Psychonomic Bulletin & Review (2018) 25:767–774
https://doi.org/10.3758/s13423-017-1427-0

BRIEF REPORT

Some further clarifications on age-related differences in Stroop interference

Maria Augustinova[1,2] · **David Clarys**[3] · **Nicolas Spatola**[1] · **Ludovic Ferrand**[1]

Abstract

Both the locus and processes underlying the age-related differences in Stroop interference are usually inferred from changes in magnitudes of standard (i.e., overall) Stroop interference. Therefore, this study addressed these still-open issues directly. To this end, a sample of younger (18–26 years old) and healthy older (72–97 years old) was administered the semantic Stroop paradigm (that assesses the relative contribution of semantic compared to response conflict both of which contribute to overall Stroop interference) combined with a single-letter coloring and cuing (SLCC) procedure. Independently of an increased attentional focus on the relevant color dimension of Stroop words induced by SLCC (as compared to all letters colored and cued, ALCC), greater magnitudes of standard Stroop interference were observed in older (as compared to younger) adults. These differences were due to greater magnitudes of response conflict whereas magnitudes of semantic conflict remained significant and unchanged by healthy aging and SLCC. Thus, this direct evidence places the locus of age-related differences in Stroop interference at the level of response conflict (as opposed to semantic and/or both conflicts). In terms of processes underlying these differences, the reported evidence shows that both age-groups are equally (in)efficient in (a) focusing on the relevant color dimension and (b) suppressing the meaning of the irrelevant word-dimension of Stroop words. Healthy older adults are simply less efficient in suppressing the (pre-)response activity primed by the fully processed meaning of the irrelevant word-dimension. Standard interpretations of age-related differences in Stroop interference and a more general issue of how attentional selectivity actually operates in the Stroop task are therefore reconsidered in this paper.

Keywords Aging · Attentional selectivity · Semantic conflict · Single-letter coloring and cueing · Stroop interference · Response conflict

Introduction

The Stroop task (Stroop, 1935) requires individuals to identify, as quickly and accurately as possible, the font color of written characters without reading them. Despite this requirement, the typical result is that individuals' identification times are longer and more error-prone for *color-incongruent* Stroop words (e.g., "BLUE" displayed in green ink; hereafter $BLUE_{\text{green}}$), than for *color-neutral* words (e.g., "DEAL" displayed in green ink, hereafter $DEAL_{\text{green}}$).

Although the magnitude of this latter difference – called *Stroop interference* (e.g., $BLUE_{\text{green}}$–$DEAL_{\text{green}}$)– remains constant across middle adulthood, it begins to increase at the age of 65 years (e.g., Comalli, Wapner, & Werner, 1962 for the first empirical demonstration). Indeed, significantly greater magnitudes of Stroop interference are habitually observed in healthy older adults as compared to their young counterparts

---

Electronic supplementary material The online version of this article (https://doi.org/10.3758/s13423-017-1427-0) contains supplementary material, which is available to authorized users.

✉ Maria Augustinova
maria.augustinova@univ-rouen.fr

✉ Ludovic Ferrand
ludovic.ferrand@uca.fr

[1] Université Clermont Auvergne, CNRS, LAPSCO, F-63000 Clermont-Ferrand, France

[2] Université de Rouen Normandie, Centre de Recherche sur les Fonctionnements et les Dysfonctionnements Psychologiques (CRFDP, EA 7475), F-76000 Rouen, France

[3] CNRS and Université de Poitiers, Poitiers, France

(e.g., Jackson & Balota, 2013; Li & Bosman, 1996, Spieler, Balota, & Faust, 1996).

The fact that greater magnitudes of Stroop interference persist in healthy older adults even after controlling for general slowing (e.g., Jackson & Balota, 2013; Li & Bosman, 1996) is taken as evidence that their selective attention declines with aging. Indeed, "The conflict between the relevant (color of the word) and irrelevant (name of the word) dimensions of the stimulus on incongruent trials presents a particularly difficult task for the selective attentional system. A system that efficiently suppresses the irrelevant dimension (i.e., the word) should exhibit faster color naming than a system in which impaired suppression of the word dimension allows greater competition between the word-name and the color name for response output." (Spieler et al., 1996, p. 461).

It should be noted that this rather consensual conceptualization of both the locus and processes underlying the age-related differences in Stroop interference remains inferred from changes in magnitudes of standard (i.e., overall) Stroop interference depicted above. Consequently, this paper attempts to shed a more direct light on just these issues.

Locus of age-related differences in Stroop interference

As mentioned above, greater magnitudes of Stroop interference observed in healthy older (as compared to younger) adults are usually attributed to less efficient suppressing of the word-dimension that leads in turn to greater magnitudes of response conflict (Spieler et al., 1996).

