

## A Sensory Origin for Color-Word Stroop Effects in Aging: Simulating Age-Related Changes in Color-vision Mimics Age-Related Changes in Stroop

Boaz M. Ben-David & Bruce A. Schneider

To cite this article: Boaz M. Ben-David & Bruce A. Schneider (2010) A Sensory Origin for Color-Word Stroop Effects in Aging: Simulating Age-Related Changes in Color-vision Mimics Age-Related Changes in Stroop, *Aging, Neuropsychology, and Cognition*, 17:6, 730-746, DOI: [10.1080/13825585.2010.510553](https://doi.org/10.1080/13825585.2010.510553)

To link to this article: <https://doi.org/10.1080/13825585.2010.510553>



Published online: 05 Nov 2010.



Submit your article to this journal [↗](#)



Article views: 693



Citing articles: 26 View citing articles [↗](#)

# A Sensory Origin for Color-Word Stroop Effects in Aging: Simulating Age-Related Changes in Color-vision Mimics Age-Related Changes in Stroop

BOAZ M. BEN-DAVID<sup>1,2,3</sup> AND BRUCE A. SCHNEIDER<sup>1</sup>

<sup>1</sup>Department of Psychology, Centre for Research on Biological Communication Systems, University of Toronto Mississauga, Ontario, Canada, <sup>2</sup>Department of Speech-Language Pathology, Oral Dynamics Lab, University of Toronto, Ontario, Canada, <sup>3</sup>Research, Toronto Rehabilitation Institute, Ontario, Canada

## ABSTRACT

An increase in Stroop effects with age can be interpreted as reflecting age-related reductions in selective attention, cognitive slowing, or color-vision. In the present study, 88 younger adults performed a Stroop test with two color-sets, saturated and desaturated, to simulate an age-related decrease in color perception. This color manipulation with younger adults was sufficient to lead to an increase in Stroop effects that mimics age-effects. We conclude that age-related changes in color perception can contribute to the differences in Stroop effects observed in aging. Finally, we suggest that the clinical applications of Stroop take this factor into account.

**Keywords:** Cognitive aging; Stroop; Selective-attention; Sensory aging; Speed of processing.

## INTRODUCTION

The need for selective attention in everyday life is ubiquitous. For example, when driving a car, one must attend to and process road signs, such as a stop sign, without interference from visual distractions such as billboards. Age-related declines in selective attention could potentially explain why older adults' performance deteriorates in a cluttered visual or auditory environment, whether it is spotting a stop sign in a busy intersection, or listening to a conversation in a noisy park.

---

Address correspondence to: Boaz M. Ben-David, Department of Speech-Language Pathology, Oral Dynamics Lab, University of Toronto, 160-500 University Avenue, Toronto, Ontario, Canada M5G 1V7. E-mail: boaz.ben.david@utoronto.ca

---

One of the most common laboratory and clinical measures of age-related changes in selective attention is the card-version of the color-word Stroop test (Golden, 1978). In this paradigm people are asked to: (1) read aloud words printed in a color-neutral context (Rn, e.g., BLUE printed white on black), (2) name the font color of a word unrelated to color or the color of a row of Xs (Cn, e.g., XXXX printed in blue), and (3) name the font color of a color-word that is incongruent with its print color (Ci, e.g., RED printed in blue). If participants can focus exclusively on the font color of the words, there should be no difference in reaction times (RT) for color-naming between Ci and Cn presentations. However, the typical result is that the RT for Ci presentations is significantly larger than it is for Cn presentations. This difference in RT is termed Stroop Interference (SI),

$$SI = RT(C_i) - RT(C_n). \quad (1)$$

The value of SI is then used to index the degree to which there is a failure of selective attention. The higher degree of SI that is found in older adults as compared to that found in younger adults is often interpreted as an age-related loss in selective attention (e.g., McDowd & Shaw, 2000; Troyer, Leach, & Strauss, 2006).

This hypothesized age-related loss in selective attention is often attributed to a decrease in the ability of older adults to inhibit irrelevant sources of information (Hartman & Hasher, 1991), which is consistent with Hasher and Zacks' (1988) theory that there is a decrease in the efficiency of inhibitory processes with aging. However, age-related changes in Stroop performance have also been attributed to changes in the speed of processing (Verhaeghen & Cerella, 2002) or in the sensory visual input (Ben-David & Schneider, 2009). The goal of this study was to see if we could produce in younger adults Stroop effects that are characteristic of older adults, by reducing the amount of color information available, thereby mimicking the losses in color information experienced by older adults.

### **Cognitive Slowing**

Several authors have suggested that age-related increases in SI do not reflect a change in selective attention but merely a generalized slowing in the speed of processing with age that is uniform across tasks. According to this view, the RTs on a task varies with task difficulty (TD) in a multiplicative fashion, that is,

$$RT[AGE,TD] = TD * f(AGE), \quad (2)$$

where  $f(\text{AGE})$  is the generalized slowing function that characterizes the effects of age on all tasks (Cerella & Hale, 1994; see also, Cerella, 1990). Taking Eq. (1) and (2) together, it is easy to see that an increase in SI with age is predicted from cognitive slowing theories (given that  $SI_{\text{young}} > 0$ ),

$$\begin{aligned} SI(\text{old}) &= RT[\text{old}, TD(C_i)] - RT[\text{old}, TD(C_n)] \\ &= [TD(C_i) - TD(C_n)]f(\text{old}) \\ SI(\text{young}) &= RT[\text{young}, TD(C_i)] - RT[\text{young}, TD(C_n)] \\ &= [TD(C_i) - TD(C_n)]f(\text{young}). \end{aligned} \quad (3)$$

$$f(\text{old}) > f(\text{young}) \rightarrow SI(\text{old}) > SI(\text{young})$$

### Sensory Degradation

Age-related increases in Stroop interference have also been attributed to age-related changes in color-vision (Ben-David & Schneider, 2009). Color-vision has been found to change rapidly after the age of 60 (Werner & Steele, 1988), which would theoretically make it more difficult to name font colors. Note that age changes in color-vision should not affect the ability to read a word, provided that the contrast and visual angle subtended by the word are sufficiently large, as Akutsu, Legge, Ross, and Schuebel (1991) have found. Thus, an age-related sensory decline in color-vision predicts an increase in the Dimensional Imbalance (DI) between the speed for color-naming a font, and for reading a word where

$$DI = RT(C_n) - RT(R_n). \quad (4)$$

Indeed, there is ample evidence that color-naming slows down with age faster than reading (e.g., Mutter, Naylor, & Patterson, 2005; Salthouse & Mein, 1995; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2006), indicating an age-related change in DI.

