$, 95% CI [37.28, 57.95], $t = 9.03$, $p < .001$). Following up the significant interaction, there was a significant congruency effect for $N - 1$ congruent trials ($\Delta M = -275.49$, 95% CI [−283.12, −267.85], $z = -70.68$, $p < .001$) and $N - 1$ incongruent trials ($\Delta M = -227.87$, 95% CI [−235.34, −220.40], $z = -59.81$, $p < .001$).

Error Rates. Finally, using the same model specifications, a generalized linear mixed-effects model with a binomial distribution was used to analyze error rates. This analysis resulted in a significant main effect of congruency ($\beta = 0.18$, 95% CI [−0.01, 0.36], $t = 1.87$, $p < .061$); no significant main effect of $N - 1$ congruency ($\beta = -0.04$, 95% CI [−0.23, 0.16], $t = -0.36$, $p < .72$); and a significant interaction between $N - 1$ congruency and congruency ($\beta = 0.26$, 95% CI [0, 0.52], $t = 1.97$, $p < .048$). Following up the significant interaction, there was a significant congruency effect for $N - 1$

Figure 2

Results From Experiment 1: Congruency Sequence Effect



Note. Balanced integration scores are plotted as a function of trial $N - 1$ congruency and congruency. Error bars represent standard errors.

Table 1
Experiment 1: Results

		Trial *N* congruency			
Dependent measure	Trial $N-1$ congruency	Incongruent	Congruent	Congruency effect	CSE
BIS	Incongruent	0.56 (0.18)	−0.52 (0.19)	1.08 (0.08)	0.36^* (0.07)
	Congruent	0.70 (0.19)	−0.74 (0.19)	1.44 (0.09)	
RT	Incongruent	1375.15 (6.92)	1147.28 (6.64)	227.87 (4.03)	47.61^* (5.71)
	Congruent	1380.27 (6.92)	1104.78 (6.57)	275.49 (4.11)	
ER	Incongruent	2.31 (0.38)	1.94 (0.33)	0.37 (0.26)	0.63^* (0.30)
	Congruent	2.87 (0.47)	1.87 (0.32)	1.00 (0.20)	

Note. RTs and ERs are estimated using generalized linear models; standard errors are presented in brackets. CSE = congruency sequence effect; BIS = balanced integration score; RT = reaction time (ms); ER = error percentage.
$^* p < .05$.

congruent trials, $\log(OR) = -0.44$, 95% CI $[-0.62, -0.26]$, $z = -4.76$, $p < .001$, but not the $N-1$ incongruent trials, $\log(OR) = -0.18$, 95% CI $[-0.36, 0.01]$, $z = -1.87$, $p < .061$.

Split-Half Reliability. Since the goals of this study were, in part, methodological, the internal consistency of the congruency and congruency sequence effects was assessed using split-half reliability estimates from the “splithalf” R package (Parsons, 2021). Using 5,000 random splits of the data, the reliability was assessed with the mean Spearman–Brown correction correlations for reaction times, error rates, and inverse-transformed reaction times. The resulting reliability scores for the congruency effects were high for the reaction times ($r_{\text{SB}} = 0.97$, 95% CI [0.96, 0.98]) and inverse-transformed reaction times ($r_{\text{SB}} = 0.98$, 95% CI [0.98, 0.99]) and lower for the error rates ($r_{\text{SB}} = 0.52$, 95% CI [0.23, 0.71]). Similarly, for the congruency sequence effects, there was modest reliability for the reaction times ($r_{\text{SB}} = 0.45$, 95% CI [0.19, 0.66]) and inverse-transformed reaction times ($r_{\text{SB}} = 0.42$, 95% CI [0.15, 0.63]) and lower reliability for the error rates ($r_{\text{SB}} = 0.3$, 95% CI $[-0.09, 0.58]$).