This reasoning is rooted in so-called *single-stage response competition accounts* (see e.g., Risko, Schmidt, & Besner, 2006 in this outlet) that consider Stroop interference as a unitary phenomenon resulting from a single source of conflict (i.e., *response conflict*; e.g., MacLeod, 1991). Within this latter view, the efficiency of suppressing of the irrelevant word-dimension determines the amount of evidence provided toward a response (e.g., blue for $BLUE_{\text{green}}$ or deal for $DEAL_{\text{green}}$). Given that for color-incongruent Stroop words ($BLUE_{\text{green}}$), this (incorrect) response is part of the response set, it competes (i.e., interferes) with the one cued by the relevant color dimension (i.e., green for $BLUE_{\text{green}}$).

Even though this latter view still dominates both psychological research and practice (Augustinova, Silvert, Spatola, & Ferrand, 2017; Risko et al., 2006), it is now challenged by multiple lines of research – that have given rise to what is now termed *multi-stage accounts* (Risko et al., 2006). Despite their differences (e.g., see Augustinova et al., 2017 for a review), these accounts share the idea that Stroop interference is a more complex phenomenon that goes beyond a single (i.e., response) conflict depicted above.

In line with this reasoning, several studies published in this journal depart from the idea that the standard (i.e., overall) Stroop interference specifically results from two kinds of conflicts: *stimulus conflict* (SC) that is semantic in its nature (hence semantic conflict) and *response conflict* (RC). In other words, these studies subscribe to so-called *SC-RC accounts of Stroop interference* (Augustinova et al., 2017 for this terminology; and e.g., Augustinova, Flaudias, & Ferrand, 2010; Augustinova, Silvert, Ferrand, Llorca, & Flaudias, 2015 for examples of these studies). In order to isolate their specific contribution to standard Stroop interference, these studies used the so-called *semantic Stroop paradigm.*

This paradigm supplements Stroop words depicted above with another kind of color-incongruent words (e.g., *SKY* displayed ingreen, hereafter SKY_{green} first used by Klein, 1964) that are only associated with a given color (i.e., SKY with blue). This addition –initially suggested by Neely and Kahan (2001) and first implemented by Manwell, Roberts, and Besner (2004) in this outlet– lies on the assumption that associated color-incongruent words only involve *semantic* conflict whereas standard color-incongruent words involve both *semantic* and *response* conflicts.

Indeed, because the meaning activated by the irrelevant word dimension of both color-incongruent words (e.g., $BLUE_{\text{green}}$and SKY_{green}) corresponds to and/or is closely related to a color (blue here), it is thought to interfere with processing of the meaning that is activated by the relevant color dimension (e.g., green) of these words. This conflict is likely to arise in an amodal semantic network because "delays of processing occur whenever distinct semantic codes are simultaneously activated, and that these delays become acute when the conflicting codes are values on a single dimension or closely related dimensions." (Seymour, 1977, p. 263; see also e.g., Augustinova et al., 2015 for corresponding N400-like evidence).

Additionally, once the irrelevant word dimension of standard color-incongruent words (e.g., $BLUE_{\text{green}}$) has been adequately processed, it primes a corresponding (pre-)response tendency that shares the same response set (hence interferes with) that the one primed by the meaning of the relevant color dimension (i.e., generates the response conflict as depicted above). Inversely, because the word dimension of associated color-incongruent words (e.g., SKY_{green}) does not activate (pre-)response tendencies linked to the associated color (e.g., press a blue button on seeing *SKY*; see Schmidt & Cheesman, 2005 for a direct demonstration), their response set does not overlap with that activated by the color dimension. Consequently, associated color-incongruent words (e.g., SKY_{green}) are, exactly like color-neutral ones ($DEAL_{\text{green}}$), free of response conflict (but see e.g., Hasshim & Parris, 2014; Klein, 1964 for a different view). This latter kind of items is also free of semantic conflict.

In line with these different assumptions, these studies consistently observed the delay in processing (i.e., interference) for both types of color-incongruent words ($BLUE_{\text{green}}$ and SKY_{green}) compared to color-neutral ones (PUT_{green}) with the magnitude of *standard Stroop interference* ($BLUE_{\text{green}}$–PUT_{green}) being significantly greater than the one of *semantic Stroop interference* (SKY_{green}–PUT_{green}). Thus these different studies successfully demonstrated that both semantic and response conflicts jointly contribute to the standard (i.e., overall) Stroop interference (Augustinova et al., 2010, 2015; Manwell et al., 2004; see also e.g., White, Risko, & Besner, 2016).

Building on these studies, the first goal of this paper was to examine the *locus* of age-related differences in Stroop interference. More specifically, it was aimed at examining the extent to which greater magnitudes of Stroop interference in healthy older adults (as compared to their younger counterparts) selectively result from the amplified semantic conflict, response conflict or whether healthy aging actually amplifies both conflicts.

