Stroop Color-Word Test Performance in Patients with Alzheimer’s Disease*

Laurel M. Fisher¹, David M. Freed¹², and Suzanne Corkin²

¹ The University of Southern California
² Massachusetts Institute of Technology: Department of Brain and Cognitive Science

ABSTRACT

Patients with Alzheimer’s disease (AD) (n = 36) and normal older adults (n = 36) were individually administered the Stroop Color-Word Test. Eight of 36 (22%) AD patients exhibited confusion between the colors blue and green, while no control subject had difficulty distinguishing among the colors. In a second experiment, a subset of the original sample (15 AD patients and 8 control subjects) was retested using the Stroop. Only 2 AD patients showed color confusion on both test occasions, while 7 AD patients exhibited color confusion on one occasion. No control subject exhibited confusion between colors the second time. These results indicate that color confusion in AD patients is inconsistent. Due to the high incidence of color confusion in AD patients, the Stroop should be used with caution in patient populations.

The Stroop Color-Word Test (the Stroop) has been found to measure cognitive processes, such as selective attention and flexibility (Glaser & Glaser, 1989; Golden, 1978; Stroop, 1935). Traditionally, the Stroop consists of three subtests. Subjects read color names, name colors, and, finally, identify the color of ink in which a word is printed (rather than reading the word itself). The Stroop Interference effect occurs when the word is the name of a color, as on the Color-Word subtest, and reading speed is slowed.

* The authors would like to thank Alfredo Sadun, M.D., Ph.D. for his comments on an earlier draft. These studies were supported by NIGMS Training Grant ST32-GM07484 to David Freed; United States Public Health Services Grants MH24433, RR00088, and R01 MH32724-07A1 (P.I.: Suzanne Corkin); and NIA Predoctoral Training Grant AG00037-12 to Laurel Fisher (P.I.: Vern Bengston).

Correspondence regarding this article should be addressed to:
Laurel M. Fisher, University of Southern California, Department of Psychology, MC 1061, Los Angeles, CA 90089-1061, USA.

Accepted for publication: February 2, 1990.
Few studies have examined changes in the Stroop Interference effect due to normal aging. In a developmental study, Comalli, Wapner, and Werner (1962) administered the Stroop to 235 subjects spanning the ages between 7 and 80 years old. Subjects were instructed to name 100 stimuli (words, colors, names of colors) as quickly as possible. Older adults and children had longer response latencies than young adults and middle aged individuals in the Color-Word subtest. Older adults were nearly as fast as young adult subjects when reading the words or naming the colors, but were disproportionately slowed in the Color-Word subtest (Comalli et al., 1962).

As part of a larger study examining cognitive functioning in older adults, Comalli, Krus, and Wapner (1965) found that the time to name all the words in each Stroop condition was significantly slower in 15 institutionalized older men (aged 67-81 years) compared to 15 community-dwelling older men (65-80 years). Institutionalized men were disproportionately slower in the Color-Word subtest compared to normal men (Comalli et al., 1965).

Comalli (1965) compared 20 community-dwelling men aged 80-90 years to the sample of normal men (aged 61-80 years) from the Comalli et al. (1965) study. The scores of the older men were not reliably lower than the younger men when reading the words and naming the colors on the Stroop. However, the older men were significantly slower relative to the younger men in the Color-Word subtest. In a footnote, Comalli (1965) reported that two of the 20 normal men (aged 80-90 years) “failed to adequately discriminate blue and green colors” on the Stroop (p. 16). However, he did not provide any description of those subjects, nor any additional data to clarify this observation. Comparing the Color-Word subtest scores across these two studies, institutionalized men were slower than the older sample of men, who in turn were slower than the younger sample. These results suggest that the Stroop Interference effect increases with age, in individuals over 80 years of age, and in patient populations.

Bettner, Jarvik, and Blum (1971) administered the Stroop to 38 older adults who had previously been a part of a longitudinal study of twins. Fourteen subjects had been diagnosed as having organic brain syndrome (OBS) and 24 did not have OBS. The OBS subjects had higher scores on the first two Stroop subtests compared to those without OBS, but the difference was statistically significant only on the Color-Word subtest. There was no relationship between disease status and Stroop Interference score. Bettner et al. (1971) concluded that the Stroop could be used to detect individuals in the early stages of OBS. The absence of a relationship between the Interference score and OBS was problematic and the authors concluded that further research was in order.