Experiment 2: List-Wide Proportion Congruency Effects

Method

Participants

Participants were 58 University of Manitoba undergraduate students who completed the experiment online for course credit. Participant ages ranged from 18 to 21 years. When asked to indicate their gender using a free response, 39 participants identified as a woman or female, 16 as a man or male, one as nonbinary, and two chose not to respond. Forty-seven participants indicated that they were right-handed, eight left-handed, two ambidextrous, and one chose not to respond. The experiment was approved by the University of Manitoba Research Ethics Board.

Stimuli

The stimuli were identical to Experiment 1.

Task and Procedure

The task and procedure were largely the same as in Experiment 1. However, the experimental phase was split into two blocks of 240 trials, one of which contained 80% congruent trials (the mostly congruent block [MC]) and the other 20% congruent trials (the mostly incongruent block [MI]). The order of blocks was randomized across participants. All other aspects of the task and procedure were the same as in Experiment 1.

Data Analysis

One participant was removed with an accuracy of less than 60%. Trials with a response time greater than 3 s or less than 300 ms were removed (removing 1.02% of observations). In addition, for response time analyses, all error trials were removed before applying the Van Selst and Jolicoeur nonrecursive outlier removal (Van Selst & Jolicoeur, 1994), removing an additional 2.94% of observations.

Results

List-Wide Proportion Congruency Effects

Balanced Integration Scores. Analyzing the balanced integration scores using a 2 × 2 repeated measures analysis of variance with proportion congruency (80% PC, 20% PC) and congruency (congruent, incongruent) as the factors resulted in a significant main effect of congruency, $F(1, 56) = 172.44$, $MSE = 0.51$, $p < .001$, $\eta_p^2 = .755$, 95% CI [0.65, 0.81]; no significant main effect of proportion congruency, $F(1, 56) = 2.61$, $MSE = 1.06$, $p < .112$, $\eta_p^2 = .045$, 95% CI [0, 0.16]; and a significant interaction between proportion congruency and congruency, $F(1, 56) = 43.77$, $MSE = 0.38$, $p < .001$, $\eta_p^2 = .439$, 95% CI [0.27, 0.56].[1] As evident in Figure 3 and Table 2, there was a significantly larger congruency effect in the mostly congruent block as compared to the mostly congruent block. Analyzing each of the proportion congruency blocks separately, there was a significant interference effect for both the mostly congruent, $\Delta M = -1.79$, 95% CI $[-2.20, -1.37]$, $t(96.94) = -8.60$, $p < .001$, and the mostly incongruent, $\Delta M = -0.70$, 95% CI $[-1.22, -0.17]$, $t(111.83) = -2.64$, $p < .009$, blocks.

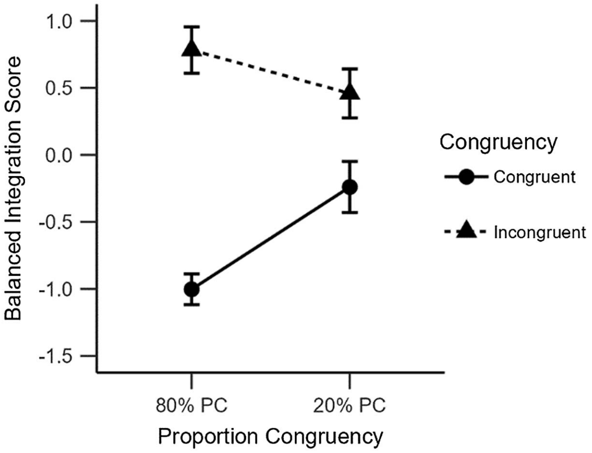
Reaction Times. Reaction times were analyzed using a generalized linear mixed-effects model with an inverse Gaussian distribution and identity link function (see Lo & Andrews, 2015), proportion congruency and congruent conditions as fixed effects (contrast coding: incongruent = 1 and congruent = 0; 80% PC = 1

---

[1] Using the full word typing accuracy, the BIS and RT interactions were significant ($p < .001$), but the ER interaction was not ($p = .08$).

