Working memory capacity and Stroop interference: Global versus local indices of executive control

By: Matt E. Meier, Michael J. Kane

Abstract:

Two experiments examined the relations among working memory capacity (WMC), congruency-sequence effects, proportion-congruency effects, and the color-word Stroop effect to test whether congruency-sequence effects might inform theoretical claims regarding WMC's prediction of Stroop interference. In Experiment 1, subjects completed either a high-congruency or low-congruency Stroop task that restricted trial-to-trial repetitions of stimulus dimensions to examine WMC's relation to congruency-sequence effects while minimizing bottom-up, stimulus-driven contributions. Congruency-sequence effects and congruency-proportion effects were significant but did not interact. WMC predicted global Stroop interference under low-congruency conditions but neither local congruency-sequence effects nor global Stroop interference under high-congruency conditions, contrary to previous studies (e.g., Kane & Engle, 2003). A high-congruency Stroop task in Experiment 2 removed the Experiment 1 task constraints, and, here, we obtained the typical, global association between WMC and Stroop interference but still no relation between WMC and congruency-sequence effects. We thus examined the methodological differences between Experiments 1 and 2 to determine whether any of these were locally responsible for the global WMC-related differences. They were not, suggesting that the changes between Experiments 1 and 2 created a general task context that engaged (or disengaged) the executive processes associated with WMC.

Keywords: Stroop | congruency sequence | executive control | individual differences | working memory | color-word Stroop effect | capacity

Article:

Working memory capacity (WMC), as measured by complex span tasks, predicts success in complex intellectual activities such as learning computer languages (Shute, 1991), multitasking (Hambrick, Oswald, Darowski, Rench, & Brou, 2010), and solving novel problems (Kane et al.,

2004; Kyllonen & Christal, 1990). Many prominent theories suggest a causal role for attention in WMC's relation with higher-order cognition (for reviews, see Conway, Jarrold, Kane, Miyake, & Towse, 2007). Evidence for this claim comes, in part, from demonstrations that WMC variation predicts performance on relatively "simple" attention tasks, such as the color-word Stroop (1935) task. Here, people must name the color in which a word is displayed while ignoring the word's identity. Interference—in the form of slowed or erroneous responding—is produced on incongruent trials, where the word conflicts with the color (e.g., *RED* in blue), compared to either congruent trials, where the word matches the color (e.g., *BLUE* in blue), or neutral trials, where the word is unrelated to the color (e.g., *POOR* in blue or *TXJF* in blue).

Not only do subjects with higher WMC experience less color-word interference than do those with lower WMC (Hutchison, 2011; Long & Prat, 2002; Unsworth & Spillers, 2010), but these WMC effects are moderated by the extent to which the task context challenges goal-maintenance capabilities (Kane & Engle, 2003): Lower-WMC subjects commit more errors than do higher-WMC subjects in unsupportive task contexts that put a premium on actively maintaining accessibility to the novel task goal against a strong but inappropriate habitual response. That is, in a Stroop task with mostly congruent trials, subjects can respond correctly on most trials by either reading the word or naming the color, and so the novel color-naming goal is infrequently demanded (or reinforced) by the context. In this permissive context, subjects may readily slip into the habitual, erroneous task set of word reading rather than color naming, and, indeed, lower WMC subjects appear to do so. In a task context with mostly incongruent trials, however, where the name-the-color goal is repeatedly reinforced by the presence of incongruent trials, lower-WMC subjects were better able to maintain the goal, reducing their error rate on incongruent trials in this goal-supportive context. Even with this contextual support, however, lower-WMC subjects still experienced more response competition than did higher WMC subjects, evidenced by their greater interference in reaction times (RTs). That is, even when lower-WMC subjects responded according to goals, thanks to the low-congruency context, they took more time to resolve the interference created by each incongruent stimulus.

WMC thus predicts aggregate performance in Stroop tasks, globally (e.g., Hutchison, 2011; Kane & Engle, 2003). Here, we examined whether individual differences at a local, trial-by-trial level might provide further insights about *how* WMC relates to Stroop performance, perhaps allowing a more specific mechanistic account of previously observed global differences. That is, although we know WMC reflects cognitive control in some situations, both WMC and cognitive control are broad constructs that need to be refined to enhance our understanding of when, where, and how they relate to one another (Kane, Poole, Tuholski, & Engle, 2006). It is possible that the previously observed global effects of WMC on Stroop task performance were the cumulative product of WMC and trial-level performance relations. One way to examine trial-level performance is via congruency transitions between trials (e.g., Gratton, Coles, & Donchin, 1992), referred to here as *congruency-sequence effects* (for a review, see Egner, 2007).

There are two dominant accounts of congruency-sequence effects: conflict-monitoring theory (Botvinick, Braver, Barch, Carter, & Cohen, 2001) and feature-integration theory (Hommel, Proctor, & Vu, 2004; Mayr, Awh, & Laurey, 2003). The conflict-monitoring account provides a

top-down explanation, where conflict (such as Stroop conflict) is detected by the anterior cingulate cortex, which signals the prefrontal cortex to bias attention toward the stimulus features required for successful performance. Accordingly, congruent trials immediately following congruent trials (cC) are the fastest of all sequences because controlled responding is unnecessary on either the preceding or current trial; incongruent trials following congruent trials (cI) are slowest because control is not deployed on the preceding congruent trial and so engaging it on the current conflict trial takes time. Congruent trials preceded by incongruent trials (iC) are slower than cC trials because control processes remain activated from the conflict on the previous trial, and incongruent trials preceded by incongruent trials (iI) are faster than cI trials, also due to carryover of control (beneficial here, due to current incongruent conflict). Congruency-sequence effects, then, may be viewed as dynamic markers of cognitive-control implementation. As such, the control processes that give rise to congruency-sequence effects may also mediate the relation between WMC and the performance of cognitive conflict tasks.

Alternatively, the feature-integration account (Hommel et al., 2004; Mayr et al., 2003), a bottom-up explanation, attributes congruency-sequence effects to the repetitions (and non-repetitions) of stimulus features across trials. That is, responding is faster and more accurate when the complete stimulus ensemble (e.g., both the Stroop color and word) repeats exactly, or changes completely, across trials. Responding is slowed and less accurate, in contrast, for ensembles with partial repetitions because the features that carry over from the prior trial are associated with that prior trial's response. Before the response on the current trial can be executed, interference from prior associations must be overcome. By definition, cC and iI trials present either complete repetitions or alternations, whereas cI and iC trials often contain partial repetitions of stimulus features from the prior trial. Thus, cI and iC responses will be slowed. Recent research has suggested that the feature-integration and conflict-monitoring accounts are not mutually exclusive and that both types of mechanism may play a role in determining congruency-sequence effects (at least in tasks that allow stimulus repetition, and thus feature integration, to occur; Notebaert, Gevers, Verbruggen, & Liefooghe, 2006).

Because congruency-sequence effects are considered markers of cognitive control, at least under some empirical conditions, and because WMC predicts cognitive control abilities, at least in some tasks argued to reflect control, researchers have recently sought associations between WMC and congruency-sequence effects. To the extent that WMC predicts only some indicators of executive control and not others (e.g., Kane et al., 2006), it may help the field specify the broad (and ill-defined) "control" construct (Kane, Conway, Hambrick, & Engle, 2007). Keye, Wilhelm, Oberauer, and van Ravenzwaaij (2009) examined WMC-related individual differences in congruency-sequence effects in flanker (Eriksen & Eriksen, 1974) and Simon tasks (Simon & Small, 1969). They found that higher-WMC subjects experienced smaller RT congruency-sequence effects than did lower-WMC subjects in the Simon task, but that WMC did not relate to congruency-sequence effects in a flanker task. Keye et al. also did not find WMC-related differences in the aggregate amount of Simon or flanker interference experienced. Perhaps WMC's influence—if any—on congruency-sequence effects will be consistent only in tasks that are generally more sensitive to WMC variation.

Here, then, we used the Stroop task to investigate congruency-sequence effects because it has frequently produced WMC-related individual differences (Hutchison, 2011; Kane & Engle, 2003; Long & Prat, 2002; Unsworth & Spillers, 2010). In work published after the studies reported here were completed, however, Unsworth, Redick, Spillers, and Brewer (2012) found no relation between WMC and congruency-sequence effects in a color-word Stroop task with a manual response. We suggest that Unsworth et al.'s findings are limited by their examining WMC and congruency-sequence effects in tasks allowing (and on trials presenting) stimulus repetitions. The bottom-up, feature-integration contributions of repetition trials may have thus masked a relation between WMC and the top-down driven components of congruency-sequence effects. WMC is thought to affect performance in a top-down manner, but congruency-sequence effects seem to be dually determined. Because of this, we cannot draw strong conclusions about the relation between WMC and top-down control processes from Unsworth et al.'s data. Together, then, the evidence from Keye et al. (2009) and Unsworth et al. presents an unclear picture of WMC and congruency-sequence-effect relations. We intended for the present study to provide some clarification.

