Psychology and Aging

Interactive Effects of Working Memory and Trial History on Stroop Interference in Cognitively Healthy Aging

Andrew J. Aschenbrenner and David A. Balota

CITATION

Interactive Effects of Working Memory and Trial History on Stroop Interference in Cognitively Healthy Aging

Andrew J. Aschenbrenner and David A. Balota
Washington University in St. Louis

Past studies have suggested that Stroop interference increases with age; however the robustness of this effect after controlling for processing speed has been questioned. Both working memory (WM) and the congruency of the immediately preceding trial have also been shown to moderate the magnitude of Stroop interference. Specifically, interference is smaller both for individuals with higher working memory capacity and following an incongruent trial. At present, it is unclear whether and how these 3 variables (age, WM and previous congruency) interact to predict interference effects in the standard Stroop color-naming task. We present analyses of Stroop interference in a large database of Stroop color-naming trials from a lifespan sample of well-screened, cognitively healthy, older adults. Our results indicated age-related increases in interference (after controlling for processing speed) that were exaggerated for individuals with low WM. This relationship between age and WM occurred primarily when the immediately preceding trial was congruent. Following an incongruent trial, interference increased consistently with age, regardless of WM. Taken together, these results support previous accounts of multiple mechanisms underlying control in the Stroop task and provide insight into how each component is jointly affected by age, WM, and trial history.

Keywords: aging, attention, working memory, congruency-sequence effect

Attentional control is a multicomponent construct that underlies our capacity to direct attention toward relevant and away from irrelevant stimuli in the environment. One of the most widely used tasks to study attentional control is Stroop color naming (MacLeod, 1991; Stroop, 1935). In the Stroop paradigm, individuals name the ink color in which color words are printed (e.g., the word RED printed in blue ink) and therefore must attend to the color (i.e., blue) while controlling the irrelevant, but highly salient, word information (i.e., red). Responses are slower and often less accurate when the color and word information are incongruent (RED in blue ink) compared with when they are congruent (RED in red ink), and the magnitude of this interference is often used as an index of attentional control. Attentional control is hypothesized to decrease in healthy aging and indeed, Stroop interference (SI) has been reported to be larger in older adults (e.g., Bugg, DeLosh, Davalos, & Davis, 2007; Jackson & Balota, 2013; Spieler, Balota, & Faust, 1996), but the existence of this effect after controlling for general slowing has been questioned (Verhaeghen & De Meersman, 1998).

Given the debate regarding age effects in the Stroop task, it is important to consider additional factors which may contribute to the observed magnitude of SI. For example, working memory (WM) has been shown to moderate SI particularly when demands on goal maintenance are high (Hutchison, 2011; Kane & Engle, 2003; Long & Prat, 2002). Goal maintenance refers to the ability to maintain the task set (e.g., "name the color") across an experiment and is typically examined by manipulating the proportion of congruent stimuli within the list. Because frequent congruent items may encourage a word-reading strategy, stronger demands are placed on internal goal representations to maintain the color-naming task set when the proportion of congruent trials is relatively high. Because WM differences primarily manifest under conditions of frequent congruent items, researchers have argued that WM is a key index of goal maintenance ability in the Stroop task (e.g., Kane & Engle, 2003).

A second factor that moderates SI is the degree to which attentional control is dynamically adjusted on a trial-by-trial basis. Specifically, characteristics of the stimulus on Trial $N-1$ may trigger a tightening or relaxing of the attentional control parameters and these new settings will carry over to influence the response dynamics on Trial N. If the updated control parameters are inappropriate for Trial N, they can be updated yet again for trial $N + 1$. This constant updating of the control system allows the responder to find the optimal settings to best complete the task. Such trial-level changes are typically measured using the congruency-sequence effect (CSE), or the "Gratton effect" (Gratton, Coles, & Donchin, 1992), which is the finding that interference is reduced following incongruent trials relative to congruent trials. More specifically, responses to congruent stimuli are faster when preceded by another congruent stimulus (CC trials) than when they are preceded by an incongruent stimulus (IC trials).

Similarly, incongruent responses are faster when preceded by an incongruent trial (II trials) than when preceded by a congruent trial (CI trials).

Several competing theories of the CSE have been proposed, the most prominent of which appeals to a conflict-monitoring mechanism (Botvinick, Braver, Barch, Carter, & Cohen, 2001). The conflict-monitoring account proposes that attentional control is dynamically adjusted in reaction to response conflict (e.g., the word information conflicts with the color information). Within this framework, researchers have argued that the amount of conflict on a given trial is detected by the anterior cingulate cortex, which then signals prefrontal brain regions to trigger appropriate adjustments in control by increasing or decreasing the relative activation of the color and word-response pathways. Thus, for example, an incongruent trial will result in a high level of response conflict, which then leads to an increase in control, which in turn reduces interference on the subsequent trial. In this way, the conflict-monitoring framework can accommodate the CSE, and indeed has been successful in accounting for a wide variety of other experimental phenomena (Botvinick, Cohen, & Carter, 2004). It is important to note, activity in the anterior cingulate cortex (i.e., the proposed conflict monitor) has been linked to the reduction in SI on the subsequent trial, lending further support to the conflict-monitoring account (Kerns et al., 2004, but see Mayr & Awh, 2009).