### *The Link between DI and Stroop Effects*

If lexical access and access to the font color are independent processes (see Lindsay & Jacobi, 1994), we would expect that it would take the same amount of time to access the font color of an incongruent word as it takes to access the font color of a non-word stimulus. In C<sub>n</sub> (e.g., naming the color of a row of Xs), the latency presumably reflects color-naming processes only. However, in C<sub>i</sub> (e.g., naming the color of the word RED printed in blue), we might expect the lexical response to interfere with the color-naming response, since lexical access is faster than access to the font color (see Posner & Snyder, 1975). Therefore, C<sub>i</sub> latencies depend not only on accessing the name of the font color, but also on the degree of interference from lexical

processing. The increase in DI with age primarily reflects an age-related decrease in the ease of access to the name of the font color (as indicated by the small age-related increase in latencies for Rn as compared to the larger age-related increase in Cn, see Ben-David & Schneider, 2009). Older adults would, therefore, have to inhibit the lexical response for a longer period of time than would younger adults until access was gained to the name of the font color. Hence, an increase in DI should lead to an increase in SI, as found by Melara and Algom (2003) and Ben-David and Schneider (2009) in separate meta-analyses.

Of course, the effects of sensory losses on cognitive processes may not be limited to increasing the period of time required to access the name of a color. Sensory processes have also been shown to affect other cognitive functions such as lexical access. For instance, Frost, Ahisar, Gotesman, and Tayeb (2003) showed that manipulating sensory intensity of printed words affected their phonological processing. Indeed, several authors have argued that losses of sensory information could have widespread effects on cognitive function (for a recent review, see Schneider, Pichora-Fuller, & Daneman, 2010; see also Ben-David et al., 2010, for a recent example regarding losses of auditory information). However, it is reasonable to expect that the cognitive functions (such as color-naming) that are directly related to the sensory loss (color-vision) would be the most affected.

### **Contrasting the Three Theories for Age-Related Increase in SI**

In a preceding paper, Ben-David and Schneider (2009) contrasted all three theories in a *cross-lab* analysis of 13 studies (including 1,825 participants) that compared groups of younger and older adults on Stroop tasks, and a *cross-sectional* analysis of an extensive lifespan study that included 1,788 participants in 12 age categories (25 to 81 years old, Van der Elst et al., 2006). The cross-lab analysis was consistent with both age related slowing and a sensory degradation theory. Conversely, the cross-sectional data did not fit the cognitive slowing explanation of Stroop effects. Contrary to the predictions of Eq. (2), the rates at which latency increased with age differed substantially among the three tasks, with latency growth rates being slowest for Rn, faster for Cn, and fastest for Ci. Hence, based on this cross-sectional study, cognitive slowing was rejected as the sole explanation of age-related changes in Stroop effects.

### **The Current Study**

If the age-related changes in SI are related to sensory decline in color-vision, it should be possible to mimic such declines in young adults, by degrading the color cues provided by the stimuli, without hampering reading. Several studies have confirmed that direct experimental manipulations of DI lead to changes in Stroop effects. For example, Melara and his colleagues (Melara & Mounts, 1993; Sabri, Melara, & Algom, 2001) have decreased the

visual angle subtended by the stimuli (by decreasing their font size) to diminish their discriminability along the word dimension (see also Algom, Dekel, & Pansky, 1996). These studies have found that an experimentally-generated decrease in DI resulted in a decrease in Stroop effects. Recently, Eidels, Townsend, and Algom (2010) were able to nearly eliminate Stroop effects after equating the speed of classification along the color and word dimensions.

However, all of these studies focused on manipulating the word dimension, rather than the color dimension. In the present study, we attempted to mimic color perception deficiency typical of older adults by decreasing the saturation of the color-set used in a Stroop test.

When the color information in a stimulus is reduced, it is reasonable to expect that it would take longer to access the color name (when asked to do so) than when the color in the display is fully saturated. This would hold for word and non-word stimuli. However, provided that the contrast and visual angle subtended by a color-neutral word is unchanged, desaturating the color should not affect the ability to read the word. Hence, desaturation of font color should lead to an increase in the latencies for Cn but not for Rn, mimicking the age-related increase in DI. According to the sensory degradation hypothesis, this increase in DI should lead to an increase in SI (i.e., larger increase in latencies for Ci than for Cn).

Unlike most studies that manipulated DI with younger adults, we simulated the card version of the Stroop test. This involves presenting multiple words on the computer screen in a blocked design rather than the single word version, in which a single word is presented in each trial (see a discussion on the differences in effects between the two versions in Balota, Tse, Hutchinson, Spieler, Duchek, & Morris, in press; Perlstein, Carter, Barch, & Baird, 1998). A recent paper (Ludwig, Borella, Tettamanti, & Ribaupierre, in press) suggested that age-effects on Stroop are larger in blocked than in item-by-item designs. Hence we chose the blocked design because it is the one most extensively used in the clinic and because it has been shown to produce the largest age effects. If sensory factors (color-vision) are indeed contributing to age-related changes in Stroop effects, we expected to find an increase in both DI and SI (and Ci) as a result of the decrease in the saturation of the color-set. On the other hand, a failure to mimic the age-related changes in Stroop effects by changing the saturation of the colors used would suggest that age-related changes in color-vision cannot be responsible for the effects of aging on Stroop performance.