Figure 3

Results From Experiment 2: List-Wide Proportion Congruency (PC) Effects



Note. Balanced integration scores are plotted as a function of PC (80% PC and 20% PC) and congruency. Error bars represent standard errors.

and 20% PC = 0) and subjects as a random effect. Here, the main effect of congruency was significant ($\beta = 166.43$, 95% CI [159.8, 173.07], $t = 49.19$, $p < .001$); the main effect of proportion congruency was significant ($\beta = -124.41$, 95% CI [−130.62, −118.19], $t = -39.22$, $p < .001$); and the interaction between proportion congruency and congruency was significant ($\beta = 137.06$, 95% CI [128.42, 145.7], $t = 31.09$, $p < .001$). Following the significant interaction, there was a significant difference between congruent and incongruent trials for both the mostly congruent block ($\Delta M = -303.49$, 95% CI [−312.24, −294.75], $z = -68.02$, $p < .001$) and the mostly incongruent block ($\Delta M = -166.43$, 95% CI [−173.07, −159.80], $z = -49.19$, $p < .001$).

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 ($\beta = -0.08$, 95% CI [−0.35, 0.2], $t = -0.55$, $p < .581$); no significant main effect of proportion congruency ($\beta = -0.46$, 95% CI [−0.75, −0.18], $t = -3.19$, $p < .001$); and a significant interaction between proportion congruency and congruency ($\beta = 1.19$, 95% CI [0.83, 1.55], $t = 6.47$, $p < .001$). Following the significant interaction, there was a significant difference between congruent and incongruent trials for both the mostly congruent block, $\log(OR) = -1.11$, 95% CI [−1.35, −0.88], $z = -9.33$, $p < .001$, but not the mostly incongruent block, $\log(OR) = 0.08$, 95% CI [−0.20, 0.35], $z = 0.55$, $p < .581$.

Split-Half Reliability. Following the exploratory analyses in Experiment 1, the internal reliability was assessed using 5,000 random splits of the data and estimating the reliability with the mean Spearman–Brown correction correlations for reaction times, error rates, and inverse-transformed reaction times. The resulting reliability scores for the congruency effects were high for the reaction times ($r_{SB} = 0.96$, 95% CI [0.94, 0.98]) and inverse-transformed reaction times ($r_{SB} = 0.98$, 95% CI [0.97, 0.99]) and lower for the error rates ($r_{SB} = 0.61$, 95% CI [0.39, 0.76]). For the list-wide proportion congruency effects, I, again, observed high reliability for reaction times ($r_{SB} = 0.76$, 95% CI [0.63, 0.85]) and inverse-transformed reaction times ($r_{SB} = 0.85$, 95% CI [0.78, 0.9]) and lower reliability for the error rates ($r_{SB} = 0.65$, 95% CI [0.35, 0.82]).

General Discussion

In the present study, a new Stroop paradigm was developed where each stimulus and response were presented on a single trial and never repeated throughout the experiment. This design effectively eliminated low-level associative regularities, including stimulus–response contingencies, stimulus feature correlations, and trial-to-trial feature/response repetitions across trials (see Braem et al., 2019). Using this paradigm, robust evidence was found for both the congruency sequence (Experiment 1) and list-wide proportion congruence (Experiment 2) effects. These experiments offer the most compelling evidence to date for proactive and reactive control in the absence of low-level learning and memory confounds. Additionally, the list-wide proportion congruency effect was observed without a contingency-biased item set (Braem et al., 2019), which suggests that the presence of low-level learning signals was not necessary for control adaptations in this paradigm (e.g., Le Pelley et al., 2013).