Despite the success and intuitive appeal of conflict monitoring, several alternative theories have questioned the need for a dedicated conflict-monitoring mechanism, at least in accounting for the CSE. For example, in many interference paradigms, such as the Simon or flanker tasks, there are only two types of stimuli (e.g., a left- or right-facing arrow in a Simon-type task). Thus, when the congruency repeats (CC or II trials), there is a .5 probability that the target and distracters will be exactly the same as on the previous trial. Therefore, the relative speeding of CC and II trials might potentially be due to simple repetition priming. Indeed, several studies have shown that when such repetitions are removed prior to analysis, the CSE is greatly reduced or eliminated (Mayr, Awh & Laurey, 2003; Mayr & Awh, 2009). However, other studies have found evidence for the CSE, even after controlling for cross-trial feature overlap (e.g., Meier & Kane, 2013; Notebaert, Gevers, Verbruggen & Liefooghe, 2006; Ullsperger, Bylsma & Botvinick, 2005), suggesting that feature repetition does not completely account for these effects at least in certain paradigms.

In addition to the contribution of stimulus repetition, other mechanisms have also been proposed to play a role in the CSE, including joint adjustment of the visuomotor activation gain and output threshold (Schlaghecken & Martini, 2012), expectation adjustment and response slowing (Lamers & Roelofs, 2011), and contingency learning (Schmidt, 2013). Thus, while the presence of a CSE indicates that adjustments are being made to the underlying response dynamics on a trial-level basis, the trigger for these adjustments remains under debate.

Importantly, one might expect that such trial-level control adjustments contribute to age or WM differences in performance on conflict tasks. A priori, one may expect individuals with better control systems (i.e., higher WM and younger age) to exhibit larger cross-trial effects, because a highly tuned attentional control system should be better able to flexibly adapt to local trial characteristics to maximize performance. For example, an individual

with a relatively compromised control system might be less able to increase his or her level of control after responding to a stimulus that produces a high degree of conflict (i.e., an incongruent stimulus). An increase in control would result in the increase of activation along the color pathway and a decrease of activation along the word pathway, which as described above, will reduce SI producing the CSE. However, if control is not adjusted in this manner, the magnitude of the CSE will be minimal.

We are aware of two studies that have examined age differences on the CSE. West and Moore (2005) utilized a modified key-press Stroop task that included a switching component. Specifically, 75% of the trials were color-identification trials and the remaining 25% were word-identification trials. It is important to note, the Age × Current Trial × Previous Trial interaction was not significant ($p > .22$), indicating similar CSEs in both younger and older adults. However, the sample size was relatively small ($N = 12$ per age group), and indeed the pattern of means indicated the CSE was larger in older adults (CSE = 120 ms) than in younger adults (CSE = 68 ms). Similarly, Puccioni and Vallesi (2012) used a manual Stroop task and specifically controlled for repetition-priming effects by ensuring that no physical characteristics of stimuli were repeated in immediately adjacent trials. These authors also failed to find any evidence of age differences in the CSE; however, it should be noted that they also did not obtain a reliable CSE overall.

Several studies have also examined the relationship between WM capacity and the CSE with mixed results. Two studies found no relationship between WM and the CSE in a young adult sample, either in the Stroop task (Meier & Kane, 2013; Unsworth, Redick, Spillers, & Brewer, 2012) or in a flanker task (Unsworth et al., 2012). On the other hand, Keye, Wilhelm, Oberauer, and van Ravenzwaaij (2009) used structural equation modeling to specify IC and CI trials as indicators of a latent "context" factor for both a flanker and a Simon task. They then correlated levels of WM with this latent factor and found a negative correlation between WM and the CSE indicating that lower WM individuals produced larger CSEs. However, this relationship was statistically significant only for the Simon task ($r = -.22$, $p < .05$) and not the flanker task ($r = -.18$, $p > .05$). Using a Stroop paradigm, Hutchison (2011) found a three way interaction among WM, previous trial and proportion congruency in predicting SI. This complex interaction indicated that low WM participants showed the standard CSE in the context of frequent incongruent items (as might be expected); however, these same participants exhibited *greater* SI following incongruent trials in the context of frequent congruent items. In contrast, for high WM individuals, the standard CSE was obtained in the frequent congruent context, but was eliminated in the frequent incongruent context.

Thus, at present there appears to be little evidence to indicate that age moderates the CSE, and the evidence regarding the effect of WM is equivocal, suggesting that the moderating effect of WM may be specific to certain tasks or contexts. However, it is important to note that the studies reviewed here typically use "nonstandard" variants on the interference tasks. For example, a proportion-congruency manipulation is often included (e.g., Hutchison, 2011; Meier & Kane, 2013) or a task-switching component might be introduced (e.g., West & Moore, 2005). It is unclear how these manipulations might influence the relationship between the CSE and measures of individual differences. Furthermore, we are not

aware of any study that has looked at the joint relationship among age, WM, and trial history in a single study. Thus, the present study was designed to examine the interplay among these three variables in predicting SI in a large sample of well-characterized, cognitively healthy, older adults. Because of the complexity of the design, it is critical to have sufficient sample size to examine these issues, which we employed for the current study ($N = 435$). It is also important to note, we used a standard, vocal Stroop task that is commonly used to investigate age differences in SI.