To the extent that WMC affects trial-to-trial congruency transitions, we should see its influence in one of two ways. On one hand, higher-WMC subjects might show *larger* congruency-sequence effects than will lower-WMC subjects. This pattern would hold if higher WMC leads either to better conflict monitoring or to more effective control-process implementation following from that monitoring. On the other hand, subjects with higher WMC might produce *smaller* congruency-sequence effects than will lower-WMC subjects because less conflict is created when the task goal is better maintained. In other words, if higher-WMC subjects keep the color-naming goal more accessible than do lower-WMC subjects, preemptively biasing attention only to task-relevant stimulus dimensions, then they will experience less conflict to react to. Indeed, Keye et al. (2009) have provided some support for this scenario with their finding of a positive correlation between interference effects and congruency-sequence effects.

Our predictions for congruency-sequence effects must also consider, however, that WMC-related differences in Stroop interference vary with overall proportions of congruency (and thus of conflict; Kane & Engle, 2003). Previous work by Stürmer, Leuthold, Soetens, Schröter, and Sommer (2002) with a Simon task, and by Purmann, Badde, and Wendt (2009) with a flanker task, indicates that list-wide proportion-congruency also moderates congruency-sequence effects. Specifically, effects were smaller in task contexts with lower proportions of congruent trials. Given that lower WMC subjects show greater Stroop interference in tasks with fewer incongruent trials, we might expect WMC differences in congruency-sequence effects to be larger in a task context with few incongruent trials. Indeed, after the present experiments were completed, Hutchison (2011) reported a three-way interaction among WMC, congruency-sequence, and proportion-congruency effects in a Stroop task: Congruency sequence effects were, in fact, smaller under low than high congruency, but WMC moderated the size and direction of the effects. Higher-WMC subjects showed equivalent Stroop interference regardless of congruency sequence and proportion congruency. Lower-WMC subjects, however, showed the typical interference reduction following incongruent trials versus congruent trials in a low-

congruency context, but they showed the opposite pattern—an increase in interference—following incongruent trials in a high-congruency context. Hutchison's unexpected high-congruency results—and his failure to find an overall congruency-sequence effect in RT interference—may have been an artifact of using an unusual Stroop method for other theoretical purposes (i.e., manipulating the congruency of particular color-word pairs as well as the congruency of the overall stimulus list). It thus remains important to test, with a more typical Stroop preparation, whether WMC relates differentially to congruency-sequence effects under different proportion-congruency manipulations.

In our first experiment, we tested whether the previously found global WMC-related individual differences in Stroop performance are related to congruency-sequence effects. Some subjects completed a high-congruency condition (70% congruent trials), whereas others completed a low-congruency condition (30% congruent trials). We manipulated proportion-congruency to test whether congruency-sequence effects were affected by list-wide congruency manipulations in a Stroop task, and whether these effects mirror those of WMC and proportion-congruency (Kane & Engle, 2003). If congruency-sequence effects explain global WMC-related performance differences, then proportion-congruency effects should vary systematically with both congruency-sequence effects and WMC. Because we theorize that WMC exerts its influence in a top-down manner, we wanted to examine congruency-sequence effects in a task context where bottom-up influences from the sequences of trials were minimized. To this end, we used a stimulus set of six colors and color-words and did not allow repetitions (partial and complete) to occur from trial to trial.

Experiment 1

Method

Subjects

One hundred and ninety-seven undergraduate students (18–35 years of age) from the University of North Carolina at Greensboro (UNCG) participated and received partial credit for a course requirement.

General procedure

Subjects completed three testing sessions with an average of 22 days ($SD = 16$) between the first and last. We individually administered the Operation Span task (OSPAN) in the first session followed by the Stroop task, and the Symmetry Span task (SSPAN) in the second session. In the third session, we administered the Reading Span task (RSPAN) in groups of up to six subjects. We administered all WMC span tasks as the first in their sessions; each session lasted 1.5 hr and presented other tasks not discussed here. Experimenters read all on-screen instructions aloud while subjects read along silently. Dell desktop computers with 17-in. (43.18-cm) liquid-crystal display (LCD) or cathode ray tube (CRT) monitors presented all task stimuli and recorded all responses, controlled by E-Prime Version 1.2 software.

We tested subjects over one full academic semester, using the end of the semester as our stopping rule for data collection (Simmons, Nelson, & Simonsohn, 2011).

WMC screening

We assessed WMC with complex span tasks (OSPAN, SSPAN, and RSPAN). Subjects were required to maintain access to memory items while completing an unrelated processing task with an individualized response deadline ($M + 2.5$ SDs) that was calculated during 15 processing-task-only items (Unsworth, Heitz, Schrock, & Engle, 2005). In OSPAN, subjects verified solutions to compound arithmetic equations [e.g., $(3 \times 2) - 1 = 4$]. In SSPAN, subjects verified the vertical symmetry of black-and-white matrix patterns. In RSPAN, subjects verified the sensibility of sentence contents (e.g., "I like to run in the sky").

The memory items in OSPAN and RSPAN were capital letters (randomly selected without replacement on each trial from 12: *F, H, J, K, L, N, P, Q, R, S, T, Y*) that each appeared for 250 ms, following 200 ms after either operation/reading verification or response deadline. After 3–7 verification-letter pairs, all 12 letters appeared onscreen, and subjects identified, via mouse click, the letters presented in serial order. In SSPAN, one square of a 4×4 grid was shaded red for 650 ms, following 200 ms after either symmetry verification or response deadline. After 2–5 verification-grid pairs, subjects recalled the locations of the shaded squares in serial order by mouse-clicking on a single, empty grid. The tasks presented each set length (3–7 in OSPAN and RSPAN; 2–5 in SSPAN) three times, randomly ordered for each subject, yielding 75 verification-letter trials in OSPAN and RSPAN, and 42 verification-grid trials in SSPAN. In all span tasks, subjects practiced with three sets of list length two.

The score for each task was the sum of items correctly recalled in serial position (Conway et al., 2005). If a subject did not achieve 85% accuracy on the processing portion of the span task, that span task was dropped from subsequent analyses. For subjects who had at least two scores after applying the 85% processing-accuracy criterion, we converted their scores to Z scores (using the means and standard deviations from our database of more than 2,500 UNCG students) and averaged them into a WMC composite. Data from subjects who did not have at least two span tasks meeting this criterion were dropped altogether. Scores correlated with rs of .60 (RSPAN $\times$ OSPAN; $N = 160$), .42 (OSPAN $\times$ SSPAN; $N = 151$), and .39 (SSPAN $\times$ RSPAN; $N = 162$); Ns varied due to individually dropped tasks. The WMC composite was unimodal and symmetrically distributed (skew = –0.47; kurtosis = –0.39).

Stroop task

The Stroop task presented six colors and words (*RED*, *GREEN*, *BLUE*, *GRAY*, *BLACK*, or *PURPLE* in red, green, blue, gray, black, or purple). We manipulated proportion-congruency between subjects (70% vs. 30% congruent) and current-trial congruency and previous-trial congruency within subjects. Subjects were randomly assigned to proportion-congruency conditions. Four unanalyzed buffer trials preceded each trial block to establish the task set. In the 70% congruency condition, these buffer trials were all congruent; in the 30% condition, they were all incongruent (Kerns et al., 2004).

We randomized the order of trials in every block for every subject, with the following constraints. Neither the color nor the word repeated on consecutive trials as either the target (color) or distracter (word). That is, if the word, *RED* was displayed in blue on one trial, the

subsequent trial could not present the words *RED* or *BLUE*, or the colors red or blue. We further constrained trial presentations to prevent clumping trial types and thus the creation of congruency sub-contexts (e.g., a long series of incongruent trials in the 70% congruency condition). The majority trial type for the condition (i.e., congruent in 70% congruency, incongruent in 30% congruency) appeared on a maximum of seven consecutive trials and the minority trial type on two. In the 70% condition, we allowed no more than 15 incongruent trials (out of 30 in the block) to occur within the first 50 trials, and we required that at least five incongruent trials occurred over the last 15 trials of a block; the same constraints held for congruent trials in the 30% condition.