## **METHOD**

### **Participants**

Eighty-eight young adults ( $M = 19.6$  years old,  $SD = 1.6$ ), undergraduates at the University of Toronto Mississauga, participated in this study.

They received either course credit or were paid \$10/hour for their participation. All participants were native English speakers, as assessed by a self report, and had achieved a minimum score on the Mill–Hill Vocabulary Test of 9/20 (Raven, 1965), corresponding to normal vocabulary levels for native-English speakers ( $M = 13/20$ ,  $SD = 1.9/20$ ). All participants had a minimum Snellen fraction of 20/20 for binocular near vision ( $M = 13.9/20$ ,  $SD = 2.7/20$ ), and a minimum score of 7 out of 8 on the Ishihara color-blindness test (see Birch, 1997), reflecting visual acuity and color-vision within the normal range.

### Stimuli, Apparatus and Design

The following font colors were used: green, blue and red. Each session consisted of two saturation blocks: *high-saturation* color-set block, in which high-contrast color stimuli were employed in a Stroop paradigm, and a *low-saturation* color-set block, in which desaturated color stimuli were employed. The color specifications of these stimuli were measured using a high performance color meter attached to the monitor (ColorCal, Cambridge Research Systems). The color values [using CIELUV 1976 with  $L$  ( $\text{cd}/\text{m}^2$ ), e.g., De Corte, 1986] for the *high-saturation* color block were  $u' = 0.441$ ,  $v' = 0.992$ ,  $L = 11.07$  for red;  $u' = 0.167$ ,  $v' = 0.375$ ,  $L = 3.56$  for blue; and  $u' = 0.112$ ,  $v' = 0.251$ ,  $L = 9.02$  for green. For the *low-saturation* color-set block, in which colors were highly desaturated, the color values were  $u' = 0.202$ ,  $v' = 0.454$ ,  $L = 90.17$  for red;  $u' = 0.189$ ,  $v' = 0.425$ ,  $L = 91.17$  for blue; and  $u' = 0.186$ ,  $v' = 0.418$ ,  $L = 95.25$  for green. The desaturated set can be best described as ‘washed-out’ versions of the saturated colors. It is important to note that we are not claiming that the color desaturation manipulation used here precisely emulates the kind of losses in color-vision experienced by older adults. Age-related changes in color-vision are gradual, progressive and varied (e.g., Nguyen-Tri, Overbury, & Faubert, 2003; Werner & Steele, 1988), and affect some color discriminations more than others. Hence, the experience of a diminution of color-vision in older adults is likely to differ from a laboratory-induced sudden decrease in color contrast in younger adults. The manipulation used here is simply a convenient way of reducing the color information available to younger adults in order to make it more difficult for them to access color names, as the sensory-degradation hypothesis postulates is happening in older adults.

In each saturation block there were three tasks: Rn, Cn and Ci. Each task consisted of the presentation of three consecutive displays. Each of the three displays contained 36 colored stimuli in a six by six format (Arial, Bold, font 24) on the black background of a 17" monitor (for a total of 108 colored stimuli per task). For Rn, participants read aloud the color-neutral words SIX, FOUR and EIGHT. Note that in the Rn condition, we used color-neutral words printed in color. This was done to check whether the degree of desaturation of the font colors was having an effect on lexical

**TABLE 1.** The order of tasks (Rn, Cn and Ci) and saturation blocks (saturated, and desaturated) for the 8 experimental groups in this study

Group	Task					
	Block 1			Block 2		
	<i>Saturated</i>			<i>Desaturated</i>		
1a	Rn	Cn	Ci	Rn	Cn	Ci
2a	Rn	Cn	Ci	Rn	Ci	Cn
3a	Rn	Ci	Cn	Rn	Cn	Ci
4a	Rn	Ci	Cn	Rn	Ci	Cn
	<i>Desaturated</i>			<i>Saturated</i>		
1b	Rn	Cn	Ci	Rn	Cn	Ci
2b	Rn	Cn	Ci	Rn	Ci	Cn
3b	Rn	Ci	Cn	Rn	Cn	Ci
4b	Rn	Ci	Cn	Rn	Ci	Cn

The order of the tasks is denoted as group-orders 1, 2, 3 and 4, and the order of saturation blocks as group-orders a, and b.

access. If saturation was affecting lexical access, we would expect to find differences in Rn latency at the two levels of saturation. For Cn: participants named aloud the font color of strings of 5 Xs. For Ci: participants named aloud the font color of color-words (RED, GREEN, BLUE) printed in incongruent font colors (e.g., BLUE printed in red). The order of the stimuli within each slide was pseudo-randomized, ensuring that no word or font color repeated more than twice in a row. Half of the participants were tested in the three conditions with the saturated colors first and the desaturated colors second. The other half were tested in the reverse order. Rn was the first task in each saturation condition. To control for the well-documented effects of practice on color-naming latencies in the Stroop paradigm (e.g., Melara & Mounts, 1993) the order of color-naming tasks (Cn and Ci) was fully counterbalanced across the two saturation conditions. This counterbalancing, when combined with the two orders for the saturation conditions generated the 8 experimental groups (11 participants per group) shown in Table 1. Participants were tested individually in a dimly lit room, seated at a distance of 60 cm from the monitor. Oral responses were recorded by a PC computer via a desk mounted, omni-directional pickup microphone (Nexxtech), which was calibrated at the beginning of the session. Each task began with a short practice trial that consisted of 12 stimuli on a single slide.