The present research was motivated by two major questions. First, are age-related changes in global SI moderated by WM? There is reason to predict this because, as mentioned, WM is hypothesized to reflect the ability to maintain task goals throughout the experiment (Kane & Engle, 2003), and to the extent that goals are better maintained, SI will be reduced. Therefore, a relatively high level of WM might offset age-related deficits in attentional control. Second, are age and/or WM differences in SI moderated by the congruency of the immediately preceding trial? Based on the conflict-monitoring hypothesis, one might expect individuals with relatively poor attentional control, as reflected by older age and/or lower WM, would produce smaller dynamic adjustments across trials. In other words, they may be less able to adjust control over the word and color pathways in response to recently presented items. However, as noted above, the available evidence regarding this hypothesis is mixed at best.

To address these questions, we conducted analyses on a large sample of Stroop color-naming trials as a function of age, WM, and trial history using linear mixed-effects models (LME). LMEs are becoming preferred to more traditional analyses of variance and are particularly useful for the present design. Specifically, LME analyses allow for the modeling of continuous covariates such as age and WM with experimental factors (current and previous congruency), and are robust to an imbalance in data that arises from the random intermixing of trial types (i.e., Current × Previous Congruency was not explicitly controlled). In addition, they allow for the modeling of the variance that arises from the specific sample and color words that we employed (Kliegl, Wei, Dambacher, Yan, & Zhou, 2011).

Method

Participants

Four hundred thirty-five healthy adults participated in this study. All participants were screened for dementia by highly trained physicians at the Charles F. and Joanne Knight Alzheimer's Disease Research Center at Washington University in St. Louis, Missouri using the sensitive Clinical Dementia Rating Scale (CDR; Morris, 1993). This research group is able to detect the very earliest stages of dementia with a high degree of accuracy (93%; Berg et al., 1998). This is important because the Stroop task has been shown to be particularly sensitive to this early stage of Alzheimer dementia (Balota et al., 2010; Spieler et al., 1996) and we wanted to insure that our participants were all cognitively healthy. Indeed, all participants in the present study were rated as CDR 0, indicating an absence of cognitive impairment. The mean Mini-Mental State Exam score (Folstein, Folstein, & McHugh, 1975) was 29 out of 30. The participants were drawn from a community-based sample (mean age = 67, $SD = 11$, range =

30–96), who completed a battery of tests designed to assess various aspects of attention and memory. Data from participants' first time completing this battery have been reported in the present study.

Stroop Task and WM Assessment

The Stroop task was based on the version of the task used in Spieler et al. (1996) and consisted of four color words printed in four different colors (*red*, *green*, *blue*, and *yellow*) with four neutral words (*bad*, *legal*, *poor*, and *deep*) included as a baseline. The list included 36 congruent trials (each word printed in the corresponding color nine times), 36 incongruent trials (each word printed in each of the three nonmatching colors three times), and 32 neutral trials (the words bad, legal, poor, and deep, each presented twice in each of the four colors), which were randomly intermixed. Participants were instructed to name the ink color of the stimulus aloud into a microphone which triggered the computer to record response latencies to the nearest millisecond. On each trial, the following events occurred: (a) fixation for 700 ms; (b) 50-ms blank screen; (c) presentation of the stimulus; (d) participant produced the response that triggered the offset of the stimulus; (e) the experimenter recorded the response using a keystroke or noted the response as a microphone error (e.g., stutters, false starts, responses that were too soft to trigger the microphone, etc.); and (f) 1750-ms intertrial interval. Participants were given a break halfway through the experimental block.

The WM measure was the computation span task used in McCabe, Roediger, McDaniel, Balota, and Hambrick (2010). In this task, participants verified the accuracy of equations (e.g., $7 - 4 = 3$) and were required to remember the middle digit. After a set number of equations (beginning with one and then increasing by one after every three correct recall trials), they were asked to recall the center digits in order. The number of equations administered before recall increased until either two out of three recall trials were in error or a maximum of seven equations (e.g., a span score of 7) was reached. The dependent variable used in the present study was the total span score, which refers to the number of trials correctly recalled in the last block for which the participant achieved at least two correct recall trials (i.e., the same block that determined the span score).

Analysis

To avoid the undue influence of extreme outliers, each individual's data were screened in the following manner. First, RTs faster than 200 ms and slower than 4000 ms were removed. Second, a mean and standard deviation were calculated from the remaining trials for each participant and any RTs that exceeded 3 SDs from the participant mean were also removed. A small subset of our participants had a relatively high number of RTs removed due to microphone errors (33 subjects had more than 10% of their trials removed). It is possible that high numbers of such trials are indicative of impending cognitive decline (e.g., Alzheimer dementia, see Balota et al., 2010), therefore we removed these subjects before conducting further analysis. Of the remaining subjects, our trimming procedure eliminated

3.6% of the total trials.[1] After this trimming procedure, we eliminated incorrect responses, the first trial (because the first trial did not have a previous congruency), and trials that occurred after an error (to avoid posterror slowing differences). After removing all of these invalid trials, there was an average of 10–11 correct RTs in each Current Congruency × Previous Congruency cell.

It is also important to also consider exact and partial repetition of stimuli characteristics when examining the CSE by either removing such trials prior to analysis or explicitly controlling them in the experimental design (cf. Mayr et al., 2003). Because we were interested in the influence of WM and age on the SI in a traditional Stroop paradigm, we first present analyses of all trials (including stimuli repetitions). We then report follow-up analyses of postcongruent trials only (i.e., trials following incongruent stimuli were excluded) with stimuli repetitions removed. Postincongruent trials were not examined due to the large number of trials that would be removed due to feature overlap. This strategy allowed us to be sure that any obtained effects were not due entirely to repetition priming while still retaining a sufficient number of trials for accurate estimates of the effects of interest.