If indeed healthy aging causes an impaired suppression of the word-dimension (see Spieler et al.'s reasoning above), healthy older adults should produce disproportionately greater amounts of both semantic and response conflicts. This pattern of results should be evidenced by respectively greater semantic (SKY_{green}–PUT_{green}) *and* standard ($BLUE_{\text{green}}$–PUT_{green}) Stroop interference.

This prediction contrasts, however, with a largely overlooked work of Li and Bosman (1996). Their healthy older (vs. younger) participants showed significantly greater magnitudes of standard ($BLUE_{\text{green}}$ – $****_{\text{green}}$) but not of semantic Stroop interference (SKY_{green} – $****_{\text{green}}$) that remained significant and of the same magnitude in both young and healthy older participants. This pattern of results suggests that age-related differences in Stroop interference result from differences in response conflict but not in sematic conflict.

However, this latter conclusion –implying that healthy older adults are *not* less efficient in suppressing the word-dimension of Stroop words– remains tentative because magnitudes of both semantic and standard interference were artificially inflated by the use of a non-word color-neutral baseline (i.e.,$****_{\text{green}}$). Therefore, the experiment reported below was aimed at replicating these latter results (or alternatively those in line with Spieler et al., 1996) while using the aforementioned semantic Stroop paradigm that uses color-neutral words as a baseline.

Processes underlying age-related differences in Stroop interference

As mentioned above, the possibility that the locus of age-related differences in Stroop interference is specifically situated at the level of response and not at the level of semantic conflict (Li & Bosman, 1996) has important implications for our current understanding of processes underlying age-related differences in Stroop interference.

Indeed, this result would directly demonstrate that there is no age-related deficit in suppression of the irrelevant word-meaning. This latter idea remains in line with the fact that processes involved in word-recognition (processing from visual features up to semantics) in healthy older adults are at least as efficient as in younger adults (e.g., Lien et al., 2006 for more efficiency in older adults).

Consequently, greater magnitudes of Stroop interference in healthy older adults (as compared to their younger counterparts) would result from less efficient suppression of an irrelevant (pre-)response activity that is activated by the word-dimension of Stroop words (i.e., response conflict). But also, and/or alternatively, these differences may result from greater difficulty of healthy older adults to focus their attention on the relevant color-dimension of Stroop words.

Jackson and Balota (2013) were the first to raise just this issue. To this end, they manipulated the time that elapses between the individual's response and the presentation of a new stimulus on a computer-screen (i.e., response-stimulus interval, hereafter RSI). In line with the idea that short RSI induces a more consistent focus on the relevant color dimension (e.g., De Jong, Berendsen, & Cools, 1999; Parris, 2014), the magnitude of the Stroop effect (e.g., $BLUE_{\text{green}}$–$BLUE_{\text{blue}}$) was reduced at the short RSI compared with the long RSI in both younger and older adults (resulting in Congruency × RSI interaction). Said differently, older adults still displayed significantly greater Stroop effect than younger adults (resulting in Congruency × Age-group interaction), but there was no evidence of an interaction between RSI and Age-group (such that the overall Congruency × RSI × Age-group interactions also remained non-significant).

This pattern of results led Jackson and Balota to conclude that the decline in ability to maintain consistent focus on the relevant color dimension is not the primary mechanism underlying age-related differences in Stroop interference. As can be seen, this conclusion remains silent with respect to the exact role of an age-related deficit in suppressing the word-dimension of Stroop words (suppression of their meaning vs. response primed by the fully processed meaning, see above). Indeed, any such reasoning would be tentative given that Stroop effect (e.g., $BLUE_{\text{green}}$–$BLUE_{\text{blue}}$) measured in this study confounds both interference (resulting from both semantic and response conflicts) and facilitation (MacLeod, 1991). Therefore, the second goal of this paper was to examine this yet unanswered issue of processes underlying age-related differences in Stroop interference directly.

To this end, the aforementioned semantic Stroop paradigm was combined with a single-letter coloring and spatial cueing (SLCC) procedure. This procedure –first introduced by

Manwell et al. (2004)– consists of using the small arrows to spatially pre-cue the position(s) that will be subsequently occupied by the target letter(s) – the color of which is to be named. This single letter (in SLCC, as opposed to all letters) is then presented in an incongruent color (e.g., green "E" in the word "*BLUE*" with "B", "L" and "U" presented in another incongruent color from the response set).

SLCC (as opposed to all letters colored and cued, hereafter ALCC) is likely to improve selection-for-action by separating perceptually the two dimensions of color-incongruent Stroop words (e.g., Augustinova et al., 2010; Augustinova et al., 2015; Manwell et al., 2004's Account 2). Indeed, the additional (i.e., unpublished) evidence from Augustinova et al. (2015), suggests that independently of Stroop words' color-incongruency, SLCC elicited greater negativity (from 185 to 245 ms) at occipito-parietal sites than ALCC. Even though this ERP evidence cannot be readily equated with cognitive processes, it is consistent with the idea that SLCC produces an early shift in attentional focus toward the relevant color dimension (e.g., Besner & Stolz, 1999).[1]

It is thus not surprising that this procedure is known to significantly reduce Stroop interference (e.g., see Küper & Heil, 2012 for a review). For instance, in the aforementioned study of Augustinova et al. (2015), SLCC (*vs.* ALCC) reduced standard but *not* semantic Stroop interference as evidenced by speeded latencies on standard color-incongruent items. Therefore, SLCC (vs. ALCC) reduced standard (i.e., overall) Stroop interference via reduction of the response but not of semantic conflict. Independently of coloring and cuing, semantic conflict remained significant and of the same magnitude (as also evidenced by corresponding amplitudes of N400-like).