## RESULTS

### Accuracy

Oral responses were digitally recorded and later their accuracy was coded. Generally, accuracy rates were very high (98.6% overall). Accuracy



was at peak for the reading task (99.7%), and was slightly reduced for Cn (98.8%), with the least accurate responses for Ci (97.1%). The accuracy data was submitted to a 2 × 3 repeated measures ANOVA with color-set (high- vs. low-saturation) and task type (Rn, Cn or Ci) as within participant factors. The analysis revealed a main effect for task type [ $F(2, 174) = 51.7, p < .001, \eta_p^2 = 0.37$ ]. There was no significant main effect for color-set [ $F(1, 87) = 2.6, p > .1$ ], nor a significant interaction between the two factors [ $F(2, 174) = 1.3, p > .1$ ]. Hence, the desaturation manipulation had no effect on accuracy. In *post-hoc* paired sample *t*-tests (Bonferroni corrected for the three possible comparisons among means), we confirmed that the ordering of accuracy was Rn > Cn > Ci when saturation conditions are collapsed [Cn–Rn,  $t(87) = 8.1, p < .001$ ; Ci–Cn,  $t(87) = 3.6, p < .01$ ; Ci–Rn,  $t(87) = 10.4, p < .001$ ].

**Latencies**

In each task, latencies were recorded for each of the three displays from the onset of the display until the last response had been terminated. The average latency on the three displays was then divided by 36 (the number of words on a display) to obtain an estimate of the average latency per word for that participant. Average latencies in the three Stroop tasks (Rn, Cn or Ci) for the two saturation conditions (high- and low-saturation) are presented in the three leftmost columns of Table 2. Across both saturation conditions there is an ordering of latency, with fastest average responses for Rn (439 ms), slower for Cn (582 ms) and slowest for Ci (801 ms). On average, across tasks, response latencies were faster in the high-saturation condition (576 vs. 638 ms for high- and low-saturation, respectively). We conducted a 2 × 3 repeated measures ANOVA, with color-set (high- vs. low-saturation) and task type (Rn, Cn or Ci) as within-participants factors, using latency as the dependent variable. The analysis revealed a main effect for color-set [ $F(1, 87) = 87.2, p < .001, \eta_p^2 = 0.50$ ], reflecting faster responses for the high-saturation

**TABLE 2.** Mean latencies in milliseconds and standard errors (in parentheses) for reading a neutral word, Rn, color-naming a string of Xs, Cn, and color-naming an incongruent color-word, Ci, averaged for 88 participants across 8 groups

	Rn	Cn	Ci	DI (Cn–Rn)	SI (Ci–Cn)	Ci–Rn
High-saturation	442 (9)	547 (13)	740 (17)	105 (9)	193 (9)	298 (13)
Low-saturation	435 (10)	617 (13)	861 (20)	182 (10)	244 (12)	426 (16)
Color effect (low – high)	–7	70**	121**	77**	51**	128**

Dimensional Imbalance (DI) is the difference between Cn and Rn latencies, and Stroop Interference (SI) is the difference between Ci and Cn latencies. In the rightmost column is the difference between Ci and Rn. The significance of the differences between the high- and low-saturation color-sets is estimated by a paired-sample *t*-test, Bonferroni corrected for 15 comparisons.

\*\* $p < .001$ .

set; a main effect for task [ $F(2, 174) = 533.6, p < .001, \eta_p^2 = 0.86$ ]; and a significant interaction of the two factors [ $F(2, 174) = 76.6, p < .001, \eta_p^2 = 0.47$ ], implying that the effects of the color-set were not equal across all tasks.

In a set of *post-hoc* *t*-tests (making Bonferroni corrections for the 15 possible comparisons among means), we first confirmed the ordering of latencies,  $Rn < Cn < Ci$  for the saturated [ $Cn-Rn, t(87) = 11.8, p < .001$ ;  $Ci-Cn, t(87) = 21.1, p < .001$ ;  $Ci-Rn, t(87) = 21.7, p < .001$ ] and desaturated conditions [ $Cn-Rn, t(87) = 18.4, p < .001$ ;  $Ci-Cn, t(87) = 19.7, p < .001$ ;  $Ci-Rn, t(87) = 24.2, p < .001$ ]. Next, to explicate the nature of the interaction of color-set and task, turn to the three leftmost columns of Table 2. The data suggest that the saturation manipulation had no effect on Rn, a substantial effect on Cn, and an even larger effect on Ci. Paired-sample *t*-tests confirmed that the effect of the saturation manipulation was significant for Cn [547 vs. 617 ms, for the high- and low-saturation conditions, respectively,  $t(87) = 10.0, p < .001$ ], and Ci [740 vs. 861 ms,  $t(87) = 10.1, p < .001$ ], but not for Rn [442 vs. 435 ms,  $t(87) = 1.0, p > .1$ ]. Furthermore, as shown in the three rightmost columns of Table 2, DI, the difference between Cn and Rn ( $DI = Cn-Rn$ ), was significantly larger in the desaturated than in the saturated condition [105 and 182 ms,  $t(87) = 8.8, p < .001$ ]. SI, the difference between Ci and Cn ( $SI = Ci-Cn$ ), was also significantly larger in the desaturated than in the saturated condition [193 and 244 ms,  $t(87) = 4.7, p < .001$ ], as was the difference between Ci and Rn [ $Ci-Rn, 298$  and  $426$  ms,  $t(87) = 11.2, p < .001$ ]. Finally, in a paired sample *t*-test we confirmed that the increase in SI and DI by desaturating the colors was not different in its extent [ $t(87) = 1.5, p > .5$ ]. To summarize, *post-hoc* analyses confirm that: (a) latencies increased from Rn to Cn to Ci; (b) the saturation manipulation affected the color-naming tasks, Ci and Cn, but not Rn, and (c) moving from high- to low-saturation condition increased both SI and DI in a similar manner.