For our primary analyses, trial-level, correct, raw RTs[2] were analyzed using the "lme4" package in R (Bates, Maechler, Bolker, & Walker, 2014). We first determined the random-effects structure by sequentially adding random intercepts for participants and items and random slopes for condition and previous condition, and assessing the increment in model fit after the inclusion of each additional random effect. Model fit increased significantly at each step and thus all random effects were included in the final model.

After this initial set of analyses, we entered the following variables into a single model: mean RT to neutral stimuli[3] (to control for differences in processing speed), trial number (to control for practice effects), age, WM, condition and previous condition, as well as all the two-, three-and four-way interactions among the four primary variables. Age and WM were standardized and entered as continuous covariates and condition and previous condition were coded as +0.5/−0.5 contrasts. We used mean naming latencies on neutral trials as a measure of processing speed, and thus neutral trials themselves were excluded from the present analysis. Finally, there is some uncertainty about how to calculate degrees of freedom in LME and as such, we relied on the procedure of a t value greater than 2.0 to indicate statistical significance (see Kliegl, Masson, & Richter, 2010). The use of this criterion led to identical inferences to when degrees of freedom were estimated using the Satterthwaite approximation, the results of which have also been provided (Kuznetsova, Brockhoff, & Christensen, 2014; Manor & Zucker, 2004)

Results

Response Latencies

Estimates of the fixed effects from our final model are presented in Table 1. There are several important points to note here. First, we observed an interaction among age, WM, and condition, indicating that age-related differences in the Stroop effect were moderated by WM. As shown in Figure 1, age-related differences in SI were particularly magnified when WM was low ($\beta = 33.4$, $t = 5.0$, $p < .0001$). Conversely, when WM was relatively high (1 SD

Table 1

Estimates of Fixed Effects on Response Latencies

Effect	Estimate	t value	p value
Intercept	827	89.9	<.0001
Neutral RT	137	43.9	<.0001
Trial number	−0.4	−7.8	<.0001
Age	−5.3	−1.7	.0909
Working memory (WM)	−1.6	−0.5	.5896
Condition (C)	166	9.5	<.0001
Previous condition (PC)	10.3	3.1	.0024
Age * WM	1.2	0.4	.6610
Age * C	20.5	4.3	<.0001
Age * PC	2.3	0.7	.6020
WM * C	−15.6	−3.2	.0014
WM * PC	−4.5	−1.3	.1878
C * PC	−43.1	−7.5	<.0001
Age * WM * C	−12.9	−2.8	.0051
Age * WM * PC	−0.2	−0.05	.9599
Age * C * PC	−2.1	−0.36	.7167
WM * C * PC	12.4	2.1	.0336
Age * WM * C * PC	19.1	3.5	.0006

above the mean), there was no significant effect of age ($\beta = 7.7$, $t = 1.2$ $p = .25$), suggesting that high WM can offset more general age-related deficits in Stroop performance.

Of primary interest is whether these targeted variables produced higher order interactions. Indeed, there was a significant four-way (i.e., age, WM, current condition, and previous condition) interaction. This interaction indicates that both the age- and WM-related differences in SI were critically moderated by the congruency of the immediately preceding trial. This interaction is displayed in Figure 2 and depicts the age-related increase in SI as a function of WM and split by previous congruent trials (top panel) and previous incongruent trials (bottom panel). As shown, large age and WM differences in SI occurred following congruent stimuli, yet these same effects were greatly reduced following incongruent stimuli. Considering just the top panel, for previous congruent trials there was a highly reliable three-way interaction among age, WM, and condition ($t = -4.3$, $p < .0001$). Specifically, there was no age-related increase in SI following congruent stimuli when WM was high ($\beta = -0.8$, $t = -0.11$), but robust increases were observed at average ($\beta = 21.6$, $t = 3.9$, $p = .0001$) and low WM ($\beta = 44.0$, $t = 5.7$, $p < .0001$). Turning to the SI following incongruent trials in the bottom panel, the three way interaction among age, WM, and condition was not reliable ($t = -0.6$, $p =$

---

[1] We repeated our analyses on the full sample of participants, including those with a higher number of outlier trials. It is important to note, the results were the same even in the larger sample.

[2] It is standard practice to transform the trial-level RTs when conducting LMEs to better approximate a normal distribution. Although such transformations can yield misleading inferences (Balota, Aschenbrenner, & Yap, 2013), when we repeated these analyses using log-transformed RTs, the critical four-way interaction remained significant.

[3] To further ensure that our effects were not due entirely to simple speed differences, we conducted identical analyses using proportional RT (RT–neutral RT) and within subject z-scored RT as the dependent measure rather than raw RT (Faust et al. 1999). Again, the four-way interaction remained reliable even in the transformed data. Furthermore, we tested for a correlation between the random intercept and random slope of the condition effect, which was not significant after including neutral RTs in the model ($p > .11$).
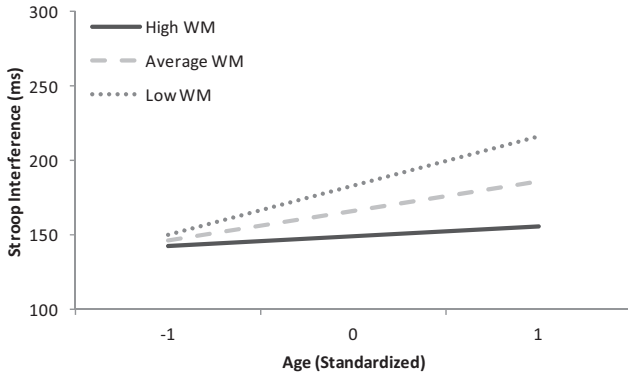
Figure 1. Regression lines of SI (incongruent RT–congruent RT) as a function of age and WM capacity.