Procedure

Each trial presented subjects with a color word in one of six colors against a white background. As in Kane and Engle (2003), the experimenter instructed subjects that the task goal was to name the color of the word, to not read the word, to go as fast as possible without making mistakes, to slow down if making mistakes, and that their performance on incongruent trials was more important than that on congruent trials (in order to establish the color-naming goal, even in a high-congruency context). Subjects first completed 12 trials of microphone practice, reading aloud color words presented in black letters; they then completed 18 trials of color-naming practice, with rows of question marks appearing in color. To assess color-blindness, experimenters monitored subjects during color-naming practice for signs of difficulty (if detected, experimenters asked subjects about their color vision).

Subjects completed three blocks of 104 trials, with short breaks (<60 s) permitted between blocks. The experimenter read aloud brief on-screen instruction reminders before the second and third blocks. Each trial began with a 3,000-ms blank interval, followed by an olive-colored fixation cross for 250 ms that was replaced by the color-word stimuli, which remained on-screen until response. A voice-activated relay (Psychological Software Tools serial response box, Model 2.0) recorded all response latencies, and the experimenter recorded each response by pressing a corresponding color-coded key on the computer keyboard during the 3,000-ms blank screen. When subjects stuttered (e.g., "b-blue"; less than 1% of trials), it was recorded as an error, and the RT was not used. Experimenters also coded any microphone errors (e.g., accidental tripping of the voice-activated relay or the relay failing to detect a subject's response; 4% of trials) so that these data would not be analyzed.

Data analyses

We computed inferential statistics for RTs and error rates with linear mixed models (LMMs) using the lme4 package (Bates, Maechler, & Bolker, 2011) in the R system for statistical computing (R Development Core Team, 2011). LMMs were used because they offered us three major advantages over analyses of variance (ANOVAs; Kliegl, Wei, Dambacher, Yan, & Zhou, 2011): (1) LMMs handle the unbalanced data produced by our proportion-congruency manipulations without a loss of power (Baayen, Davidson, & Bates, 2008; Pinheiro & Bates, 2000; Quené & van den Bergh, 2008), (2) LMMs allowed us to treat WMC as a continuous variable in the context of a repeated-measures experimental design, and (3) LMMs control for

overall RT and error-rate differences between subjects as well as for nuisance variance created by specific effects of particular color-words and colors. For RTs, we interpreted estimates with t values greater than 2 as significant; with a large number of observations, the t distribution converges on the standard normal distribution, where absolute values greater than 2 are less than the traditional alpha of .05. We computed inferential statistics for error rates using generalized linear mixed models (GLMMs), also in the lme4 package (Bates et al., 2011). GLMMs offer all of the advantages covered above for LMMs but also enabled us to model error rates as binomially distributed (Dixon, 2008). The betas for the GLMMs are reported in terms of log odds ratios, with higher log odds ratios meaning that an error is more likely to be committed. Unless otherwise noted, the fixed effects of current-trial congruency, previous-trial congruency, and proportion-congruency were all coded as –.5/+.5 contrasts, so the parameter estimates for these experimental manipulations represent the experimental effect. WMC was centered and entered as a continuous variable. We specified subjects, colors, and words as uncorrelated random effects. For RTs, we did not analyze data from the first trial of every block, from incorrect trials, or from trials with latencies shorter than 200 ms or longer than 1,500 ms, resulting in a loss of less than 1% of trials in both Experiment 1 and Experiment 2.

Results

We dropped all data from eight subjects (six who did not meet the processing-accuracy criterion for the WMC composite, one for vision problems, and one with 0% incongruent-trial accuracy in Block 1 of the Stroop task). This left data from 189 subjects: 96 in the 70% congruency condition and 93 in the 30% condition. The WMC means for the low-congruency (0.04, $SD = 0.75$) and high-congruency (0.08, $SD = 0.72$) conditions did not differ, $t(187) = 0.40$, $p = .69$. Means and standard errors for RTs and error rate proportions are displayed in Tables 1 and 2.

Table 1. *Means (and Standard Errors) of Subject Means of Reaction Times by Experiment, Condition, Previous-Trial Congruency, and Current-Trial Congruency*

	Experiment 1				Experiment 2			
	30% congruency (N = 93)		70% congruency (N = 96)		3 color (N = 130)		6 color (N = 121)	
Congruency type	Current congruent	Current incongruent	Current congruent	Current incongruent	Current congruent	Current incongruent	Current congruent	Current incongruent
Previous congruent	688 (11)	792 (10)	655 (9)	812 (10)	640 (7)	814 (9)	660 (8)	831 (10)
Previous incongruent	709 (11)	781 (10)	675 (9)	806 (10)	677 (7)	784 (9)	687 (9)	808 (10)

Table 2. *Means (and Standard Errors) of Subject Means of the Proportion of Errors by Experiment, Condition, Previous-Trial Congruency, and Current-Trial Congruency*

	Experiment 1				Experiment 2			
	30% congruency (N = 93)		70% congruency (N = 96)		3 color (N = 130)		6 color (N = 121)	
Congruency type	Current congruent	Current incongruent	Current congruent	Current incongruent	Current congruent	Current incongruent	Current congruent	Current incongruent
Previous congruent	.004 (.001)	.020 (.002)	.004 (.001)	.036 (.004)	.005 (.001)	.048 (.004)	.004 (.001)	.047 (.005)
Previous incongruent	003 (.001)	.016 (.002)	.003 (.001)	.025 (.004)	.005 (.001)	.023 (.003)	.004 (.001)	.025 (.003)

Response times

Global Stroop interference effects

Subjects showed significant Stroop interference overall, with congruent trials yielding faster responses than incongruent trials ($b = 116$ ms, $SE = 1$ ms, $t = 90.1$). As expected, and as can be seen in Figure 1, Stroop interference was reduced in the 30% congruency versus the 70% congruency condition ($b = -55$ ms, $SE = 3$ ms, $t = -21.3$). Overall, the amount of Stroop interference did not differ by subjects' WMC ($b = -3$ ms, $SE = 2$ ms, $t = -1.6$), but the three-way interaction among WMC, proportion-congruency, and Stroop interference was significant, ($b = -9$ ms, $SE = 4$ ms, $t = -2.4$). We conducted follow-up LMMs, one for each congruency condition, to determine the locus of this three-way interaction. In the 30% congruency condition— consistent with Kane and Engle (2003)—Stroop interference was reduced for subjects with higher WMC composites ($b = -7$ ms, $SE = 3$ ms, $t = -2.6$). However, inconsistent with the findings of Kane and Engle, in the 70% congruency condition, Stroop interference was not reduced for subjects with higher WMC ($b = 1$ ms, $SE = 2$ ms, $t = 0.6$).
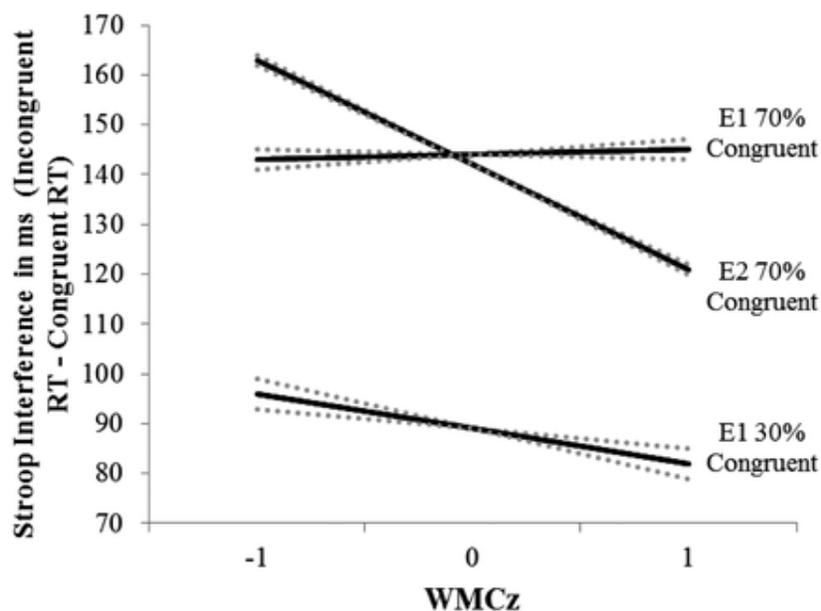
Figure 1. Experiment 1 (E1) and Experiment 2 (E2) model parameters of Stroop interference as a function of working memory capacity (WMCz) by proportion-congruency condition. The dotted lines represent the standard error of the estimate. The line for E2 is collapsed over stimulus-set-size conditions. RT = reaction time.