### **Correlations**

The correlation matrix of response latencies amongst the six conditions (three tasks  $\times$  two levels of saturation) is presented in Table 3. All six conditions are highly inter-correlated (with a mean correlation coefficient of 0.673). A factor analysis shows that a single factor (Eigen-value = 4.4) that loads all six conditions can account for 73.0% of the variance, with no other factor approaching significance. The most likely candidate for this single factor is individual differences in speed of processing. To see if speed of processing could account for the extent of the Stroop effects in the saturated and desaturated conditions, we used Rn as an independent measure of speed of processing and correlated Rn and SI in the saturated condition, and in the desaturated condition, separately. Neither correlation was significant [ $r_p(86) = .047, p > .5$ , and  $r_p(86) = .16, p > .1$ , for the saturated and desaturated

**TABLE 3.** Correlation matrix ( $r_p$ ) of average response latencies among the six conditions (three tasks Rn–Cn–Ci, X two saturation conditions. saturated-desaturated)

	Saturated			Desaturated		
	Rn	Cn	Ci	Rn	Cn	Ci
<i>Saturated</i>						
Rn	1	0.71**	0.56**	0.74**	0.60**	0.40**
Cn	0.71**	1	0.84**	0.66**	0.85**	0.72**
Ci	0.56**	0.84**	1	0.49**	0.79**	0.81**
<i>Desaturated</i>						
Rn	0.74**	0.66**	0.49**	1	0.63**	0.49**
Cn	0.60**	0.85**	0.79**	0.63**	1	0.81**
Ci	0.40**	0.72**	0.81**	0.49**	0.81**	1

The *N* in all cases was 88.  
 \*\* $p < .001$ .

conditions, respectively]. Hence, speed of processing cannot account for the degree of Stroop interference in either saturation condition.

**Order Effects**

In Table 1 we portray how we fully counterbalanced the order of presentation of saturation blocks (saturated block presented first or last; denoted in Table 1 as group-orders a and b) and the temporal order in which the tasks were presented across these blocks (group-orders 1, 2, 3 and 4). To test whether these factors indeed had an impact on latencies in our study, we conducted a repeated measures ANOVA with color-set (high- vs. low-saturation) and task type (Rn, Cn or Ci) as within-participant factors, and saturation block order (group-orders a, or b) and task order (group-orders 1, 2, 3 or 4) as between-participants factors. Again, we found main effects for task type [ $F(2, 160) = 540.7, p < .001, \eta_p^2 = 0.87$ ], and color set [ $F(1, 80) = 126.1, p < .001, \eta_p^2 = 0.61$ ], and a significant interaction of the two [ $F(2, 160) = 91.6, p < .001, \eta_p^2 = 0.53$ ]. The order of color-naming tasks was not found to have a significant effect on latencies [ $F(3, 80) < 1$ ] or interact significantly with any of the other effects ( $F_s < 2, p > .1$ ) and will not be discussed further. In contrast, even though no main effect was found for saturation block order [ $F(1, 80) = 1.0, p > .1$ ], it was found to interact significantly with the color-set factor [ $F(1, 86) = 39.4, p < .001, \eta_p^2 = 0.33$ ], and finally there was a three-way interaction of block order with the color-set and the task factors [ $F(2, 160) = 19.2, p < .001, \eta_p^2 = 0.19$ ].

To understand the source of the three-way interaction for each task condition (Rn, Cn and Ci) we conducted a separate ANOVA with color-set as a within-participants factor (high- vs. low-saturation), and saturation block order (whether the high-saturation block was presented first or second,

group-orders a, or b) as a between-participants factor. To correct for conducting three additional ANOVAs on the same data set we applied a Bonferroni correction. For Cn, color-set was found to have a significant main effect [ $F(1, 86) = 100.5, p < .001, \eta_p^2 = 0.54$ ], but the main effect of practice was not significant [ $F(1, 86) < 1, p > .5$ ], and there was no interaction between the two [ $F(1, 86) < 1, p > .5$ ]. Thus, the order of saturation blocks had no effect on Cn latencies. For Ci, color-set again had a significant main effect [ $F(1, 86) = 144.3, p < .001, \eta_p^2 = 0.63$ ], but block order did not ( $F < 1, p > .1$ ). However, we found a significant interaction of the effects of saturation and block order [ $F(1, 86) = 36.9, p < .001, \eta_p^2 = 0.30$ ]. This interaction reflects the fact that the difference in response latencies for Ci between the low- with the high-saturation blocks ( $Ci_{low} - Ci_{high}$ ) was *larger* when the low-saturation block was presented first (182 and 60 ms). For Rn, the color-set had no main effect on latencies [ $F(1, 86) = 1.9, p > .1, \eta_p^2 = 0.02$ ], neither did the block order [ $F(1, 86) = 3.3, p > .1, \eta_p^2 = 0.04$ ]. However, the two factors interacted significantly [ $F(1, 86) = 50.7, p < .001, \eta_p^2 = 0.37$ ]. Here, the interaction indicates that moving from a low- to a high-saturation block ( $Rn_{low} - Rn_{high}$ ) increased latencies for Rn *only* when the low-saturation block was presented first. However, this effect was reversed when the high-saturation block was presented first (31 vs. -46 ms). In other words, participants were faster the second time they were asked to perform the Rn task.

Taken together, these *post-hoc* tests show that the effect of the order of saturation blocks interacted significantly with the effect of color-set only on latencies in Rn and Ci. Note that both of these tasks involved stimuli that carry lexical content, whereas in the Cn task, the stimuli consisted of strings of X's. Therefore, practice appears to have a larger impact on Stroop tasks with lexical content with younger-adult participants.