A joint consideration of these behavioral data along with ERP evidence mentioned above suggests that the benefit in (late) response processing (i.e., reduced response conflict) is still associated with changes in early processing (from 185 to 245 ms). Therefore, this manipulation was expected to shed some additional light on processes underlying greater magnitudes of Stroop interference in older (as compared to younger) adults. Building on Jackson and Balota, we do not a priori expect greater Stroop interference in older adults to result specifically from the difficulty to focus their attention on the relevant color dimension of Stroop words. Therefore, as in their study, Coloring × Age-Group interaction should remain non-significant and the a priori expected Interference-Type × Coloring along with Interference-Type × Age-group interactions should *not* be included in the overall Interference-Type × Coloring × Age-Group interaction.

To sum up, the study reported below was aimed at examining both the locus and processes underlying age-related differences in Stroop interference. To this end, the semantic Stroop paradigm combined with a single-letter coloring and spatial cueing (SLCC) procedure was administered to a sample of younger (18–26 years old) and healthy older (72–97 years old) adults. Given that some predictions tested in this study imply null effects, the usual frequentist statistical approach was extended to include a Bayesian approach which makes it possible to quantify the evidence in favor of the alternative and of the null hypotheses (e.g., Robidoux & Besner, 2015).

Method

Participants

Twenty-nine younger (17 females and 12 males; M_{age}=20.92) and 29 healthy older (22 females and 7 males; M_{age}=79.07) native French-speakers reporting normal or corrected-to-normal vision volunteered to take part in this experiment (see Table 1A in Supplementary Materials for further demographic and psychometric information).

Design and stimuli

The data were collected using a 3 (Stimulus-Type: incongruent vs. color-associated vs. neutral) × 2 (Coloring: ALCC vs. SLCC) × 2 (Age-Group: younger vs. older adults) design, with the first two factors being used as within-participants factors. There were 30 trials in each condition of these latter factors.

The stimuli (presented in lowercase Courier font, size 18, on a black background subtending an average visual angle of 0.9° high × 3.0° wide) consisted of four color words: *rouge* [red], *jaune* [yellow], *bleu* [blue], and *vert* [green]; four color-associated words: *tomate* [tomato], *maïs* [corn], *ciel* [sky], and *salade* [salad]; and four color-neutral words: *balcon* [balcony], *robe* [dress], *pont* [bridge] and *chien* [dog]. In each condition, all the stimuli were similar in length and frequency. In half the trials (i.e., ALCC-condition), the entire letter string appeared in incongruent target colors. In the other half of the trials (i.e., SLCC-condition), a single letter appeared in one incongruent target color, with the remaining letters appearing in another incongruent color from the response set.

Apparatus and procedure

The participants first completed the computerized version of the Stroop task described above. DMDX software (Forster & Forster, 2003) running on a portable PC (Dell Precision) was used for stimulus presentation and data collection. The

---

[1] Note that this pattern of ERP results is also consistent with the idea that spatial attention is a necessary preliminary to lexical processing of words (e.g., see Besner et al., 2016 for a review).

participants were seated approximately 50 cm from a 17-in. screen. They were asked to name the color of a letter indicated by small arrows as quickly and accurately as possible while ignoring everything else in the display. To this end, and as in Augustinova et al. (2010, Experiment 2), the participants were instructed to concentrate on the fixation cross ("+") that appeared for 500 ms in the center of the screen at the beginning of each trial. This was then replaced by small arrows (height of 1.2° of visual angle) displayed 0.6° above and below the position previously occupied by the fixation cross and subsequently occupied by the target letter. The arrows (i.e., spatial cues) remained on screen for 150 ms, after which the stimulus was displayed. In order to control for the letter-position effects in SLCC (e.g., Parris, Sharma, & Weekes, 2007), the spatially pre-cued letter was randomly located at the initial, middle, final, or the optimal viewing position. Because of this latter variation, the stimuli shifted horizontally from trial to trial in such a way that the spatially pre-cued letter always appeared in the same location as the preceding central fixation cue. The stimulus continued to be displayed until the participant responded or until 2,000 ms had elapsed. The participants' responses were recorded via a Koss 70-dB microphone headset and stored on the hard drive.