.54), indicating that SI increased consistently with age regardless of WM.

To insure that our obtained effects were not due to repetition-priming effects, we analyzed postcongruent trials after removing exact and partial repetitions, which eliminated 36% of the data. Postincongruent trials (i.e., II and IC trials) were not analyzed for the reasons mentioned above. It is important to note, even with this smaller number of trials[4], we obtained a reliable Age $\times$ WM $\times$ Condition interaction ($\beta = -23.0$, $t = -3.5$, $p = .0005$). This interaction is plotted in Figure 3 and displays the same pattern that

Previous = Congruent

Previous = Incongruent

Figure 2. Regression lines of SI (incongruent RT–congruent RT) as a function of age and WM capacity and split by previous trial type. Top panel = previous congruent and bottom panel = previous incongruent.
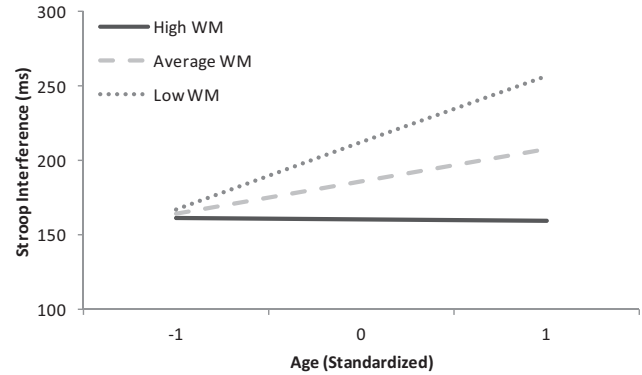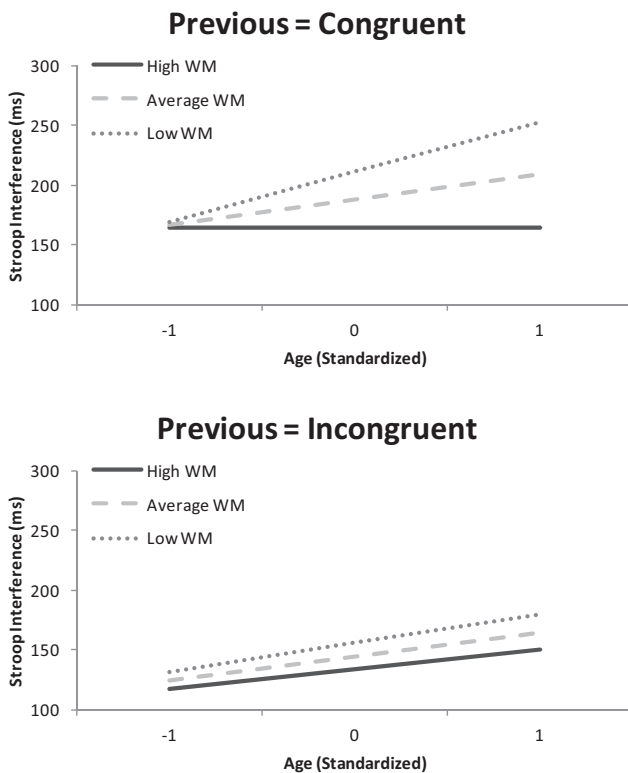

Figure 3. Regression lines of SI (incongruent RT–congruent RT) as a function of age and WM capacity after removing stimuli repetitions. Data are from trials following a congruent stimulus (i.e., CC and CI trials) only.

was previously noted. Specifically, SI increased significantly with age at low levels of WM ($\beta = 45.0$, $t = 4.7$, $p < .0001$) and average WM ($\beta = 22.0$, $t = 3.2$, $p = .001$), but not high WM ($\beta = -1.0$, $t = -0.11$, $p = .91$).

Accuracy Analysis

Although the primary interest in the present study was response latency, we also explored the same effects on accuracy. Specifically, we modeled the probability of producing an incorrect response as a function of current condition, previous condition, age, and WM with random intercepts across subjects. Errors due to stutters, false starts, or microphone errors were excluded prior to analysis. Briefly, as shown in Table 2, our results indicated a main effect of condition ($p < .001$) and a main effect of WM ($p = .007$). The main effect of condition reflects a greater number of errors produced in the incongruent condition (3%) than in the congruent condition (0%). The main effect of WM indicates that the probability of making an error decreased significantly with higher WM, as would be expected. No other main effects or higher order interactions were significant. Note that we failed to find evidence for a CSE in our accuracy measure (i.e., the Current Condition $\times$ Previous Condition interaction was not reliable, $p = .08$), probably due to the floor effects in the congruent condition.

Discussion

To date, the available literature has been mixed regarding SI and age, which suggests that there are additional factors potentially moderating age differences in SI. We have highlighted two of these factors in the present report. First, we demonstrated that WM had a moderating influence on age differences in SI. Specifically, SI increased with age at both average and low levels of WM, but not when WM was relatively high. This finding can be accommo-

---

[4] Although a large number of postcongruent trials have been removed due to feature overlap, we also examined the CSE after removing repetition trials. It is important to note, even with a smaller number of trials, we obtained a smaller but still significant condition by previous condition interaction ($\beta = -26.9$, $t = -2.9$, $p = .004$). The higher order interactions including age and WM were also still reliable ($ps < .05$).