Local congruency-sequence effects

Congruency-sequence effects are indicated by an interaction between Stroop interference (current-trial congruency) and previous-trial congruency. Figure 2 displays Stroop interference following congruent and incongruent trials as a function of the overall proportion-congruency condition. There was a main effect of previous-trial congruency ($b = 6$ ms, $SE = 1$ ms, $t = 4.7$), but more importantly, as shown in Figure 2, we found reliable congruency-sequence effects, with interference reduced following an incongruent trial versus a congruent trial ($b = -29$ ms, $SE = 3$ ms, $t = -11.3$). However, congruency-sequence effects did not interact with WMC ($b = 1$ ms, $SE = 4$ ms, $t = 0.2$) or proportion-congruency ($b = -6$ ms, $SE = 5$ ms, $t = -1.2$).
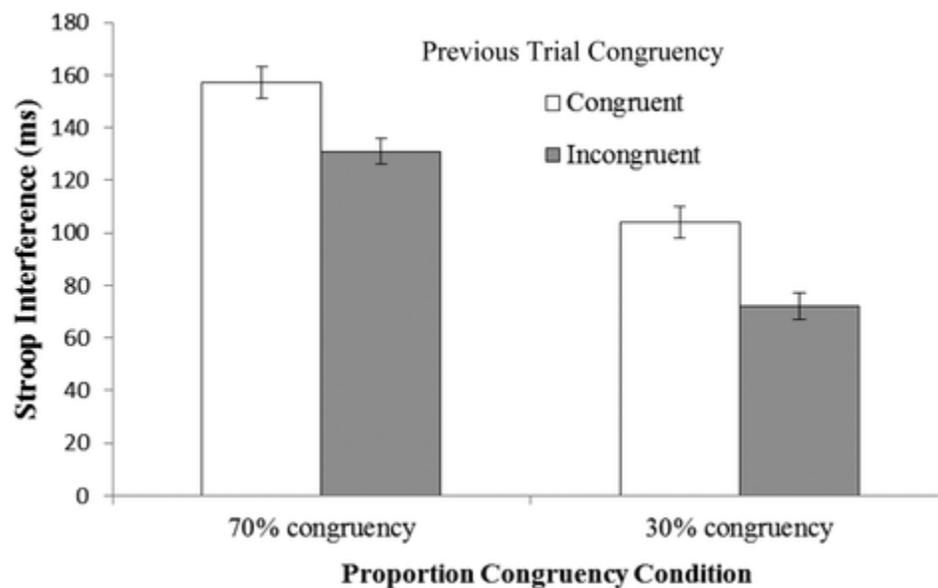

Figure 2. Experiment 1 observed means (of subjects' means) for Stroop interference (incongruent-trial reaction time minus congruent-trial reaction time) as a function of previous trial congruency by proportion-congruency condition. Error bars represent 1 SEM.

Errors

Global Stroop interference effects

Subjects showed Stroop interference, with more errors committed on incongruent than congruent trials ($b = 2.08$, $SE = 0.16$, $Z = 13.0$), but interference did not significantly differ by proportion-congruency ($b = -0.17$, $SE = 0.32$, $Z = 0.60$). Inconsistent with the findings from Kane and Engle (2003), WMC was not associated with interference magnitude ($b = -0.04$, $SE = 0.22$, $Z = -0.20$) or with the interaction of interference with proportion-congruency ($b = -0.15$, $SE = 0.44$, $Z = -0.34$).

Local congruency-sequence effects

Error interference was reduced following incongruent compared to congruent trials ($b = -0.68$, $SE = 0.32$, $Z = -2.14$). These congruency-sequence effects did not interact with WMC ($b = -0.28$, $SE = 0.44$, $Z = -0.64$) or proportion-congruency ($b = -0.84$, $SE = 0.64$, $Z = -1.32$).

Discussion

Subjects exhibited Stroop interference in RTs and errors in both the 70% and 30% congruency conditions, with weaker interference in the 30% condition. Regarding WMC-related individual differences, we replicated Kane and Engle (2003) in the low-congruency—but not the high-congruency—condition. In the 30% congruency task, subjects with higher WMC experienced slightly (but significantly) less Stroop RT interference than did subjects with lower WMC, with no WMC-related effects on errors. Although we expected that subjects with higher WMC would experience substantially less Stroop interference in either RTs or errors in the 70% congruent condition, they did not. We thus designed Experiment 2 to explore why the 70% congruency condition did not replicate the results of Kane and Engle (or Hutchison, 2011; Unsworth & Spillers, 2010) and to again test for the relation of WMC and congruency-sequence effects.

We found significant congruency-sequence effects in both RTs and errors, with subjects experiencing less Stroop interference in RT and errors after incongruent than after congruent trials. To our knowledge, this is the first report of congruency-sequence effects in a color-word Stroop task that did not permit immediate repetitions of colors or words (thus ruling out bottom-up priming effects based on feature-integration, and seeming to support the influence of top-down contributions). In contrast to findings from flanker and Simon tasks (Purmann et al., 2009; Stürmer et al., 2002), but consistent with Funes, Lupianez, and Humphrey's (2010) assertion that congruency-sequence effects and proportion congruent effects reflect independent types of control, congruency-sequence effects were not moderated by the proportion of congruent trials. That is, we found the control engaged by the overall task context to be different than the control engaged by a recent detection of conflict.

Perhaps unsurprisingly, in a task that did not produce robust WMC-related differences overall, we did not see any relation between WMC and congruency-sequence effects. If there is little WMC-related variation in performance, there is little to look at as its causes or effects. Recall that this study attempted to improve upon and add clarity to the previous work of Keye et al. (2009) and Unsworth et al. (2012) by examining the relation between WMC-and congruency-sequence effects in a task with a history of producing WMC-related differences, and with the feature-integration contributions to congruency-sequence effects removed. The constraints we placed on the task in Experiment 1 seem to have eliminated the expected WMC-related variation in performance.

Experiment 2

We were surprised that Experiment 1 failed to yield WMC-related differences in Stroop interference in a 70% congruency condition (cf. Kane & Engle, 2003; see also Hutchison, 2011; Unsworth & Spillers, 2010). With a sample size of 96 subjects, a lack of power was not a

likely culprit. We therefore suspected that any of three methodological differences may have accounted for these null findings: (1) the number of congruent trials allowed to occur consecutively, (2) the presence versus absence of stimulus-feature repetitions from trial-to-trial, and (3) the size of the stimulus-response set.

First, by limiting the number of consecutive congruent trials in Experiment 1, we may have eliminated a critical element of the task context for promoting goal neglect. Prior conceptions of goal neglect have characterized it as occurring when—for sustained periods—guidance from the task goal is not needed for correct responding (Duncan, Emslie, Williams, Johnson, & Freer, 1996), as in long streaks of congruent Stroop trials (Kane & Engle, 2003). Second, by preventing repetitions of stimulus features across trials, we may have reduced the influence of response competition, and such competition may be important to WMC-related variation in performance (Engle & Kane, 2003). That is, stimulus features of the previous trial are associated with the response of the previous trial; when these same features appear on the current trial, the prior stimulus-response association must be overcome if the current response differs. Third, and finally, the six color-words used in Experiment 1 represent a larger stimulus-response set than the three color-words used by Kane and Engle (2003). We thought that a smaller stimulus-response set may not only allow for more stimulus-feature repetitions (as discussed previously), but it may also lead to the task being more monotonous, thereby promoting mind wandering and lapses of goal maintenance (e.g., McVay & Kane, 2009).

In Experiment 2, then, we sought to reintroduce WMC-related performance differences in a 70% congruency Stroop task and, in that context, to test for WMC-related differences in congruency-sequence effects. To do so, we removed the constraints on repetitions of stimuli and trial types on consecutive trials, and we also manipulated the size of the stimulus-response set between subjects.

Method

Subjects

Two hundred and sixty-seven UNCG undergraduates (18–35 years of age), who had not participated in Experiment 1, participated in Experiment 2 for partial credit toward a course requirement.

General procedure

Subjects completed three testing sessions over the course of one semester, with an average of 21 days ($SD = 15$) between first and last sessions. We individually administered the SSPAN at the beginning of the first session followed by the Stroop task, the OSPAN to groups of 1–4 at the beginning of the second session, and the RSPAN in groups of up to 6 at the beginning of the third. All sessions lasted 1.5 hr and included other tasks not discussed here. We used the same computer hardware, software, and general procedures as in Experiment 1, and we used the same stopping rule (one full semester) for data collection.

WMC screening

We assessed WMC the same way as in Experiment 1. Scores correlated with rs of .57 (RSPAN $\times$ OSPAN; $N = 212$), .49 (OSPAN $\times$ SSPAN; $N = 203$), and .47 (SSPAN $\times$ RSPAN; $N = 218$). The WMC composite was unimodal and symmetrically distributed (skew $= -0.64$; kurtosis $= 0.02$).