After completing 32 practice trials consisting of strings of asterisks (presented in the four colors described above), the participants performed the experimental task, which consisted of a single block of 180 experimental trials (see above). After a break, they completed a psychometric test battery designed to assess various aspects of psychological functioning (see Table 1A in Supplementary Material).

Results and discussion

Latencies greater than 3 *SD*s above or below each participant's mean latency for each condition (1.24% of the total data) were excluded from the analyses. Given the important general slowing in older (compared to younger) adults, the analyses of mean correct latencies (see Tables 1 and 2 and Supplementary Materials) cannot be meaningfully interpreted (e.g., Faust, Balota, Spieler, & Ferraro, 1999). To control for this issue, these latencies were transformed into percentages of standard ($[(\mathrm{M}_{\text{standard color-incongruent}}\ \mathrm{RT}–\mathrm{M}_{\text{color-neutral}}\ \mathrm{RT})/\mathrm{M}_{\text{color-neutral}}\ \mathrm{RT}]*100$) and semantic Stroop interference ($[(\mathrm{M}_{\text{color-associated incongruent}}\ \mathrm{RT}–\mathrm{M}_{\text{color-neutral}}\ \mathrm{RT})/\mathrm{M}_{\text{color-neutral}}\ \mathrm{RT}]*100$) as in Li and Bosman's data (1996). These percentages of standard and semantic Stroop interference observed in both Coloring conditions were subsequently analyzed using both traditional frequentist and Bayesian analyses.

The Bayes factor (BF) corresponding to the Bayesian probability of occurrence of a hypothesis (H1) and the null hypothesis (H0) was calculated in JASP (JASP Team, 2017). All Bayesian analyses were conducted on a BF_{10} comparison (i.e., H1 was compared with H0). All priors were equal. The BF for each parameter was estimated as a ratio of the likelihood of the model including the parameter and the likelihood of the model excluding it, and Jeffreys' (1961) classification was used to interpret the ensuing results as representing anecdotal, moderate or strong evidence.

Interference-Type × Coloring × Age-Group mixed ANOVA revealed main effects of Interference-Type [$F(1,56)=128.01$; $p<.001$, $\eta_p^2=.70$], Coloring [$F(1,56)=29.00$; $p<.001$, $\eta_p^2=.34$] and Age-Group [$F(1,56)=8.46$; $p<.001$, $\eta_p^2=.13$]. It further revealed a significant Interference-Type × Coloring [$F(1,56)=27.25$; $p<.001$, $\eta_p^2=.33$] interaction and a marginally significant Interference-Type × Age-Group [$F(1,56)=3.78$; $p=.057$, $\eta_p^2=.06$] interaction. The Coloring × Age-Group [$F(1,56)=1.85$; $p=.179$, $\eta_p^2=.03$] and the overall Interference-Type × Coloring × Age-Group interactions were not significant [$F(1,56)=.70$; $p=.407$, $\eta_p^2=.01$].

As indicated by a marginally significant Interference-Type × Age-Group interaction (see Table 1) on percentages of interference, age-related differences in the processing of Stroop words persisted even after controlling for general slowing.

The decomposition of this latter interaction showed that the simple main effect of Age-Group was significant for the magnitude of standard Stroop interference [$F(1,56)=9.45$; $p=.003$, $\eta_p^2=.14$]. It was significantly greater in the older ($M=12.08$; $SD=.90$; $CI[10.29, 13.87]$) than in the younger adults ($M=8.18$; $SD=.90$; $CI[6.38, 9.97]$). Conversely, the simple main effect of Age-Group did not reach the conventional level of significance for the magnitude of semantic Stroop interference [$F(1,56)=1.80$; $p=.185$, $\eta_p^2=.03$]. It was of a similar magnitude in both the older ($M=3.08$; $SD=.67$; $CI[1.72, 4.42]$) and younger ($M=1.82$; $SD=.67$; $CI[.48, 3.15]$) participants. As far as the magnitude of standard Stroop interference is concerned, Bayes factor (hereafter BF) provided moderate (between 3 and 10) evidence for the Age-Group effect hypothesis ($\mathrm{BF}_{10}=3.384$). Additionally, $\mathrm{BF}_{10}=0.439$ provided anecdotal (between 1/3-1) evidence for the null hypothesis of no Age-Group effect on semantic Stroop interference.

In line with Li and Bosman's past findings (1996), the aforementioned results seem to place the locus of the well-established age-related differences in Stroop interference at the level of response conflict and not at the level of semantic conflict (or at level of both conflicts, Spieler et al., 1996).

To address further processes underlying these age-related differences, a significant Interference-Type × Coloring interaction (see Table 2) was first decomposed.