Table 2
Estimates of Fixed Effects in the Analysis of Accuracy

Effect	Estimate	*z* value	*p* value
Intercept	−6.0	−23.6	<.0001
Age	0.31	1.37	.1714
Working Memory (WM)	−0.69	−2.7	.007
Condition (C)	3.33	7.7	<.0001
Previous Condition (PC)	0.60	1.4	.1604
Age*WM	−0.03	−0.15	.8776
Age*C	−0.05	−0.11	.9141
Age*PC	−0.17	−0.40	.6899
WM*C	0.16	0.34	.7330
WM*PC	0.41	0.84	.4007
C*PC	−1.48	−1.73	.087
Age*WM*C	−0.07	−0.17	.868
Age*WM*PC	−0.09	−0.22	.8269
Age*C*PC	0.26	0.29	.7667
WM*C*PC	−0.73	−0.76	.4456
Age*WM*C*PC	−.19	−0.24	.8099

dated by accounts of Stroop performance that contain two different processes, goal maintenance and conflict resolution (Kane & Engle, 2003). WM is hypothesized to influence goal maintenance (i.e., name the color, not the word), which allows for sustained attention away from the distracting dimension of the stimulus. This is typically examined by manipulating the relative frequency of congruent items in the list. As mentioned in the introduction, frequent congruent items place greater demands on goal maintenance. However, the effect of age on SI may not depend critically on this maintenance mechanism, as evidenced by a lack of age differences in the magnitude of interference as a function of the stimulus-to-response interval (Jackson & Balota, 2013). Hence, it is possible that the observed moderating influence of WM on the Age × SI interaction may reflect a link between WM measures and the quality of the task-set parameters.

Although the mechanisms by which WM influences Stroop performance is not entirely clear, our data indicated that a high level of WM can offset changes in age-related SI. For example, if the effect of age is to diminish control over the prepotent word dimension (Spieler et al., 1996), a high WM could compensate for this by proactively maintaining attention away from that dimension before the stimulus is even presented. This would lead to overall less influence of the word pathway, thereby diminishing SI.

A second critical factor is the congruency of the immediately preceding trial. We replicated the well-established phenomenon that SI is reduced following an incongruent trial relative to following a congruent trial, a finding that has typically been accommodated by assuming that the attentional control system can be dynamically adjusted on a trial-level basis. A priori, one would expect that those individuals who have relatively good attentional control systems (e.g., younger adults with high WM ability) would show the largest cross-trial changes. More specifically, high WM and younger individuals should be better able to dynamically adjust their control systems to the ongoing demands of the task. However, precisely the opposite pattern was found in these data. Specifically, it was the individuals with relatively poor attentional control systems (i.e., lower WM and older age) who exhibited the largest trial-level changes in interference. Indeed, a greater CSE for individuals with worse control systems is more consistent with the extant literature (see, e.g., West & Moore, 2005).

Importantly, age differences in SI were particularly magnified following congruent trials. Such trials cause relatively little response conflict, suggesting that something other than magnitude of response conflict on Trial $N-1$ seems to be increasing SI on Trial N, particularly for older adults, which is inconsistent with the conflict-monitoring account. These findings are similar to those of Lamers and Roelofs (2011), who showed that neutral trials trigger a CSE similar in size to incongruent trials, whereas congruent trials lead to a significantly larger CSE. Because incongruent trials cause a relatively large degree of response conflict, the largest differences in cross-trial adjustments should occur after these trials, which was not the case in these data. Thus, alternative mechanisms are likely to underlie the CSE in the Stroop task.

We have already discussed several such alternative theories. However, we propose that the present results may be best accommodated by an additional mechanism, which we refer to as pathway priming. This account relies on the notion that there are two pathways (color and word) to making a response in the Stroop task, along with a control system that biases the use of one pathway over the other (e.g., Cohen, Dunbar, & McClelland, 1990). Typically, one must exert control over the faster and prepotent word pathway to make a color response. The pathway-priming account extends this framework by assuming the relative contribution of each pathway can be moderated by the utility of that pathway in making a response on the immediately preceding trial. Specifically, when Trial $N-1$ is incongruent, the color pathway is primed because color is the selected aspect of an incongruent stimulus. This priming of the color pathway would benefit subsequent incongruent trials, wherein color is again the most relevant dimension, and potentially slow subsequent congruent trials, wherein the word pathway could be beneficial on some percentage of trials (see MacLeod, 1991). Thus, the result of priming the color pathway on incongruent $N-1$ trials would be a net decrease in the Stroop effect on trial N.

In contrast, when Trial $N-1$ is a congruent trial, the word pathway is primed due to the utility of the fast word pathway in making a correct response. This priming of the word pathway would benefit subsequent congruent trials wherein the word dimension is again useful in making the correct response, and slow down the subsequent incongruent trials when the highly activated word pathway needs to be controlled to produce the correct color response. Thus, the SI effect would be increased following congruent trials, precisely as observed in the present data. Indeed, such a priming mechanism was proposed by Spieler, Balota, and Faust (2000) to account for the unique effect of congruent stimuli at the level of the RT distribution. Spieler et al. further suggested that reliance on the word dimension could be modulated by task demands (e.g., proportion congruency manipulations) and the present results indicate that these effects can occur even after a single trial. Thus, age and WM differences in SI as a function of previous congruency can be attributed to a reduced ability to control cross-trial pathway priming, which results in greater interference.