Stroop task

We presented some subjects with three colors and color-words and others with six colors and color-words (the latter as in Experiment 1). The three-color condition used red, blue, and green (as in Kane & Engle, 2003). All subjects saw 70% congruent trials. As in Experiment 1, we manipulated current-trial and previous-trial congruency within-subjects.

We randomized the order of trials in every block for every subject and removed all of the Experiment 1 constraints, aside from proportion-congruency. Here, we allowed color and word repetitions across consecutive trials, and we allowed incongruent trials and congruent trials to occur as many times consecutively as the randomization of trial types produced. Subjects stuttered on fewer than 1% of trials, and microphone errors occurred on 3% of trials.

Results

We dropped the data from 16 subjects (11 who did not meet the WMC processing criterion, two who reported color-blindness, two who looked away from the monitor during Stroop, and one who did not initially understand Stroop instructions [89% error rate for Block 1 incongruent trials]). This left data from 251 subjects: 130 in the three-color condition and 121 in the six-color condition. The WMC means for the three-color (-0.04, $SD = 0.82$) and six-color (-0.08, $SD = 0.82$) conditions did not differ significantly, $t(249) = 0.42$, $p = .68$, and the overall mean WMC composite for Experiment 2 (-0.06, $SD = 0.82$) did not differ from the mean for Experiment 1 (0.06, $SD = 0.73$), $t(438) = 1.52$, $p = .13$. Means and standard errors for RTs and error rate proportions are displayed in Tables 1 and 2.

Response times

Global Stroop interference effects

Subjects experienced Stroop interference ($b = 142$ ms, $SE = 1$ ms, $t = 134.2$), which did not differ by stimulus set size ($b = 3$ ms, $SE = 2$ ms, $t = 1.2$). As seen in Figure 1, and now replicating Kane and Engle's (2003) findings, subjects with higher WMC scores exhibited less interference than did subjects with lower WMC scores ($b = -21$ ms, $SE = 1$ ms, $t = -15.7$). This WMC effect did not differ by stimulus set size ($b = 3$ ms, $SE = 3$ ms, $t = 1.3$).

Congruency-sequence effects

We again found significant congruency-sequence effects, with reduced interference following incongruent trials ($b = -59$ ms, $SE = 2$ ms, $t = -27.6$; see Figure 3). Congruency-sequence effects were also smaller in the six-color condition than in the three-color condition ($b = 15$ ms, $SE = 4$ ms, $t = 3.6$). Because the three-color condition contained more repetitions of stimulus elements from trial-to-trial than did the six-color condition, the greater congruency-sequence effects in this condition are suggestive of bottom-up processing contributions highlighted in the feature-integration account (Hommel et al., 2004; Mayr et al., 2003). Of most importance, congruency-

sequence effects did not vary with WMC ($b = 1$ ms, $SE = 3$ ms, $t = 0.5$), even in a task context producing WMC-related differences in global Stroop interference.
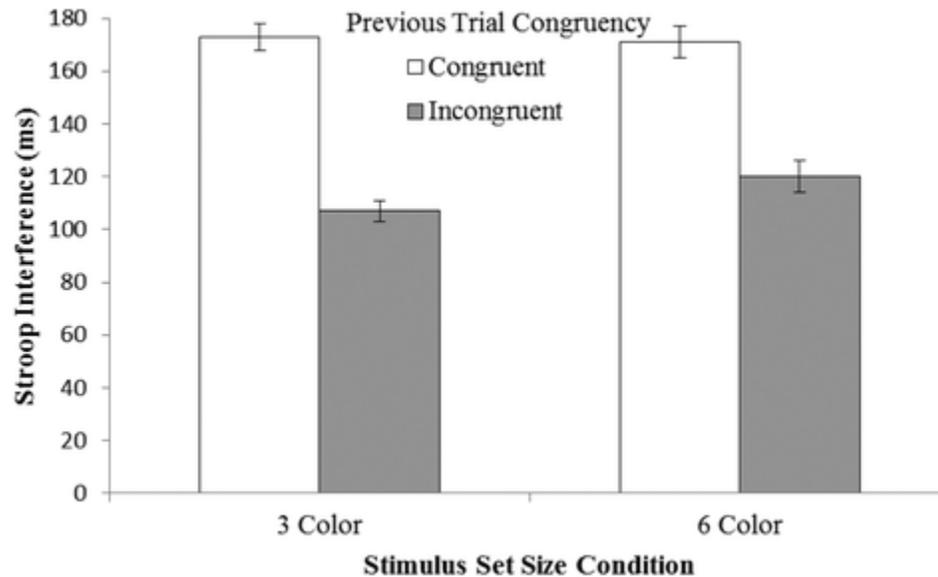

Figure 3. Experiment 2 observed means (of subjects' means) for Stroop interference (incongruent-trial reaction time minus congruent-trial reaction time) as a function of previous trial congruency by stimulus-set-size condition. 3 color = the three-color stimulus-set-size condition; 6 color = the six-color set-size condition. Error bars represent 1 SEM.

Errors

Global Stroop interference effects

Subjects showed interference here, too, with more errors committed on incongruent than congruent trials ($b = 2.06$, $SE = 0.09$, $Z = 23.1$). Also replicating Kane and Engle (2003), subjects with higher WMC showed less interference in errors than did subjects with lower WMC ($b = -0.52$, $SE = 0.11$, $Z = -4.66$). Error interference did not differ by stimulus set size ($b = 0.03$, $SE = 0.18$, $Z = 0.18$).

Local congruency-sequence effects

Stroop interference was reduced following incongruent compared to congruent trials ($b = -0.80$, $SE = 0.18$, $Z = -4.41$). Again, of central importance, congruency-sequence effects did not interact with WMC ($b = 0.09$, $SE = 0.22$, $Z = 0.43$). We found no effect of stimulus set size on the congruency-sequence effects ($b = -0.84$, $SE = 0.64$, $Z = -1.32$).

Additional analyses

We conducted additional analyses to determine whether any of the methodological changes from Experiment 1 to Experiment 2, in isolation and in the moment, were responsible for producing the WMC-related individual differences in Experiment 2. We also used these analyses to

examine the relation between WMC and congruency-sequence effects with the influence of bottom-up contributions from stimulus-repetition trials controlled.

Congruent trial runs

The 70% congruency condition in Experiment 1—where WMC did not predict interference—allowed a maximum of seven consecutive congruent trials. Different outcomes between Experiments 1 and 2 thus could have been driven by WMC-related differences in Stroop interference locally on trials that followed relatively long congruent runs. We assessed this possibility by examining the Stroop Interference × WMC interaction in Experiment 2, with a dichotomous predictor variable representing run length in the model. We took the following steps to create our run-count variable. Run count for an incongruent trial reflected the number of consecutive congruent trials occurring prior to that trial. For congruent trials, in addition to the preceding consecutive congruent trials, we counted the current trial as belonging to run count. We dropped trials not preceded by a congruent trial from this analysis. We then split the run count into a dichotomous variable, with runs of seven or less dummy coded as the reference level, and runs of greater than seven as the comparison level. We collapsed across stimulus-set-size conditions and entered run count, current-trial congruency, and WMC as fixed effects in a LMM. Subjects, words, and colors were entered as uncorrelated random effects.

RT interference (b = 57 ms, SE = 4 ms, t = 13.9) and error interference (b = 0.69, SE = 0.24, Z = 2.87) were greater after runs of longer than seven congruent trials than after congruent trial runs of less than seven trials. If these immediate effects of long congruent-trial runs accounted for the WMC-related RT and error differences between Experiments 1 and 2, then we should find either that the WMC × Interference interaction from Experiment 2 disappears with run count included in the model, or that the interaction is moderated by run count. Neither was the case for RTs, as interference was still moderated by WMC (b = –21 ms, SE = 3 ms, t = 8.3) but was not further moderated by run count (b = 1 ms, SE = 5 ms, t = 0.2). Nor was it the case for accuracy, with the WMC × Interference interaction still significant in the model (b = –0.55, SE = 0.15, Z = –3.64) and not differing by run count (b = 0.03, SE = 0.30, Z = 0.10). From this, we conclude that the local effects of experiencing long congruent-trial runs in Experiment 2 were not solely responsible for producing the global WMC-related differences we observed in Stroop interference.