### ***Controlling for Order Effects: A Between-Groups Analysis***

The preceding analysis indicated that the order in which participants experienced the two saturation conditions had an effect on response latencies. To see if the effect of the saturation condition on the three Stroop tasks observed when the data included both saturation orders also holds *only in the first saturation block* (see Table 1 where in order a, 44 participants were tested first in the saturation condition with the remaining 44 participants, order b, tested first in the desaturation condition), we conducted a repeated-measures ANOVA with task as a within-participants factor (Rn, Cn or Ci) and color-set as a between-participants factor (high- vs. low-saturation) for the first block data only.

The analysis revealed the same pattern of results as the analysis of the full data set: a main effect for task [ $F(2, 172) = 448.7, p < .001, \eta_p^2 = 0.84$ ]; a main effect for color-set [ $F(1, 86) = 13.5, p < .001, \eta_p^2 = 0.14$ ], reflecting faster responses for the group who performed in the high-saturation set; and

a significant interaction of the two factors [ $F(2, 172) = 12.6, p < .001, \eta_p^2 = 0.13$ ], showing that the effects of the color-set were not equal across all tasks. Hence, the results of the between-participants analysis replicated the findings in the within-participants analysis. In a set of  $t$ -tests, we again found the same pattern of effects, Rn was not affected by color manipulation [446 vs. 470, for high- and low-saturation, respectively,  $t(86) = 1.3, p > .1$ ], but Cn was [544 vs. 623,  $t(86) = 3.3, p = .001$ ], producing an increase in DI when comparing the high-saturation group to the low-saturation group [98 vs. 153 ms,  $t(86) = 3.3, p < .005$ ]. As we found in the within-participant analyses, Ci was affected by the color-manipulation [756 vs. 906,  $t(86) = 3.4, p < .001$ ], more than Cn, as indicated by a significant increase in SI when moving between the two saturation groups [212 vs. 283 ms,  $t(86) = 4.1, p < .005$ ]. In sum, even though practice was found to affect RTs, the effects of color-set on Stroop performance (an increase in DI and SI) persisted.

### *Sensory Degradation*

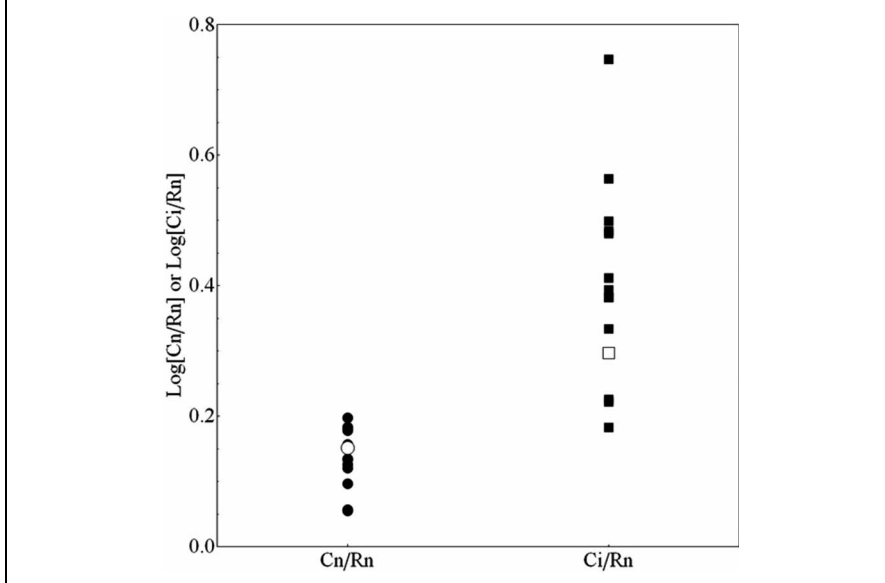
In accordance with the sensory degradation account for Stroop and aging, we found that latencies for Ci were positively correlated with DI across the 88 participants, both in the high-saturation condition ( $r_p = .61, p < .001$ ) and in the low-saturation condition ( $r_p = .55, p < .001$ ). Moreover, we noted that a third of our participants (29) had a slight color-vision deficiency as manifested by a score of 7 out of 8 in an Ishihara color-test (approximately evenly distributed across experimental conditions). This slight color deficiency increased latencies for Ci more than for Cn in the saturated color set, resulting in a moderate increase in SI in a between-participants comparison [180 vs. 217 ms,  $t(86) = 1.9, p < .05$ , one-tail test<sup>1</sup>]. Similar correlations between Stroop effects and color-vision have been reported for both younger (e.g., Laeng, Låg & Brennen, 2005) and older adults (Anstey, Dain, Andrews, & Drobny, 2002).

To determine the extent to which desaturating the colors in young adults mimics the effects of aging, we have compared the results of this manipulation to those reported in a meta-analysis of age effects in color-word Stroop tests (Ben-David & Schneider, 2009). First we note that latencies for Rn in the color desaturation condition fall within the inter-quartile range of older adults' Rn latencies in the 13 studies examined in the meta-analysis. Hence in terms of reading, our younger adults were reasonably matched with respect to older adults (see a similar example in Hartman & Hasher, 1991). Because Rn latencies are generally found to be correlated with both Cn and Ci latencies (e.g., Salthouse & Meinz, 1995), we have

---

<sup>1</sup>Because there was no reason to believe that a reduction in color-vision would result in faster color-naming responses, a one-tail test was indicated. In addition, a non-parametric test (Mann-Whitney) also found a significant difference between the two groups [Mann-Whitney  $U(N_1 = 29, N_2 = 59) = 616, p < .05$ , two-tailed].