This decomposition revealed that the simple main effect of Coloring was significant for the magnitude of standard Stroop interference [$F(1,56)=45.66$; $p<.001$, $\eta_p^2=.45$]. It was considerably reduced by SLCC ($M=6.26$; $SD=.62$; $CI[5.01, 7.51]$) as compared to ALCC (i.e., standard coloring condition;

Table 1 Color-naming performance observed as a function of stimulus- or interference-type and age-group

	Age group							
	Younger participants			Older participants			Age effect (RT)	Age effect (%ER)
	M (SE)	CI	%ER	M (SE)	CI	%ER		
Standard incongruent $BLUE_{red}$	854 (28)	[798, 910]	.95	1023 (28)	[967, 1079]	5.12	+169**	+4.18*
Color-associated incongruent SKY_{red}	804 (25)	[754, 854]	.41	939 (25)	[889, 989]	.70	+135**	$+.29^{ns}$
Color-neutral PUT_{red}	791 (23)	[744, 838]	.25	911 (23)	[864, 958]	.49	+120**	$+.25^{ns}$
	M (SE)	CI		M (SE)	CI		Age effect (% interf)	
% of standard Stroop interference	8.18 (.90)	[6.38, 9.97]	<	12.08 (.90)	[10.29, 13.87]		+3.90**	
% of semantic Stroop interference	1.82 (.67)	[.48, 3.15]	≈	3.08 (.67)	[1.75, 4.42]		$+1.26^{ns}$	

Note. Percentages of interference (controlling for age-related difference in processing speed) were calculated automatically using unrounded response times

* significant at $p < .05$

** significant at $p < .01$

M=13.99; *SD*=1.04; *CI*[11.92, 16.06]). SLCC-procedure left the semantic Stroop interference unaffected [*F*(1,56)=.71; *p*=.402, η_p^2=.01] – it was comparable in both the ALCC (*M*=2.83; *SD*=.59; *CI*[1.65, 4.02]) and SLCC (*M*=2.07; *SD*=.71; *CI*[.64, 3.50]) conditions. Additionally, BF_{10}>300 provided decisive evidence for the Coloring-effect hypothesis on standard Stroop interference and BF_{10}=0.293 moderate evidence (i.e., between 1/10-1/3) for the null hypothesis of no Coloring-effect on semantic Stroop interference.

The absence of Coloring × Age-Group and Interference-Type × Coloring × Age-Group interactions suggests that the age-related differences in response conflict reported above remain constant under both ALCC and SLCC. Yet, the results of further planned comparisons of the Age-Group effects on the standard Stroop interference nuance this latter idea. Indeed, the age-related differences in standard Stroop interference were significant in the ALCC-condition [*F*(1,56)=7.15; *p*=.010, η_p^2 =.11] but – due to the SLCC benefit described above – only marginally significant in the SLCC-condition [*F*(1,56)=3.31; *p*=.074, η_p^2=.056]. The BF evidence for the Age-effect hypothesis on standard Stroop interference was moderate in ALCC (BF_{10}=4.801) but anecdotal in SLCC (BF_{10}=1.045).

These additional analyses also suggest that both young and older participants are equally and highly responsive to SLCC and if there are any age-related differences in terms

Table 2 Color-naming performance observed as a function of stimulus- or interference-type and coloring

	Coloring							
	All letters colored (ALC)			Single letter colored (SLC)			SLC effect (RT)	SLC effect (ER)
	M (SE)	CI	%ER	M (SE)	CI	%ER		
Standard incongruent $BLUE_{red}$	963 (23)	[917, 1010]	3.82	913 (18)	[877, 950]	2.25	-50**	-1.58*
Color-associated incongruent SKY_{red}	866 (18)	[830, 903]	.46	877 (18)	[841, 913]	.65	$+11^{ns}$	$-.19^{ns}$
Color-neutral PUT_{red}	843 (18)	[808, 878]	.19	859 (16)	[827, 892]	.55	+17*	$-.36^{ns}$
	M (SE)	CI		M (SE)	CI		SLC effect (% interference)	
% of standard Stroop interference	13.99 (1.04)	[11.92, 16.06]	>	6.26 (.62)	[5.01, 7.51]		-7.73**	
% of semantic Stroop interference	2.83 (.59)	[1.65, 4.02]	≈	2.07 (.71)	[.64, 3.50]		$-.77^{ns}$	

Note. Percentages of interference (controlling for age-related difference in processing speed) were calculated automatically using unrounded response times (RTs)

* significant at $p < .05$

** significant at $p < .01$

of the benefit from this procedure, then it is the older participants who potentially benefit from this procedure more than younger ones. The results of planned comparisons of the SLCC-effect on standard Stroop interference seem generally consistent with this latter idea. Indeed, this effect was significant in both samples but somewhat smaller in the younger [$F(1,56)=14.22$; $p<.001$, $\eta_p{}^2=.20$] than in the older adults [$F(1,56)=33.48$; $p<.001$, $\eta_p{}^2=.37$] (see Table 1). Despite this, the BF evidence for the Coloring-effect hypothesis on standard Stroop interference was decisive in both samples: younger ($BF_{10}>300$) and older ($BF_{10}>300$) adults.