Immediate repetitions of colors and words

To assess the role of local stimulus repetitions on interference and WMC effects, we examined only trials that contained no repetitions. In Experiment 2, where repetitions could occur, higher WMC subjects may have better handled the potentially disruptive influence of a color from a previous trial becoming the word in the current trial (or vice versa), or they may have experienced more facilitation from complete repetitions of stimuli from trial-to-trial than did subjects with lower WMC (Long & Prat, 2002). If the immediate effect of trials with repetitions was solely responsible for the WMC-related differences in interference in Experiment 2, then WMC differences should not persist in their absence. We thus removed from analysis all trials featuring complete repetitions (e.g., the second of two consecutive trials where *RED* appeared in

blue) and all trials reflecting partial repetitions of stimulus features from trial to trial (e.g., *RED* in the color blue following *GREEN* in red). The RT analysis removed means of 74 and 145 trials per subject in the six- and three-color conditions, respectively.

Regarding global interference effects, subjects still showed RT interference ($b = 137$ ms, $SE = 2$ ms, $t = 78.6$), and subjects with higher WMC still showed less interference than did subjects with lower WMC ($b = -22$ ms, $SE = 2$ ms, $t = -10.3$). There was an interaction between stimulus set size, WMC, and Stroop interference ($b = 9$ ms, $SE = 4$ ms, $t = 2.1$), with the WMC slope being less steep in the six-color condition than in the three-color condition; that is, WMC-related differences in interference were smaller with six than with three colors. However, as verified with separate LMMs for each stimulus-set-size condition, the slope for WMC was significantly negative in both the three-color condition ($b = -27$ ms, $SE = 4$ ms, $t = 7.3$) and the six-color condition ($b = -18$ ms, $SE = 2$ ms, $t = 7.8$).

The removal of repetitions also did not substantially affect analysis of errors, with Ms = 148 and 76 trials per person removed in the three- and six-color conditions, respectively (these means are greater than the means removed for the above RT analyses because error trials were not included in the RT analyses). Subjects showed significant Stroop interference, with more errors committed on incongruent trials than congruent trials ($b = 2.04$, $SE = 0.14$, $Z = 14.18$). Moreover, as in the original analysis, subjects with higher WMC showed less interference than did those with lower WMC ($b = -0.59$, $SE = 0.17$, $Z = -3.37$). Error interference did not significantly differ by stimulus set size ($b = 0.08$, $SE = 0.29$, $Z = 0.27$). From these analyses, we can conclude, as we did with congruent-trial runs, that any local effects caused by immediate stimulus repetitions did not drive the global WMC-related differences in Stroop interference that we observed between subjects with lower versus higher WMC scores.

Regarding local congruency-sequence effects, they also remained significant after removing stimulus repetitions ($b = -46$ ms, $SE = 3$ ms, $t = -13.4$). WMC still did not interact with congruency-sequence effects, here with the bottom-up contributions of repetition trials controlled ($b = -3$ ms, $SE = 4$ ms, $t = -0.7$). The significant interaction found between congruency-sequence effects and stimulus set size that was found when all of the trials (repetitions included) were analyzed above was now non-significant ($b = 3$ ms, $SE = 7$ ms, $t = 0.40$). It is not surprising that, when repetition trials were initially left in the analysis, we found greater congruency-sequence effects in the three-color condition than in the six-color condition. The three-color condition, with fewer stimuli than the six-color condition, necessarily produced more complete and partial repetitions of stimulus elements. As stated previously, complete repetitions facilitate responding, whereas partial repetitions slow responding. The combination of these two effects leads to larger congruency-sequence effects.

Cross-experiment comparisons

We conducted additional LMMs to examine the locus of the changes in WMC-related Stroop interference effects from the 70% congruency condition in Experiment 1 to Experiment 2 (collapsing Experiment 2 across stimulus-set-size conditions). That is, the between-experiment differences we saw in the association between WMC and interference may have come primarily

from changes in performance on congruent trials, incongruent trials, or both. For this analysis, we dummy coded experiment and current-trial congruency so that Experiment 1 congruent trials were the reference level. As in the previous analyses, we centered WMC as a continuous variable, and we treated subjects, words, and colors as uncorrelated random effects.

For congruent trials, the RT slope for WMC did not differ between Experiments ($b = –5$ ms, $SE = 14$ ms, $t = –0.4$), but the change in error slope was close to our significance threshold ($b = 0.45$, $SE = 0.23$, $Z = 1.97$), weakly indicating that the higher the subject's WMC, the more likely he or she was to make a congruent-trial error in Experiment 2 compared to Experiment 1. We found much more robust effects on incongruent trials, where the negative WMC slope was significantly steeper in Experiment 2 versus Experiment 1 (see Figure 4 for the RTs), in both RTs ($b = –22$ ms, $SE = 2$ ms, $t = –9.0$) and in errors ($b = –0.64$, $SE = 0.20$, $Z = –3.18$).
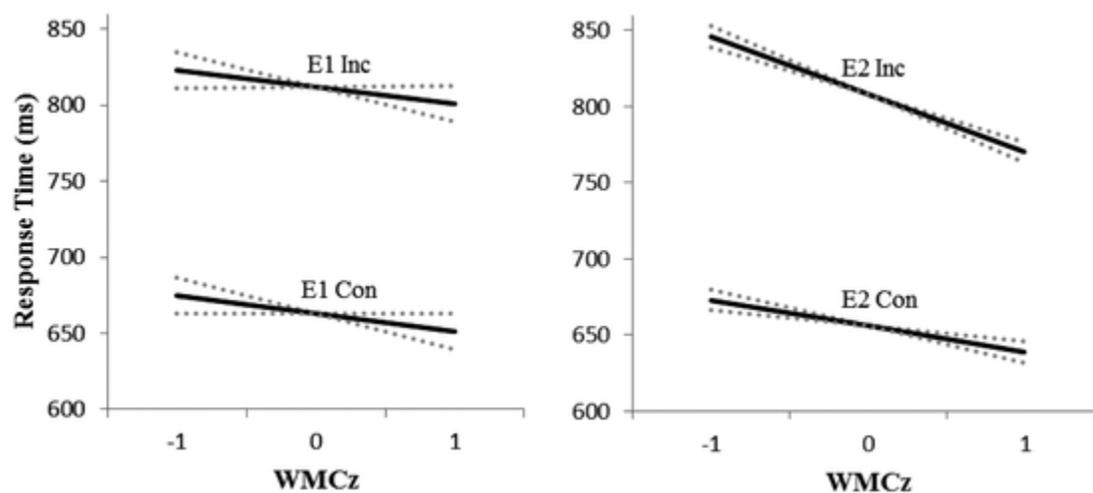

Figure 4. Experiment 1 (E1) and Experiment 2 (E2) model parameters for congruent (Con) and incongruent (Inc) reaction times in a high-congruency condition as a function of working memory capacity (WMCz). The dotted lines represent the standard error of the estimate. E2 parameters are collapsed over the stimulus-set-size conditions.

Discussion

In Experiment 2, we observed a relation between WMC and Stroop interference in both RTs and errors, commensurate with findings from Kane and Engle (2003). Looking at Figure 4, it appears that lower-WMC subjects experienced more interference in Experiment 2, whereas higher-WMC subjects experienced less. Yet, in a task producing these robust WMC-related individual differences in Stroop interference, we still did not observe any relation between WMC and congruency-sequence effects.

The between-subject manipulation of stimulus set size had no impact on the relations between WMC, Stroop interference, and congruency-sequence effects, and thus Experiment 2 yielded two independent replications of the WMC-interference correlation under high-congruency contexts (along with four separate high-congruency groups in Kane & Engle, 2003; see also Hutchison, 2011; Unsworth & Spillers, 2010). However, stimulus set size did alter the magnitude of

congruency-sequence effects on RTs, with the effects being greater in the three-color condition than in the six-color condition (most likely due to the greater number of repetitions in a condition with fewer stimuli). Neither the immediate, local effects of congruent-trial runs, nor of stimulus repetitions, seemed to be responsible for the WMC-related differences we observed. A cross-experiment comparison also indicated that the WMC-related changes in RT and error interference across experiments were mostly the product of performance on incongruent trials.

General Discussion

We assessed whether WMC-related individual differences in a color-word Stroop task are related to the control processes associated with congruency-sequence effects. Moreover, we tested whether this relation—if any—is moderated by proportion-congruency. We addressed the limitations of previous work by Keye et al. (2009) and Unsworth et al. (2012) by using a color-word task that has repeatedly produced WMC-related individual differences (Hutchison, 2011; Kane & Engle, 2003; Long & Prat, 2002; Unsworth & Spillers, 2010) and that, critically, allows for the dissociation of top-down and bottom-up contributions to congruency-sequence effects. If congruency-sequence effects varied with WMC, it might help specify how people with higher WMC are able to achieve their superior Stroop performance. That is, if WMC was related to congruency-sequence effects, with the bottom-up contributions of stimulus feature repetitions controlled, we could attribute higher-WMC subjects' better performance, at least in part, to superior conflict monitoring or cognitive-control implementation on a trial-by-trial basis.