**FIGURE 1.** Scatter plot for the logarithmic increments in latency in going from reading a color-neutral word (Rn) to color-naming a neutral word (Cn, circles), and from Rn to color-naming incongruent words (Ci, squares) for older adults in 13 color-word Stroop studies. These data points are adapted from Table 1 in Ben-David and Schneider (2009, p. 511). The empty symbols represent the points corresponding to the logarithmic increment in the color-desaturated condition in the current study. Note that our simulated data is within the inter-quartile range of older adults' data in the 13 studies.



expressed the changes in latency in Cn and Ci relative to the Rn latency. In effect, Rn latency served as a baseline for speed of processing. Hence, in Figure 1, we show a scatter plot of the logarithmic increments in latency in going from Rn to Cn (circles), and from Rn to Ci (squares) in these 13 studies along with the logarithmic increment in our color-desaturated condition (empty symbols). These plots indicate that our simulation of color desaturation mimics the proportional increments in Cn and Ci relative to Rn found in the Stroop performance of older adults.

## GENERAL DISCUSSION

The present results indicate that simulating a color deficiency in younger adults can accurately mimic the color-word Stroop performance characteristic of older adults. In a previous paper (Ben-David & Schneider, 2009), we proposed a model in which age-related increments in DI (the advantage of reading over color-naming) mediates the effects of age on Stroop performance. Age-related sensory deficits presumably increase the time it takes to access the print color of the word, thereby requiring the older adult to inhibit the lexical response for a longer time until sufficient access is gained to the

print color of the word (i.e., inhibit the 'red' response to the word RED printed in blue). This, in turn, raises the question as to whether the poorer performance of older adults on color-word Stroop truly reflects loss of inhibitory control. Since the same participants were tested in both the saturated and desaturated versions of the Stroop paradigm in our study, the larger Stroop interference found for desaturated colors cannot be due to a loss of inhibitory control. Therefore, our study shows that it is possible to mimic age effects without a loss of inhibitory control.

Since one would expect color desaturation to increase latencies for both Cn and Ci, it is possible to interpret the experimental manipulation conducted here as a simulation of age-related slowing with respect to these two tasks. It should be noted, however, that desaturating the colors did not slow down latencies for Rn. Hence, there is no evidence in this study that the saturation of the font colors had an effect on lexical access. A full simulation of cognitive slowing would require that latencies on *all* three tests increase by the *same proportional* amount (see Eq. 2). In our simulation, Rn did not increase at all, and the ratio of desaturated-Ci to saturated-Ci latencies was significantly larger than the ratio of desaturated-Cn to saturated-Cn latencies [ $Ci_{low}/Ci_{high}$  vs.  $Cn_{low}/Cn_{high}$ ,  $t(87) = 2.1$ ,  $p < .05$ ]. Our study, thus, further shows that it is possible to simulate aging effects in Stroop by changing discriminability along the color dimension (mimicking age-related sensory degradation), when the simulation violates the assumptions of a cognitive slowing model.

Of course, successfully simulating the effects of aging by desaturating the colors does not prove that age-related changes in Stroop effects are solely due to age-related changes in perception. Age-related changes in cognitive functioning (e.g., the inhibitory deficit hypothesis) and/or generalized slowing are also likely to contribute to Stroop effects, and the ways in which these factors interact with age-related changes in color-vision are likely to vary from individual to individual. However, our simulation does show that making it more difficult for younger adults to process color information increases the time it takes to name of the color for the Ci conditions more than it does for the Cn condition. Hence, there is reason to believe that age-related declines in color-vision will have a similar effect in older adults, thereby leading to an age-related increase in Stroop effects.

Our results do, however, show that the age-related changes that occur are consistent with a perceptual degradation model, and they suggest that age-related changes in color-vision need to be considered when the Stroop test is used to measure selective attention as part of a battery to construct a neuropsychological profile for older adults. We note that the effects of *normal* age-related changes in color-vision on Stroop may be mitigated by using age-corrected norms (see Troyer et al., 2006, for an example on normative data for a standardized Stroop test), and that sensory functions are sometimes

included in neuropsychological test batteries. However, considering the high variability in sensory function in older age (see Schneider & Pichora-Fuller, 2000, for an overview), and the large impact of color information on the Stroop task, as indicated by our data (see also Anstey et al., 2002, for an estimate of the effect of individual color-vision differences on Stroop in older age), one is advised to be mindful of how potential individual differences in color-vision (and other perceptual functions) could affect Stroop performance in older adults.

## SUMMARY

A recent meta-analysis of Stroop and aging (Ben-David & Schneider, 2009) proposed that a sensory source – deterioration in color-vision with age – can account, in part, for age-related changes in Stroop effects. We have complemented the theoretical approach with empirical data, showing that mimicking an age-related decrease in color perception in younger adults is sufficient to lead to Stroop effects that are similar to those found in older adults. As the color-word Stroop test is widely used in clinical, neuropsychological and experimental screening tests as a measure of inhibitory control in older participants, it might be advisable to take into account individual changes in color-vision and other perceptual factors that might influence Stroop performance in older adults.

## ACKNOWLEDGEMENTS

This study was partially supported by a group grant on Sensory and Cognitive Aging, funded by Canadian Institutes of Health Research Grants (STP-53875, MGC-42665, & MOP-15359), and a research opportunity program grant from the Faculty of Arts & Science at the University of Toronto Mississauga. The first author was partially supported by a grant from the Ontario Neurotrauma Foundation (2008-ABI-PDF-659). We wish to thank Wu Yan (Lulu) Li, Linh Le Truc Nguyen, Nicole Durham and Amanda Dydynski for their assistance in collecting the data.