Taken together, these different results are thus in line with Jackson and Balota (2013) suggesting that greater magnitudes of standard Stroop interference in older adults are unlikely to result from an impaired ability to focus their attention on the relevant color dimension of Stroop words. Thus, planned comparisons of the Age-Group and Coloring-effects on semantic Stroop interference were conducted to examine further the extent to which these greater magnitudes can be attributed to an age-related deficit in ignoring the irrelevant word-dimension of Stroop words.

The planned age-group effects were non-significant in both the ALCC [$F(1,56) = 2.27$; $p = .137$, $\eta p^2 = .039$] and SLCC conditions [$F(1,56) = .27$; $p = .603$, $\eta p^2 = .005$]. This conclusion is supported by BF evidence that was anecdotal ($BF_{10} =0.685$) in ALCC and moderate ($BF_{10} =0.298$) in SLCC for the null hypothesis of no age-effect on semantic interference. Similarly, planned comparisons of Coloring-effects on this interference showed no significant reduction either in the younger [$F(1,56)=.04$; $p=.845$, $\eta_p{}^2=.001$] or the older [$F(1,56)=.997$; $p=.322$, $\eta_p{}^2=.017$] participants. BF provided respectively moderate evidence ($BF_{10}=0.272$) in younger and anecdotal evidence ($BF_{10}=0.419$) in older adults for the null hypothesis of no Coloring-effect on semantic conflict.

These latter results along with those presented above suggest that greater magnitudes of standard Stroop interference in older (as compared to younger) adults are solely due to less efficient suppression of an irrelevant response that is activated by the word-dimension of Stroop words (i.e., response conflict) and not to less efficient suppression of the irrelevant meaning of these words (i.e., semantic conflict).

General discussion and conclusions

The direct empirical evidence reported in this paper runs counter to consensually held views on both the locus and processes underlying age-related differences in Stroop interference that were inferred from mere changes in magnitudes of standard (i.e., overall) Stroop effect/interference (e.g., see Li & Bosman, 1996 for an only exception). Indeed, it clearly places the locus of these differences at the level of response as opposed to semantic conflict (Li & Bosman, 1996) or as opposed to both conflicts (Spieler et al., 1996).

With respect to processes underlying these age-related differences, the empirical evidence reported above shows that both younger and older adults are equally efficient in their focus on the relevant color dimension (Jackson & Balota, 2013; e.g., see also Ruthruff & Lien, 2016) and equally (in)efficient in suppressing the meaning of the irrelevant word-meaning of Stroop words. Therefore, this direct evidence also has several important implications for a more general issue of how selective attention actually operates in the Stroop task.

Selective attention is "the ability to focus on one thing [e.g., to attend to and process a relevant color-dimension of Stroop words] while ignoring other things [e.g., an irrelevant word-dimension] excluding to-be-ignored information from deeper processing and control over action." (Ruthruff & Lien, 2016, p.3, text in brackets added, see also Spieler and colleagues' reasoning in the Introduction section).

The dissociative pattern of both Age- and SLCC-effects on semantic vs. standard Stroop interference suggests that (1) an amplitude of response conflict is independent of people's efficiency insuppressing the word-dimension of Stroop words mainly because (2) this suppression is itself unlikely. Indeed, (3) the to-be-ignored information in the Stroop task is never excluded from deeper processing (as suggested by constantly significant semantic Stroop interference, the magnitude of which remained unchanged by Aging and SLCC). Rather, this to-be-ignored information (4) is more or less efficiently excluded from control over action (as suggested by standard Stroop interference involving the response conflict, the magnitude of which varied as a function of Aging and SLCC). As emphasized by multiple-stage models of Stroop interference, (5) the exclusion of to-be-ignored information from deeper processing and its exclusion from control over action constitute two different classes of processes. Historically favored single-stage response accounts –in which the customary implementations of Stroop inference/effect ($BLUE_{\text{green}}$–$DEAL_{\text{green}}$/$BLUE_{\text{blue}}$) are rooted– are therefore likely to be obsolete. These two classes of processes (6) are likely to be different in nature. Those involved in the control of the irrelevant semantic information over (response-related) action seem clearly controllable and thus evolve with aging and are permeable to moderators (e.g., SLCC, RSI). Even though we are inclined to conclude that processes involved in the semantic (i.e., conceptual) processing of Stroop words are automatic –therefore preserved in healthy aging and not permeable to moderators (Augustinova & Ferrand, 2014), this conclusion would still remain unwarranted in light of other related studies (e.g., Labuschagne & Besner, 2015; see Besner et al., 2016 for a review). Thus, future research needs to address this issue directly.

Meanwhile, perhaps the most immediate conclusion to be drawn from the present study is that the processes involved in selective attention as well as their modulation by different variables might remain unseen and/or be misinterpreted when observed using the standard Stroop paradigm.