From a methodological perspective, the trial-unique Stroop task is a compelling alternative to confound-minimized designs (Braem et al., 2019) with a few noteworthy advantages. For one, the trial-unique Stroop task does not require idiosyncratic changes to the stimulus set or trial list to investigate different control effects. Here, for example, nearly identical paradigms were used to demonstrate the congruency sequence and list-wide proportion congruent

Table 2

Experiment 2: Results

Dependent measure	PC block	Congruency Incongruent	 Congruent	Congruency effect	LWPC
BIS	MI	0.46 (0.18)	−0.24 (0.19)	0.70 (0.08)	1.09* (0.16)
	MC	0.78 (0.17)	−1.00 (0.11)	1.79 (0.16)	
RT	MI	1376.10 (9.06)	1209.67 (9.43)	166.43 (4.92)	137.06* (7.45)
	MC	1388.75 (9.28)	1085.26 (8.81)	303.49 (5.44)	
ER	MI	1.79 (0.23)	1.93 (0.32)	−0.14 (0.26)	2.55* (0.51)
	MC	3.63 (0.51)	1.22 (0.16)	2.41 (0.42)	

Note. RTs and ERs are estimated using generalized linear models; standard errors are presented in brackets. PC = proportion congruency; LWPC = list-wide proportion congruency effect; BIS = balanced integration score; MI = mostly incongruent block; MC = mostly congruent block; RT = reaction time (ms); ER = error percentage.
$^*p < .05$.

effects—without the need for separate subsets of trials such as "inducer" and "diagnostic." The task could easily be extended further to examine other effects such as the context-specific proportion congruency effect (Crump et al., 2017) with minimal changes to the experimental design or stimuli. Another important advantage in using this task is how easily learning signals, such as stimulus–response contingencies, can be reintroduced into the experimental design. Thus, one can investigate the relationship between control, learning, and memory processes by manipulating conflict and learning signals orthogonally.

The trial-unique Stroop task is similar in concept to one employed by Spinelli et al. (2019), in which participants used vocal responses to categorize and name images, while ignoring superimposed words. Spinelli et al. used this vocal naming task to demonstrate the presence of list-wide proportion congruency effects. Despite similarities with this previous paradigm, however, the present study introduces some key differences.[2] Firstly, the present study investigates the use of cognitive control in a cross-modal context. Although it is commonly assumed that cognitive phenomena will generalize across modalities, this assumption is not always valid (e.g., Spence & Driver, 1997), and examining cognitive control across modalities warrants investigation in its own right (e.g., Roelofs, 2005). Secondly, to create task conflict, as previously discussed, responses were collected via typing rather than vocalizing (Tang et al., 2023). This approach is more practical for collecting responses in both online and laboratory settings. Successfully demonstrating the effectiveness of typed responses in this context represents a noteworthy methodological advancement. Thirdly, using words rather than pictures as targets also provides practical advantages. The number of easily nameable pictures is limited compared to typable words. In addition, the relationships between word stimuli are less ambiguous than those between pictures and words, and stimulus features such as semantic similarity are more readily controlled for and manipulated compared to pictures, which can have multiple interpretations and may be less clear-cut. Finally, Spinelli et al. used categories of pictures/words that were predictive of the congruency proportion, which has been a point of contention in the interpretation of their results (see Schmidt, 2021). However, this issue is not applicable in the present study, which adds to its methodological strengths.

Moreover, the split-half reliability estimates were remarkably encouraging. The reliability of the reaction time congruency effect, which is a difference score, was consistently high in both experiments, ranging from .96 to .98. Of greater interest, however, is the fact that both the congruency sequence ($r_{\mathrm{SB}} = .42$) and list-wide proportion ($r_{\mathrm{SB}} = .85$) effects exhibited relatively high reliability—particularly for being a difference of difference scores. To put these results into context, it is worth noting that the issue of poor reliability has been a topic of concern in recent research using conflict paradigms (Hedge et al., 2018; Rouder & Haaf, 2019). For instance, Hedge et al. (2018) conducted a study on congruency effects across various tasks and observed reliability estimates from 0.36 to 0.77. Another study by Whitehead et al. (2019) found that the reliability of congruency sequence effects within a traditional Stroop task was "nonexistent," with r_{SB}s ranging from −0.1 to 0.01. Although the reliability of the list-wide proportion effect has not been previously reported, Snijder et al. (2023) reported the split-half reliability of the congruency effect within a mostly incongruent condition to be 0.68. Similarly, Bejjani and Egner (2021) reported the correlations between diagnostic and inducer congruency effects in mostly congruent and mostly incongruent lists to be 0.32 and 0.29, respectively. Therefore, a reliability of 0.85 for the full list-wide proportion congruency effect is extremely promising and puts it above the clinical standard for a measure to be considered excellent. The split-half reliability method, however, evaluates only the internal consistency of a single administration of the task, potentially not capturing the complete reliability of the task over time. Subsequent research could consider using the test–retest reliability method to determine its appropriateness for individual differences approaches.