WMC and Stroop Interference

In Experiment 1, we created a modified six color-word Stroop task that did not permit stimulus feature repetitions or any long runs of consecutive congruent (in the high-congruency condition) or incongruent (in the low-congruency condition) trials. We did not find the expected WMC-related individual differences in a high-congruency condition. Rather than ignore this failure to replicate (and risk contributing to a "false positive psychology" or to inflating the field's estimates of the relevant effect sizes; Ioannidis, 2008; Simmons et al., 2011), we took the approach that null WMC correlations to other variables can be informative to theories of executive control, as a form of discriminant validity (e.g., Kane et al., 2006; Poole & Kane, 2009). That is, executive control remains a poorly understood and variably defined construct, and so identifying the boundaries of the associations between WMC and ostensible executive measures can help illuminate the cognitive processes that underlie these constructs.

We therefore created three- and six-color high-congruency Stroop tasks for Experiment 2, and we removed the constraints imposed on the Experiment 1 task. Removing these constraints created a task context where individual differences in WMC now affected Stroop performance, for subjects in both the three- and six-color conditions, providing two replications of prior findings (e.g., Hutchison, 2011; Kane & Engle, 2003). To better understand the influences of the Experiment 1 task constraints, we assessed in Experiment 2 the local impact of allowing trial-to-trial stimulus-feature repetitions, and of removing limits on the number of current consecutive congruent trials. We found that neither of these task parameters was responsible for producing local, in-the-moment effects that subsequently drove global WMC-related individual differences.

That is, we found an association between performance and WMC at the overall-task level (i.e., averaging across all trials in the task), but this WMC-interference relation could not be decomposed to find a specific, local context element that was responsible.

Thus, the methodological changes between Experiments 1 and 2 appear to have worked at a global, task level—either together or separately—to create a context that differentially engaged the attentional-control components of working memory for those people who had the capability to use them effectively. The cross-experimental analysis we conducted suggests that subjects with higher WMC were able to tune their performance to respond particularly quickly and accurately on the rare incongruent trials in Experiment 2, whereas lower WMC subjects' performance suffered. Our findings suggest that either the overall presence of occasional long congruent-trial runs in Experiment 2, or the frequent repetition of stimulus features across trials in Experiment 2 (or the influence of both combined), led lower-WMC subjects to more frequently neglect the task goal than did the Experiment 1 task context. Indeed, Kane and Engle (2003) originally reported that WMC's apparent influence on goal-neglect differences was significantly modified by aspects of the experimental context. Most obvious was the influence of proportion congruency: High- versus low-WMC differences in Stroop interference were large in high-congruency tasks but were quite modest in low-congruency tasks.

However, Kane and Engle (2003) also discovered much more subtle task influences that may help contextualize the present Experiment 1 results. WMC-related differences in error interference varied not only with current-task congruency but also with prior-task congruency (in most of their experiments, subjects completed two blocked Stroop tasks varying in congruency proportion). In Stroop tasks with 75% congruent trials, lower-WMC subjects showed exaggerated error interference relative to higher-WMC subjects only when the high-congruency task stood alone or was completed first (Experiments 1 and 2). WMC no longer predicted error interference when the 75%-congruent task followed a 0%-congruent task (Experiments 3A and 3B). That is, extended practice in a 0%-congruent Stroop task, which arguably reinforced the *ignore-the-word* goal, minimized subsequent goal neglect in lower-WMC subjects when they transferred to the unsupportive 75%-congruent task. In fact, this 0%-congruent task practice helped subjects with lower WMC to perform as accurately as those with higher WMC. However, note further that this practice effect seems limited to the 0%-congruent task context. When a high-congruent Stroop task followed a 20%-congruent trial block (Kane & Engle, 2003, Experiment 4), lower-WMC subjects again showed more goal neglect in the subsequent high-congruent task than did higher-WMC subjects (with lower-WMC subjects reverting to their no-practice levels of neglect). The simple—and modest—increase of incongruent trials, from "none" (0%) to "some" (20%), then, was enough to disrupt lower-WMC subjects' implementation of executive control. Clearly, relatively minor changes in congruency proportions can have measurable effects on WMC-related differences in Stroop interference and goal neglect. It thus seems less surprising that the similarly modest task constraints we imposed in Experiment 1 also affected lower-WMC subject's performance, even if fully explaining their influence will require further research.

WMC and Congruency-Sequence Effects

In Experiment 1, we found significant congruency-sequence effects and proportion-congruency effects, but no interaction between them. These findings seem to align with those of Funes et al. (2010). They presented subjects with two types of conflict trials (one Stroop-like, the other Simon-like), with conflict type randomly determined form trial to trial. The key result was a dissociation between the types of control exhibited by congruency-sequence effects and performance in differing proportion-congruency contexts. Proportion-congruency manipulations engaged a more general type of control than congruency-sequence effects and affected both types of conflict trials (i.e., smaller conflict effect with lower proportion-congruency and greater conflict effects with higher proportion-congruency for both Stroop-like and Simon-like trials). Congruency-sequence effects, in contrast, were only present between trials that presented the same type of conflict (i.e., reductions in conflict effects for incongruent Stroop-like trials following incongruent Stroop-like trials but not for incongruent Stroop-like trials following incongruent Simon-like trials). Funes et al. thus suggested that the types of control associated with congruency-sequence effects and proportion-congruency effects are independent: Congruency-sequence effects reflect the work of an attentional system specific to the resolution of in-the-moment conflict, whereas proportion-congruency effects result from an attentional system that is responsive to contexts where conflict is present.

Of central importance, we did not find any relation between WMC and congruency-sequence effects in Experiment 1, but, of course, we did not find any overall WMC effects there, either. In Experiment 2, after removing task constraints and finding significant WMC effects in overall Stroop interference, we still did not find any WMC-related variation in congruency-sequence effects (whether or not we included in the analysis those trials that presented stimulus-feature repetitions). Following Funes et al. (2010), then, we conclude that congruency-sequence effects and proportion-congruency effects reflect independent levels (or processes, or circuits) of executive control and that WMC predicts variation in only the control processes that respond to global proportion-congruency effects. To be clear, proportion-congruency effects are evident for both higher- and lower-WMC subjects (see Figure 1), but higher-WMC subjects seem better able to engage additional preemptive control processes when triggered by a task context that challenges goal maintenance and resolving response competition (Kane & Engle, 2003).

Our conclusion, that WMC does not predict congruency-sequence effects, also jibes with recent work by Unsworth et al. (2012), who found null WMC × Congruency-Sequence correlations in flanker and manual-Stroop tasks, and by Keye et al. (2009), who found null WMC correlations in a flanker task. We extended Unsworth et al. by distinguishing the potential contributions of feature-integration and conflict-monitoring processes to congruency-sequence effects. Unsworth et al.'s task design did not allow for the experimental or statistical control of repetitions (and so may have minimized top-down influences and maximized bottom-up influences; e.g., Mayr et al., 2003) and left open the possibility that WMC might predict top-down conflict-monitoring influences on congruency-sequence effects when bottom-up processes were controlled. Our experiments indicate that it does not. Indeed, our results show that WMC's association with global performance in Stroop tasks does not arise from more efficient or effective trial-to-trial transitions. In contrast, Keye et al. (2009) found WMC to be related to congruency-sequence effects in a vertical Simon task (RTs only) but not in a horizontal flanker task. In Keye et al.,

higher WMC subjects did not outperform lower WMC subjects overall in either task. We tested whether Keye et al.'s discrepant between-task findings resulted from using tasks that did not produce overall WMC-related individual differences. This was not the case. We found no difference in the relations between WMC and congruency sequence effects in a task that did (Experiment 2) and did not (Experiment 1 high-congruency condition) produce WMC-related individual differences.