Original manuscript received 1 November 2009

Revised manuscript accepted 30 June 2010

## REFERENCES

- Akutsu, H., Legge, G. E., Ross, J. A., & Schuebel, K. J. (1991). Psychophysics of reading—X. Effects of age-related changes in vision. *Journal of Gerontology*, *46*, 325–331.
- Algom, D., Dekel, A., & Pansky, A. (1996). The perception of number from the separability of the stimulus: The Stroop effect revisited. *Memory and Cognition*, *24*, 557–572.
- Anstey, K., Dain, S., Andrews, S., & Drobny, J. (2002). Visual abilities in older adults explain age-differences in Stroop and fluid intelligence but not face recognition: Implications for the vision-cognition connection. *Aging Neuropsychology and Cognition*, *9*, 253–265.



- Balota, D. A., Tse, C.-S., Hutchison, K. A., Spieler, D. H., Duchek, J. M., & Morris, J. C. (in press). Predicting conversion to dementia of the alzheimer type in a healthy control sample: The power of errors in Stroop color naming. *Psychology and Aging*.
- Ben-David, B. M., & Schneider, B. A. (2009). A sensory origin for aging effects in the color-word Stroop task: An analysis of studies. *Aging, Neuropsychology, and Cognition*, *16*, 505–534.
- Ben-David, B. M., Chambers, C., Daneman, M., Pichora-Fuller, M. K., Reingold, E., & Schneider, B. A. (2010). Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements. *Journal of Speech, Language, and Hearing Research*. (in press).
- Birch, J. (1997). Efficiency of the Ishihara test for identifying red-green colour deficiency. *Ophthalmic and Physiological Optics*, *17*, 403–408.
- Cerella, J. (1990). Aging and information processing rate. In J. Birren & K. Schaie (Eds.), *Handbook of the psychology of aging* (3 ed., pp. 201–221). San Diego, CA: Academic Press.
- Cerella, J., & Hale, S. (1994). The rise and fall in information-processing rates over the life span. *Acta Psychologica (Amst)*, *86*, 109–197.
- De Corte, W. (1986). Finding appropriate colors for color displays. *Color Research and Application*, *11*, 56–61.
- Eidels, A., Townsend, J. T., & Algom, A. (2010). Comparing perception of Stroop stimuli in focused versus divided attention paradigms: Evidence for dramatic processing differences. *Cognition*, *114*, 129–150.
- Frost, R., Ahissar, M., Gotesman, R., & Tayeb, S. (2003). Are phonological effects fragile? The effects of luminance and exposure duration on from priming and phonological priming. *Journal of Memory and Language*, *48*, 346–378.
- Golden, C. (1978). *The Stroop Color and Word Test*. Chicago: IL: Stoelting Company.
- Hartman, M., & Hasher, L. (1991). Aging and suppression: Memory for previously relevant information. *Psychology and Aging*, *6*, 587–594.
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.
- Laeng, B., Låg, T., & Brennen, T. (2005). Reduced Stroop interference for opponent colors may be due to input factors: Evidence from individual differences and a neural network simulation. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 438–452.
- Lindsay, D., & Jacoby, L. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 219–234.
- Ludwig, C., Borella, E., Tettamanti, M., & de Ribaupierre, A. (in press). Adult age differences in the Color Stroop Test: A comparison between an Item-by-item and a Blocked version. *Archives of Gerontology and Geriatrics*, (2009), doi: 10.1016/j.archger.2009.09.040
- McDowd, J., & Shaw, R. (2000). Attention and aging: A functional perspective. In F. Craik & T. Salthouse (Eds.), *The handbook of aging and cognition* (2 ed., pp. 221–292). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, *110*, 422–471.
- Melara, R. D., & Mounts, J. R. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition*, *21*, 627–645.
- Mutter, S. A., Naylor, J. C., & Patterson, E. R. (2005). The effects of age and task context on Stroop task performance. *Memory and Cognition*, *33*, 514–530.

- Nguyen-Tri, D., Overbury, O., & Faubert, J. (2003). The role of lenticular senescence in age related color vision changes. *Investigative Ophthalmology & Visual Science*, *44*, 3698–3704.
- Perlstein, W., Carter, C., Barch, D., & Baird, J. (1998). The Stroop task and attention deficits in schizophrenia: A critical evaluation of card and single-trial Stroop methodologies. *Neuropsychology*, *12*, 414–425.
- Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Raven, J. C. (1965). *The Mill Hill vocabulary scale*. London: H. K. Lewis.
- Sabri, M., Melara, R. D., & Algom, D. (2001). A confluence of contexts: Asymmetric versus global failures of selective attention to Stroop dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 515–537.
- Salthouse, T. A., & Meinz, E. J. (1995). Aging, inhibition, working memory, and speed. *Journals of Gerontology: Psychological Sciences*, *50*, 297–306.
- Schneider, B. A., & Pichora-Fuller, M. K. (2000). Implications of perceptual deterioration for cognitive aging research. In F. I. M. Craik, & T. A. Salthouse (Eds.), *The Handbook of Aging and Cognition* (2 ed., pp. 155–219). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Schneider B. A., Pichora-Fuller, M. K., & Daneman M. (2010). The effects of senescent changes in audition and cognition on spoken language comprehension (pp. 167–210). In S. Gordon-Salant, R. D. Frisina, A. N. Popper & R. R. Fay (Eds.), *The aging auditory system* (pp. 167–210). New York, NY: Springer.
- Troyer, A. K., Leach, L., & Strauss, E. (2006). Aging and response inhibition: Normative data for the Victoria Stroop test. *Aging, Neuropsychology, and Cognition*, *13*(1), 20–35.
- Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2006). The Stroop color-word test: Influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment*, *13*, 62–79.
- Verhaeghen, P., & Cerella, J. (2002). Aging, executive control, and attention: A review of meta-analyses. *Neuroscience and Biobehavioral Reviews*, *26*, 849–857.
- Werner, J. S., & Steele, V. G. (1988). Sensitivity of human foveal color mechanisms throughout the life span. *Journal of the Optical Society of America. A, Optics and Image Science*, *5*, 2122–2130.