Another important limitation of the present study is that the experiments were conducted online, and thus, we cannot verify how participants executed the experiment. For example, in the mostly incongruent block of Experiment 2, participants might not have been actively looking at the screen, resulting in smaller average Stroop effects. However, several reasons suggest that this was not the case. First, the task was structured to ensure participants remained engaged with the screen throughout the task. Participants had to type the correct number of letters before submitting, and visual feedback was given to signify a keypress; the background colour of each letter would change as it was typed. This means participants needed to monitor the screen to confirm their keypress was registered and to know when they could submit their responses.

Further supporting this argument is the data itself. In the mostly incongruent block, there was a significant and large Stroop effect (166 ms). If participants had routinely looked away from the screen, we would anticipate similar performance on both congruent and incongruent trials. The presence of a large Stroop effect demonstrates that participants performed markedly better on congruent trials and suggests that they were actively reading the words, even in the mostly incongruent trials. Moreover, a secondary analysis was conducted, specifically examining participants who displayed the most pronounced congruency effects during the mostly incongruent block (>150 ms, $N = 26$). Large MI congruency effects suggest sustained reading throughout the block for these individuals. If the noted interaction effect largely results from participants looking away from the screen, then this interaction should disappear when analyzing only those with pronounced MI congruency effects. However, upon reanalysis with this subset of participants, a strong interaction effect remained ($p < .001$). This persistent interaction effect among these participants indicates that the phenomenon is not solely attributed to those with the smallest congruency effects in the block. Nevertheless, future studies should validate this task in a laboratory environment to ensure participants engage with the task as intended.

Lastly, in the current version of the task, participants typed responses to an auditory target while ignoring a visual distractor. Typing responses were chosen because they are simple to collect online and have previously been used in the context of the Stroop paradigm (Crump et al., 2017; Logan & Zbrodoff, 1998). Presumably, however, one could instead use vocal responses to a visual target while ignoring an auditory distractor, retaining the target–response modality incompatibility. Such a design might be better suited for special populations, like children or clinical patients where vocal responding is more appropriate

---

[2] The author thanks Dr. Giacomo Spinelli for his helpful comments and suggestions on a previous version of this article.

(see Strauss et al., 2006). Still, more research is required to fully understand the psychometric properties of the task and the necessary conditions for using vocal responses in this context. Future research using typing would also benefit from making use of the substantial body of typing research to develop more accurate response models (see Logan, 2018). For example, by considering individual differences in typing ability (Behmer & Crump, 2016) or patterns in language structure and learning (Crump et al., 2019), researchers may be able to better isolate response variability resulting from conflict-driven control processes.