Our results do seem to contradict the more complex pattern of results reported by Hutchison (2011). Recall that WMC moderated the association between RT congruency-sequence effects and proportion congruency (i.e., WMC interacted with previous-trial congruency and overall congruency proportion in predicting Stroop interference effects). For higher-WMC subjects, Stroop interference was equivalent following congruent and incongruent trials in high- and low-congruency lists; subjects with higher WMC were thus unaffected by previous-trial congruency and overall congruency proportion. For lower-WMC subjects, in contrast, Stroop interference was greater following congruent than incongruent trials in a low-congruency context (as would be expected if a previous incongruent trial triggered top-down control), but interference was *smaller* following congruent than incongruent trials in a high-congruency context. Under low congruency, lower-WMC subjects seemed to adjust their control setting on a trial-by-trial basis as expected. Under high congruency, however, they suffered more RT interference following incongruent trials than under any other task condition. We are not sure what to make of this unexpected, if not unique, finding, but we failed to replicate it as a three-way interaction among WMC, congruency-sequence effects, and proportion-congruency in our Experiment 1 data ($b = 1$ ms, $SE = 7$ ms, $t = 0.1$). Perhaps we did not replicate Hutchison's finding because he used a heavily modified Stroop task, with the congruency of particular color-words manipulated within list-level manipulations of congruency proportion. (For example, within a high-congruency list, two critical words were presented in their own color 67% of the time, whereas two other critical words were presented in their own color only 33% of the time; two filler words were presented in their own color 100% of the time to create the overall congruency of the list [67%].) Note that Hutchison's results were also unusual in that, contrary to our results and other Stroop findings in the literature (e.g., Kerns et al., 2004; Unsworth et al., 2012), they yielded no overall congruency-sequence effects in RT. Thus, the manipulation of particular color-words may have led to a different type of processing than is normally used in a more prototypical color-word Stroop task (or other commonly used conflict tasks). As we have shown here, in Experiment 1 (see also Kane & Engle, 2003), WMC-related differences in interference-control processes are rather sensitive to the task context, and therefore Hutchison's unique findings may not generalize outside of that specific task environment.

Conclusions

Using color-word Stroop tasks, we found that individual differences in WMC are not associated with those in congruency-sequence effects, reflecting the trial-to-trial adjustments that subjects appear to make in response to experienced conflict (see also Keye et al., 2009; Unsworth et al., 2012). Higher- and lower-WMC subjects responded similarly, in the moment, to the prior exposure to incongruent versus congruent Stroop trials, despite the facts that (1) our task

preparations minimized the impact of bottom-up contributions and maximized the potential for top-down contributions to congruency-sequence effects, and (2) our second experiment yielded WMC-related differences in overall Stroop interference. Thus, although WMC is broadly associated with many indices of cognitive control, such as Stroop interference (e.g., Kane & Engle, 2003; Long & Prat, 2002), flanker interference (e.g., Heitz & Engle, 2007; Redick & Engle, 2006), antisaccade performance (e.g., Kane, Bleckley, Conway, & Engle, 2001; Unsworth, Schrock, & Engle, 2004), and vulnerability to involuntary mind-wandering experiences (e.g., McVay & Kane, 2009), it is not associated with *all* phenomena that arguably reflect some form of executive functioning (see also Kane et al., 2007, 2006). Indeed, our additional finding that WMC did not predict Stroop interference in a high-congruency task that restricted consecutive trial types and stimulus-dimension overlap (Experiment 1), but did in two high-congruency tasks that allowed these prototypical trial-to-trial transitions (Experiment 2), also demonstrates the selective nature of WMC's affiliation with control-related behaviors. By continuing to delineate the boundaries around the WMC-control association, we argue that the field will continue to better understand the mental processes underlying both WMC and cognitive control.

Footnotes

1 Kerns et al. (2004) and Notebaert et al. (2006) similarly found congruency-sequence effects in a Stroop task after statistically removing partial and complete repetition trials from their analyses.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. doi:10.1016/j.jml.2007.12.005

Bates, D., Maechler, M., & Bolker, B. (2011). *lme4: Linear mixed-effects models using S4 classes*. Retrieved from http://CRAN.R-project.org/package=lme4

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652. doi:10.1037/0033-295X.108.3.624

Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. (Eds.). (2007). *Variation in working memory*. New York, NY: Oxford University Press.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769–786. doi:10.3758/BF03196772

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language*, *59*, 447–456. doi:10.1016/j.jml.2007.11.004

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology*, *30*, 257–303. doi:10.1006/cogp.1996.0008

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

Engle, R. W., & Kane, M. J. (2003). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of Learning and Motivation*, *44*, 145–199. doi:10.1016/S0079-7421(03)44005-X

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*, 143–149. doi:10.3758/BF03203267

Funes, M. J., Lupiáñez, J., & Humphreys, G. (2010). Sustained vs. transient cognitive control: Evidence of a behavioral dissociation. *Cognition*, *114*, 338–347. doi:10.1016/j.cognition.2009.10.007

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. doi:10.1037/0096-3445.121.4.480

Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology*, *24*, 1149–1167. doi:10.1002/acp.1624

Heitz, R. P., & Engle, R. W. (2007). Focusing the spotlight: Individual differences in visual attention control. *Journal of Experimental Psychology: General*, *136*, 217–240. doi:10.1037/0096-3445.136.2.217

Hommel, B., Proctor, R. W., & Vu, K.-P. L. (2004). A feature-integration account of sequential effects in the Simon task. *Psychological Research*, *68*, 1–17. doi:10.1007/s00426-003-0132-y

Hutchison, K. A. (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 851–860. doi:10.1037/a0023437

Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology*, *19*, 640–648. doi:10.1097/EDE.0b013e31818131e7

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*, 169–183.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory as variation in executive attention and control. In A. R. A.Conway, C.Jarrold, M. J.Kane, A.Miyake, & J. N.Towse (Eds.), *Variation in working memory* (pp. 21–48). New York, NY: Oxford University Press.

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

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*, 189–217. doi:10.1037/0096-3445.133.2.189

Kane, M. J., Poole, B. J., Tuholski, S. W., & Engle, R. W. (2006). Working memory capacity and the top-down control of visual search: Exploring the boundaries of "executive attention."*Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 749–777. doi:10.1037/0278-7393.32.4.749

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004, February13). Anterior cingulate conflict monitoring and adjustments in control. *Science*, *303*, 1023–1026. doi:10.1126/science.1089910

Keye, D., Wilhelm, O., Oberauer, K., & van Ravenzwaaij, D. (2009). Individual differences in conflict-monitoring: Testing means and covariance hypothesis about the Simon and the Eriksen Flanker task. *Psychological Research*, *73*, 762–776. doi:10.1007/s00426-008-0188-9

Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2011). Experimental effects and individual differences in linear mixed models: Estimating the relationship between spatial, object, and attraction effects in visual attention. *Frontiers in Psychology*, *1*, 238. doi:10.3389/fpsyg.2010.00238

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?!*Intelligence*, *14*, 389–433. doi:10.1016/S0160-2896(05)80012-1

Long, D. L., & Prat, C. S. (2002). Working memory and Stroop interference: An individual differences investigation. *Memory & Cognition*, *30*, 294–301. doi:10.3758/BF03195290

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

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 196–204. doi:10.1037/a0014104

Notebaert, W., Gevers, W., Verbruggen, F., & Liefooghe, B. (2006). Top-down and bottom-up sequential modulations of congruency effects. *Psychonomic Bulletin & Review*, *13*, 112–117. doi:10.3758/BF03193821

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York, NY: Springer.

Poole, B. J., & Kane, M. J. (2009). Working memory capacity predicts the executive control of visual search among distractors: The influences of sustained and selective attention. *Quarterly Journal of Experimental Psychology*, *62*, 1430–1454.

Purmann, S., Badde, S., & Wendt, M. (2009). Adjustments to recent and frequent conflict reflect two distinguishable mechanisms. *Psychonomic Bulletin & Review*, *16*, 350–355. doi:10.3758/PBR.16.2.350

Quené, H., & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, *59*, 413–425. doi:10.1016/j.jml.2008.02.002

R Development Core Team. (2011). *R: A language and environment for statistical computing*. Retrieved from http://www.R-project.org/

Redick, T. S., & Engle, R. W. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology*, *20*, 713–721. doi:10.1002/acp.1224

Shute, V. J. (1991). Who is likely to acquire programming skills?*Journal of Educational Computing Research*, *7*, 1–24. doi:10.2190/VQJD-T1YD-5WVB-RYPJ

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359–1366. doi:10.1177/0956797611417632

Simon, J. R., & Small, A. M., Jr. (1969). Processing auditory information: Interference from an irrelevant cue. *Journal of Applied Psychology*, *53*, 433–435. doi:10.1037/h0028034

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

Stürmer, B., Leuthold, H., Soetens, E., Schröter, H., & Sommer, W. (2002). Control over location-based response activation in the Simon task: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1345–1363. doi:10.1037/0096-1523.28.6.1345

Unsworth, N., Heitz, R. C., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*, 498–505. doi:10.3758/BF03192720

Unsworth, N., Redick, T. S., Spillers, G. J., & Brewer, G. A. (2012). Variation in working memory capacity and cognitive control: Goal maintenance and microadjustments of control. *Quarterly Journal of Experimental Psychology*, *65*, 326–355. doi:10.1080/17470218.2011.597865

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1302–1321. doi:10.1037/0278-7393.30.6.1302

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language*, *62*, 392–406. doi:10.1016/j.jml.2010.02.001