It is important to note that the pathway-priming account nicely accommodates the observation that both age and WM modulate the size of this effect in the observed direction. Specifically, the participants with relatively poor attentional control systems (i.e., older adults with low WM ability) produced the largest CSE following congruent trials. This would be predicted because the irrelevant word pathway is more accessible following congruent trials, which demands greater control to produce the correct re-

sponse. Thus, individuals with less attentional control should have particular difficulty with subsequent incongruent trials.

Such a priming perspective places a key role on pathway utility, rather than on response conflict, in determining trial-level response dynamics, which is a subtle but important distinction from the conflict-monitoring account. We find it critical that both mechanisms predict reduced SI following incongruent trials but for different reasons. Pathway priming predicts reduced SI because there is no additional priming from the color pathway that must be controlled. Conflict monitoring, on the other hand, predicts this because the high level of response conflict from the stimulus on trial $N-1$ signals an increase of control (i.e., increase color-pathway activation and reduce word-pathway activation). Due to the similarity in the predicted outcomes, it is critical to examine additional variables, such as individual difference measures (e.g., age and WM), that may adjudicate between these theories, which is a key contribution of the present work.

Other accounts have similarly extended the basic Cohen, Dunbar, and McClelland (1990) architecture in accounting for the CSE. Specifically, Mayr and Awh (2009) suggested that, because it takes more time to respond to an incongruent stimulus, the control system, which is constantly feeding activation into the relevant pathway, will have more time to operate to increase the activation in the color pathway. Thus, after responding to an incongruent stimulus, the system will be primed to utilize the color pathway on subsequent trials. It is unclear however if this mechanism can accommodate the patterns uncovered here. Specifically, if the critical variable is the degree to which the control system can influence the two pathways per unit of time, one might expect that individuals with stronger control (and thus a greater influx of activity to the relevant pathway from the control system) would benefit more from a previous incongruent trial, and thus show a larger CSE. This was not observed in the current study. Future studies should seek to differentiate between these accounts.

Although there are a number of strengths of the present study, including the large sample of well-characterized participants, which afforded the opportunity to model the effect of multiple individual difference measures, there were also a few differences from past studies which may limit generalization. First, the present study included approximately equal proportions of congruent, incongruent, and neutral stimuli. Often, the proportion congruency is manipulated and it is possible that cross-trial effects are moderated by the relative probabilities of each stimulus type. However, it is important to note that a recent study found no relationship between proportion congruency and the magnitude of the CSE (Meier & Kane, 2013), so this may not be a strong limitation of the present study. A second limitation is that, although our sample encompassed much of the adult life span, the ages were somewhat skewed toward older participants. Both of the previous studies regarding age and the CSE (Puccioni & Vallesi, 2012; West & Moore, 2005) included a group of younger adults (college students) as a comparison group. It is possible that the differences we obtained here may not be observed when comparing such extreme groups.

Finally, as mentioned in the introduction, it is critically important to consider feature repetitions when investigating the CSE. Due to the number of trials in our database, we focused on analyses eliminating repetitions following congruent trials. The presence of a significant interaction allowed us to infer that modulation of the Stroop effect after congruent trials is not due entirely to feature overlap. However, this does not directly speak to the magnitude of the CSE. Furthermore, it is possible that the relatively low number of trials included in our design contributed to the observed magnitude of the CSE. Specifically, Mayr and Awh (2009) showed that the largest CSE occurred within the first 170 trials of the experiment, with relatively little CSE occurring after that. However, 170 trials would be on the upper end of the number of trials that might be administered in a standard computer-based Stroop task.

In summary, we observed robust age-related changes in SI indicating deficient attentional control processes in older individuals. It is important to note, this relationship was moderated both by levels of WM and also by the congruency of the preceding trial. Specifically, WM moderated age-related changes in SI primarily after congruent stimuli. This suggests that different processes are brought online based on trial history, and that these processes are subsequently influenced differentially by WM. These findings have potential implications for the assessment of attentional control of older adults using the Stroop task. Specifically, age differences in SI may be exaggerated if there are unequal proportions of congruent and incongruent stimuli tested. However, the CSE may be tapping a fundamentally different aspect of the attentional control system, and whether it differs from other measures (such as Stroop errors) in predicting cognitive outcomes (e.g., cognitive decline, Balota et al., 2010 or sensitivity to Alzheimer-related biomarkers, Duchek et al., 2013), remains to be determined. What is more important, however, is that these results provide useful insight into attentional control systems across the adult lifespan and the potential mechanisms underlying the CSE.

References

Balota, D. A., Aschenbrenner, A. J., & Yap, M. J. (2013). Additive effects of word frequency and stimulus quality: The influence of trial history and data transformations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 1563–1571. http://dx.doi.org/10.1037/a0032186

Balota, D. A., Tse, C. S., Hutchison, K. A., Spieler, D. H., Duchek, J. M., & Morris, J. C. (2010). Predicting conversion to dementia of the Alzheimer's type in a healthy control sample: The power of errors in Stroop color naming. *Psychology and Aging, 25,* 208–218. http://dx.doi.org/10.1037/a0017474

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7, http://CRAN.R-project.org/package=lme4