Résumé

Les comportements orientés vers un but sont habituellement conceptualisés comme un juste équilibre trouvé entre deux états de contrôle cognitif antagonistes, par exemple le contrôle proactif et le contrôle réactif, comme démontré par des phénomènes liés aux conflits, comme l'effet de congruence proportionnelle dans l'ensemble de la liste (*list-wide proportion congruency*) et celui de séquence de congruence. Cependant, les explications fondées sur le contrôle de ces phénomènes ont suscité des critiques en raison de régularités associatives de faible niveau qui sont fréquemment confondues avec des manipulations du conflit au sein de ces conceptions expérimentales. Aux fins de la présente étude, un paradigme de Stroop novateur appelé « tâche Stroop unique à l'essai » a été élaboré pour examiner si ces effets pouvaient être observés en l'absence de régularités associatives de faible niveau. Lors de chaque essai, les participants devaient taper un mot qu'ils avaient entendu prononcé à voix haute tout en ignorant un mot affiché à l'écran. Il importe de noter que chaque mot n'apparaissait que dans un seul essai tout au long de l'expérience. Aussi, comme les stimuli et les réponses n'étaient jamais répétés, il n'existait aucune régularité associative de faible niveau d'un essai à l'autre. En utilisant ce paradigme, nous avons observé à la fois des effets de séquence de congruence (expérience 1) et de congruence proportionnelle dans l'ensemble de la liste (expérience 2), fournissant ainsi les preuves les plus concluantes à ce jour quant aux explications fondées sur le contrôle de ces phénomènes. Des analyses par partage ont mis en lumière un degré beaucoup plus élevé de fiabilité que celui fourni par des tâches Stroop couleur-mot traditionnelles pour l'effet de congruence ($r_{SB} = 0{,}98$), l'effet de séquence de congruence ($r_{SB} = 0{,}42$), et l'effet de congruence proportionnelle dans l'ensemble de la liste ($r_{SB} = 0{,}85$). De plus, les avantages méthodologiques de la tâche Stroop unique à l'essai permettent la manipulation indépendante des caractéristiques de la tâche liée aux processus de contrôle, d'apprentissage et de mémorisation. Cette étude et ses résultats prometteurs soutiennent l'application de la tâche Stroop unique à l'essai dans ce contexte, et ouvrent de nouvelles perspectives pour la recherche.

Mots-clés : contrôle cognitif, adaptation aux conflits, attention, effet de congruence proportionnelle, effet de séquence de congruence

References

Algom, D., & Chajut, E. (2019). Reclaiming the Stroop effect back from control to input-driven attention and perception. *Frontiers in Psychology*, *10*, Article 1683. https://doi.org/10.3389/fpsyg.2019.01683

Algom, D., Fitousi, D., & Chajut, E. (2022). Can the Stroop effect serve as the gold standard of conflict monitoring and control? A conceptual critique. *Memory & Cognition*, *50*(5), 883–897. https://doi.org/10.3758/s13421-021-01251-5

Augustinova, M., Parris, B. A., & Ferrand, L. (2019). The loci of Stroop interference and facilitation effects with manual and vocal responses. *Frontiers in Psychology*, *10*, Article 1786. https://doi.org/10.3389/fpsyg.2019.01786

Behmer, L. P., Jr., & Crump, M. J. C. (2016). Crunching big data with fingertips: How typists tune their performance toward the statistics of natural language. In M. N. Jones (Ed.), *Big data in cognitive science* (pp. 329–345). Psychology Press.

Bejjani, C., & Egner, T. (2021). Evaluating the learning of stimulus-control associations through incidental memory of reinforcement events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(10), 1599–1621. https://doi.org/10.1037/xlm0001058

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Evaluating the demand for control: Anterior cingulate cortex and conflict monitoring. *Psychological Review*, *108*(3), 624–652. https://doi.org/10.1037/0033-295X.108.3.624

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

Braem, S., & Egner, T. (2018). Getting a grip on cognitive flexibility. *Current Directions in Psychological Science*, *27*(6), 470–476. https://doi.org/10.1177/0963721418787475

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010

Brosowsky, N. P. (2023, November 1). *Cross-modal Stroop with trial-unique stimuli*. Open Science Framework. https://doi.org/10.17605/OSF.IO/6JWE2

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

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. (2023). 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), 1329–1347. https://doi.org/10.1037/xhp0000942

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

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, *84*(4), 822–848. https://doi.org/10.1037/0022-3514.84.4.822

Bugg, J. M. (2014). Conflict-triggered top-down control: Default mode, last resort, or no such thing? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 567–587. https://doi.org/10.1037/a0035032

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*, Article 367. 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*(8), 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), 1029–1050. https://doi.org/10.1037/xhp0000801

Carriere, J. S., Seli, P., & Smilek, D. (2013). Wandering in both mind and body: Individual differences in mind wandering and inattention predict fidgeting. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *67*(1), 19–31. https://doi.org/10.1037/a0031438