References

Augustinova, M., & Ferrand, L. (2014). Automaticity of word reading: Evidence from the semantic Stroop paradigm. *Current Directions in Psychological Science, 23*, 343–348.

Augustinova, M., Flaudias, V., & Ferrand, L. (2010). Single-letter coloring and spatial cueing do not eliminate nor reduce a semantic contribution to the Stroop effect. *Psychonomic Bulletin & Review, 17*, 827-833.

Augustinova, M., Silvert, L., Ferrand, L., Llorca, P. M., & Flaudias, V. (2015). Behavioral and electrophysiological investigation of semantic and response conflict in the Stroop task. *Psychonomic Bulletin & Review,22,* 543-549.

Augustinova, M., Silvert, L., Spatola, N., & Ferrand, L. (2017). Further investigation of distinct components of Stroop interference and of their reduction by short response-stimulus intervals. *Acta Psychologica*, https://doi.org/10.1016/j.actpsy.2017.03.009.

Besner, D., Risko, E. F., Stolz, J. A., White, D., Reynolds, M., O'Malley, S., & Robidoux, S. (2016). Varieties of attention: Their roles in visual word identification. *Current Directions in Psychological Science, 25*, 162-168.

Besner, D., & Stolz, J. A. (1999). What kind of attention modulates the Stroop effect?. *Psychonomic Bulletin & Review,* 6, 99-104.

Comalli, P. E., Jr., Wapner, S., & Werner, H. (1962). Interference effects of Stroop colour-word test in childhood, adulthood, and ageing. *Journal of Genetic Psychology, 100*, 47–53.

De Jong, R., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations: Dissociable causes of interference effects in conflict situations. *Acta Psychologica, 101*, 379–394.

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin, 125*, 777-799.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116–124.

Hasshim, N., & Parris, B. A. (2014). Two-to-one color-response mapping and the presence of semantic conflict in the Stroop task. *Frontiers in Psychology, 5*.

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.

Jeffreys, H. (1961). *The theory of probability* (3rd ed.). Oxford, England: Oxford University Press.

Klein, G. S. (1964). Semantic power measured through the effect of words with color-naming. *American Journal of Psychology, 77*, 576–588.

Küper, K. & Heil, M. (2012). Attentional focus manipulations affect naming latencies of neutral but not of incongruent Stroop trials. *Swiss Journal of Psychology,71*, 93–100.

Labuschagne, E. M., & Besner, D. (2015). Automaticity revisited: When print doesn't activate semantics. *Frontiers in Psychology, 6*.

Li, K. Z., & Bosman, E. A. (1996). Age differences in Stroop-like interference as a function of semantic relatedness. *Aging, Neuropsychology, and Cognition ,3*, 272-284.

Lien, M.-C., Allen, P. A., Ruthruff, E., Grabbe, J., McCann, R. S., & Remington, R. W. (2006). Visual word recognition without central attention: Evidence for greater automaticity with advancing age. *Psychology and Aging, 21,* 431-447.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163-203.

Manwell, L. A., Roberts, M. A., & Besner, D. (2004). Single letter coloring and spatial cueing eliminates a semantic contribution to the Stroop effect. *Psychonomic Bulletin and Review*, *11*, 458–462.

Neely, J. H., & Kahan, T. (2001). Is semantic activation automatic? A critical re-evaluation. In Roediger, H. L., Nairne, J. S., Neath, I., and Surprenant, A. M. (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 69-93). Washington, DC: American Psychological Association.

Parris, B. A. (2014). Task conflict in the Stroop task: When Stroop interference decreases as Stroop facilitation increases in a low task conflict context. *Frontiers in Psychology*, 5,1182, https://doi.org/10.3389/fpsyg.2014.01182.

Parris, B. A., Sharma, D., & Weekes, B. (2007). An optimal viewing position effect in the Stroop task when only one letter is the color carrier. *Experimental Psychology, 54*, 273-280

Risko, E. F., Schmidt, J. R., & Besner, D. (2006). Filling a gap in the semantic gradient: Color associates and response set in the Stroop task. *Psychonomic Bulletin & Review, 13*, 310-315.

Robidoux, S., & Besner, D. (2015). Conflict resolved: On the role of spatial attention in reading and color naming tasks. *Psychonomic Bulletin & Review, 22,* 1709-1716.

Ruthruff, E., & Lien, M.-C. (2016). Aging and attention. In Pachana, N. A. (Ed.), *Encyclopedia of geropsychology*. Springer: New York.

Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the Stroop task. *Canadian Journal of Experimental Psychology, 59*, 132-138.

Seymour, P. H. (1977). Conceptual encoding and locus of the Stroop effect. *Quaterly Journal of Experimental Psychology, 29*, 245-265.

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.

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

White, D., Risko, E. F., & Besner, D. (2016). The semantic Stroop effect: An ex-Gaussian analysis. *Psychonomic Bulletin & Review, 23*, 1576-1581.