Berg, L., McKeel, D. W., Jr., Miller, J. P., Storandt, M., Rubin, E. H., Morris, J. C., . . . Saunders, A. M. (1998). Clinicopathologic studies in cognitively healthy aging and Alzheimer's disease: Relation of histologic markers to dementia severity, age, sex, and apolipoprotein E genotype. *Archives of Neurology, 55,* 326–335. http://dx.doi.org/10.1001/archneur.55.3.326

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. http://dx.doi.org/10.1037/0033-295X.108.3.624

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences, 8,* 539–546. http://dx.doi.org/10.1016/j.tics.2004.10.003

Bugg, J. M., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2007). Age differences in Stroop interference: Contributions of general slowing and task-specific deficits. *Neuropsychology, Development, and Cognition: Section B. Aging, Neuropsychology and Cognition, 14,* 155–167. http://dx.doi.org/10.1080/138255891007065

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the

Stroop effect. *Psychological Review, 97,* 332–361. http://dx.doi.org/10.1037/0033-295X.97.3.332

Duchek, J. M., Balota, D. A., Thomas, J. B., Snyder, A. Z., Rich, P., Benzinger, T. L., . . . Ances, B. M. (2013). Relationship between Stroop performance and resting state functional connectivity in cognitively normal older adults. *Neuropsychology, 27,* 516–528. http://dx.doi.org/10.1037/a0033402

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin, 125,* 777–799. http://dx.doi.org/10.1037/0033-2909.125.6.777

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

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

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. http://dx.doi.org/10.1037/a0023437

Jackson, J. D., & Balota, D. A. (2013). Age-related changes in attentional selection: Quality of task set or degradation of task set across time? *Psychology and Aging, 28,* 744–753. http://dx.doi.org/10.1037/a0033159

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. http://dx.doi.org/10.1037/0096-3445.132.1.47

Kerns, J. G., Cohen, J. D., MacDonald, A. W., III, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303,* 1023–1026. http://dx.doi.org/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. http://dx.doi.org/10.1007/s00426-008-0188-9

Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition, 18,* 655–681. http://dx.doi.org/10.1080/13506280902986058

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,* 238–249. http://dx.doi.org/10.3389/fpsyg.2010.00238

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2014). lmerTest: Tests for random and fixed effects for linear mixed effects models (lmer objects of lme4 package). R Package Version 2.0–11. http://CRAN.R-project.org/package=lmerTest

Lamers, M. J. M., & Roelofs, A. (2011). Attentional control adjustments in Eriksen and Stroop task performance can be independent of response conflict. *The Quarterly Journal of Experimental Psychology, 64,* 1056–1081. http://dx.doi.org/10.1080/17470218.2010.523792

Long, D. L., & Prat, C. S. (2002). WM and stroop interference: An individual differences investigation. *Memory & Cognition, 30,* 294–301. http://dx.doi.org/10.3758/BF03195290

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203. http://dx.doi.org/10.1037/0033-2909.109.2.163

Manor, O., & Zucker, D. M. (2004). Small sample inference for the fixed effects in the mixed linear model. *Computational Statistics & Data Analysis, 46,* 801–817. http://dx.doi.org/10.1016/j.csda.2003.10.005

Mayr, U., & Awh, E. (2009). The elusive link between conflict and conflict adaptation. *Psychological Research, 73,* 794–802. http://dx.doi.org/10.1007/s00426-008-0191-1

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

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working-memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology, 24,* 222–243. http://dx.doi.org/10.1037/a0017619

Meier, M. E., & Kane, M. J. (2013). Working-memory capacity and Stroop interference: Global versus local indices of executive control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 748–759. http://dx.doi.org/10.1037/a0029200

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology, 43,* 2412–2414. http://dx.doi.org/10.1212/WNL.43.11.2412-a

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. http://dx.doi.org/10.3758/BF03193821

Puccioni, O., & Vallesi, A. (2012). Conflict resolution and adaptation in normal aging: The role of verbal intelligence and cognitive reserve. *Psychology and Aging, 27,* 1018–1026. http://dx.doi.org/10.1037/a0029106

Schlaghecken, F., & Martini, P. (2012). Context, not conflict, drives cognitive control. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 272–278. http://dx.doi.org/10.1037/a0025791

Schmidt, J. R. (2013). The parallel episodic processing (PEP) model: Dissociating contingency and conflict adaptation in the item-specific proportion congruent paradigm. *Acta Psychologica, 142,* 119–126. http://dx.doi.org/10.1016/j.actpsy.2012.11.004

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 461–479. http://dx.doi.org/10.1037/0096-1523.22.2.461

Spieler, D. H., Balota, D. A., & Faust, M. E. (2000). Levels of selective attention revealed through analyses of response time distributions. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 506–526. http://dx.doi.org/10.1037/0096-1523.26.2.506

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

Ullsperger, M., Bylsma, L. M., & Botvinick, M. M. (2005). The conflict adaptation effect: It's not just priming. *Cognitive, Affective & Behavioral Neuroscience, 5,* 467–472. http://dx.doi.org/10.3758/CABN.5.4.467

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. *The Quarterly Journal of Experimental Psychology, 65,* 326–355. http://dx.doi.org/10.1080/17470218.2011.597865

Verhaeghen, P., & De Meersman, L. (1998). Aging and the Stroop effect: A meta-analysis. *Psychology and Aging, 13,* 120–126. http://dx.doi.org/10.1037/0882-7974.13.1.120

West, R., & Moore, K. (2005). Adjustments of cognitive control in younger and older adults. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 41,* 570–581. http://dx.doi.org/10.1016/S0010-9452(08)70197-7