Crump, M. J. C., Brosowsky, N. P., & Milliken, B. (2017). Reproducing the location-based context-specific proportion congruent effect for frequency unbiased items: A reply to Hutcheon and Spieler (2016). *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *70*(9), 1792–1807. https://doi.org/10.1080/17470218.2016.1206130

Crump, M. J. C., Lai, W., & Brosowsky, N. P. (2019). Instance theory predicts information theory: Episodic uncertainty as a determinant of keystroke dynamics. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *73*(4), 203–215. https://doi.org/10.1037/cep0000182

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dishon-Berkovits, M., & Algom, D. (2000). The Stroop effect: It is not the robust phenomenon that you have thought it to be. *Memory & Cognition*, *28*(8), 1437–1449. https://doi.org/10.3758/BF03211844

Dreisbach, G. (2012). Mechanisms of cognitive control: The functional role of task rules. *Current Directions in Psychological Science*, *21*(4), 227–231. https://doi.org/10.1177/0963721412449830

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective & Behavioral Neuroscience*, *7*(4), 380–390. https://doi.org/10.3758/CABN.7.4.380

Goh, W. D., Yap, M. J., & Chee, Q. W. (2020). The Auditory English Lexicon Project: A multi-talker, 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

Goschke, T. (2013). Volition in action: Intentions, control dilemmas and the dynamic regulation of intentional control. In W. Prinz, A. Beisert, & A. Herwig (Eds.), *Action science: Foundations of an emerging discipline* (pp. 409–434). MIT Press. https://doi.org/10.7551/mitpress/9780262018555.003.0024

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*(4), 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*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Hommel, B. (2015). Between persistence and flexibility: The Yin and Yang of action control. In A. J. Elliot (Ed.), *Advances in motivation science* (Vol. 2, pp. 33–67). Elsevier.

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

Kim, S., & Cho, Y. S. (2014). Congruency sequence effect without feature integration and contingency learning. *Acta Psychologica*, *149*, 60–68. https://doi.org/10.1016/j.actpsy.2014.03.004

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

Le Pelley, M. E., Vadillo, M., & Luque, D. (2013). Learned predictiveness influences rapid attentional capture: Evidence from the dot probe task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6), 1888–1900. https://doi.org/10.1037/a0033700

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

Liesefeld, H. R., Fu, X., & Zimmer, H. D. (2015). Fast and careless or careful and slow? Apparent holistic processing in mental rotation is explained by speed–accuracy trade-offs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1140–1151. https://doi.org/10.1037/xlm0000081

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

Lindsay, D. S., & Jacoby, L. L. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 219–234. https://doi.org/10.1037/0096-1523.20.2.219

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*, Article 1171. https://doi.org/10.3389/fpsyg.2015.01171

Logan, G. D. (2018). Automatic control: How experts act without thinking. *Psychological Review*, *125*(4), 453–485. https://doi.org/10.1037/rev0000100

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*(3), 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), 978–992. https://doi.org/10.1037/0096-1523.24.3.978

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), 684–700. https://doi.org/10.1037/h0080661

Mann, Y. (2015). *Interactive music with tone.js* [Conference session]. Proceedings of the International Web Audio Conference (WAC '15), Paris, France. http://webaudioconf.com/_data/papers/pdf/2015/2015_40.pdf

Mayr, U., Awh, E., & Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature Neuroscience*, *6*(5), 450–452. https://doi.org/10.1038/nn1051

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, *110*(3), 422–471. https://doi.org/10.1037/0033-295X.110.3.422

Notebaert, W., & Verguts, T. (2007). Dissociating conflict adaptation from feature integration: A multiple regression approach. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1256–1260. https://doi.org/10.1037/0096-1523.33.5.1256

Parsons, S. (2021). splithalf: Robust estimates of split half reliability. *Journal of Open Source Software*, *6*(60), Article 3041. https://doi.org/10.21105/joss.03041

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

Roelofs, A. (2005). The visual-auditory color-word Stroop asymmetry and its time course. *Memory & Cognition*, *33*(8), 1325–1336. https://doi.org/10.3758/BF03193365

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin & Review*, *26*(2), 452–467. https://doi.org/10.3758/s13423-018-1558-y

Schmidt, J. R. (2019). Evidence against conflict monitoring and adaptation: An updated review. *Psychonomic Bulletin & Review*, *26*(3), 753–771. https://doi.org/10.3758/s13423-018-1520-z

Schmidt, J. R. (2021). When data transformations are appropriate or even necessary: A response to Cohen-Shikora, Suh and Bugg (2019). *Timing & Time Perception*, *9*(2), 161–197. https://doi.org/10.1163/22134468-bja10019

Schmidt, J. R., & Besner, D. (2008). The Stroop effect: Why proportion congruent has nothing to do with congruency and everything to do with contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 514–523. https://doi.org/10.1037/0278-7393.34.3.514

Schmidt, J. R., & De Houwer, J. (2011). Now you see it, now you don't: Controlling for contingencies and stimulus repetitions eliminates the Gratton effect. *Acta Psychologica*, *138*(1), 176–186. https://doi.org/10.1016/j.actpsy.2011.06.002

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

Sharma, D., & McKenna, F. P. (1998). Differential components of the manual and vocal Stroop tasks. *Memory & Cognition*, *26*(5), 1033–1040. https://doi.org/10.3758/BF03201181

Snijder, J.-P., Tang, R., Bugg, J. M., Conway, A. R. A., & Braver, T. S. (2023). On the psychometric evaluation of cognitive control tasks: An investigation with the Dual Mechanisms of Cognitive Control (DMCC) battery. *Behavior Research Methods*. Advance online publication. https://doi.org/10.3758/s13428-023-02111-7

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, *59*(1), 1–22. https://doi.org/10.3758/BF03206843

Spinelli, G., & Lupker, S. J. (2021). Proactive control in the Stroop task: A conflict-frequency manipulation free of item-specific, contingency-learning, and color-word correlation confounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(10), 1550–1562. https://doi.org/10.1037/xlm0000820

Spinelli, G., Perry, J. R., & Lupker, S. J. (2019). Adaptation to conflict frequency without contingency and temporal learning: Evidence from the picture-word interference task. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(8), 995–1014. https://doi.org/10.1037/xhp0000656

Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). Oxford University Press.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Tang, R., Bugg, J. M., Snijder, J.-P., Conway, A. R., & Braver, T. S. (2023). The Dual Mechanisms of Cognitive Control (DMCC) project: Validation of an online behavioural task battery. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *76*(7), 1457–1480. https://doi.org/10.1177/17470218221114769

Tomat, M., Wendt, M., Luna-Rodriguez, A., & Jacobsen, T. (2021). Adjustments of selective attention to response conflict—Controlling for perceptual conflict, target–distractor identity, and congruency level sequence pertaining to the congruency sequence effect. *Attention, Perception & Psychophysics*, *83*(6), 2531–2550. https://doi.org/10.3758/s13414-021-02294-1

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

Viviani, G., Visalli, A., Montefinese, M., Vallesi, A., & Ambrosini, E. (2022). *The Stroop legacy: A cautionary tale on methodological issues and a proposed spatial solution*. PsyArXiv. https://doi.org/10.31234/osf.io/qejum

Weissman, D. H., Jiang, J., & Egner, T. (2014). Determinants of congruency sequence effects without learning and memory confounds. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(5), 2022–2037. https://doi.org/10.1037/a0037454

White, B. W. (1969). Interference in identifying attributes and attribute names. *Perception & Psychophysics*, *6*(3), 166–168. https://doi.org/10.3758/BF03210086

Whitehead, P. S., Brewer, G. A., & Blais, C. (2019). Are cognitive control processes reliable? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(5), 765–778. https://doi.org/10.1037/xlm0000632

Received June 6, 2023
Revision received November 1, 2023
Accepted December 15, 2023 ■