Koss, Ober, Delis, and Friedland (1984) administered the Stroop to both mildly- and moderately-impaired Alzheimer’s disease (AD) patients, patients with dementia other than AD, and normal older adults. The Stroop was administered in the form of a reaction time task. Normal older subjects were faster on the Color-Word subtest than all patient groups. Mild AD and non-AD patients had slower reaction times on the Color-Word subtest than the moderate AD
patients, a counter-intuitive finding. Koss et al. (1984) suggested that the increased interference seen in the mildly demented AD patients may be due to a speed/accuracy trade-off. The mild AD patients were still able to monitor their performance and were aware of errors they made. The authors speculated that an awareness of mistakes significantly slowed the performance of mild AD patients relative to the moderately impaired AD patients, who were no longer able to realize when an error had been made (Koss et al., 1984). Errors in the naming of the ink color tended to increase with the severity of dementia. Koss et al. (1984) concluded that performance on the Stroop may be an index of dementia severity and that higher order cognitive processes are impaired early in the progression of AD.

Cohn, Dustman, and Bradford (1984) administered the Stroop to subjects across the life-span and found that older adults (aged 61-90 years) were slower when naming the standard three ink colors on the Color-Word subtest compared to younger adults (21-50). In addition, the oldest subjects (71-90 years old) were significantly slower than the 60-70-year-old group on the Color-Word subtest. Noting that most older adults experience yellowing of the lens and that this may have caused some confusion when discriminating among the colors, Cohn et al. (1984) analyzed the color identification errors made by older adults. They did not find a pattern of errors by color, i.e., the older adults had more or less random errors in color identification. These findings suggest that normal older adults do not have difficulty discriminating among different colors, specifically blues and yellows. Cohn et al. (1984) suggest that the disproportionate increase in response time on the Interference task observed in the older adults is primarily due to slowing of cognitive processing rather than perceptual difficulties.

Panek, Rush, and Slade (1984) administered the Stroop to 50 young adults (18-23 years) and 31 community-dwelling older adults (61-85 years). The subjects were healthy and screened for color blindness, a step previous investigators had not taken. The Stroop procedure used was identical to that used by Comalli et al. (1962). Older adults were slower in all three Stroop conditions and were disproportionately slower on the Color-Word subtest. Panek et al. (1984) suggested that age differences on the Color-Word subtest may be due to differences in response dominance (i.e., the word reading response is stronger than the color naming response). The authors speculate that the larger interference effects seen with increasing age could be due to increased complexity of the task (see also Rush & Panek, 1983).

Overall, the findings from these studies confirm the notion that the Stroop measures cognitive processing and that the effects of interference are increased with the normal aging process. Response latencies on the Color-Word subtest are disproportionately increased in institutionalized men and AD patients. However, several problems with previous studies need to be addressed. In one of the two studies involving dementia patients, Koss et al. (1984) administered the Stroop in a reaction time setting. Reaction time tasks tend to confound the normal slowing seen in older adults (Birren, 1964) with the Stroop Interference effect. The other
studies of normal older adults did not use reaction time measures. Thus, the conclusion drawn by Koss et al. (1984) that the Stroop Interference score is sensitive to dementia may be confounded with the method of measurement. In contrast, Bettner et al. (1971) found no relationship between disease status and Stroop Interference score, suggesting that the Stroop is not sensitive to the presence of AD. Further, Bettnner et al. (1971) found no significant difference between subjects with OBS and those without OBS on the Stroop subtests, with the exception of the Color-Word subtest. Additional research is needed to clarify the relationship between performance on the Stroop and dementia.

In the studies reviewed here, no attempt was made to detect color blindness in the subjects, except for Panek et al. (1984). It is interesting to note that only Comalli (1965) reported that normal older adults (albeit 2 subjects) displayed confusion between the colors blue and green used on the Stroop subtests. Cohen, Cronin-Golomb, Growdon, and Corkin (1988) noted that of the 67 AD patients administered the Stroop, 15% reported difficulty discriminating between the colors blue and green (blue/green color confusion) during the color naming portion of the test. Cohen et al. (1988) tested the color vision of an independent group of AD patients and normal controls. The AD patients made significantly more errors in color identification than the controls on the City University Colour Vision Test. It was concluded that AD patients have more difficulty with color naming than normal older adults.

The present studies were conducted to further investigate the use of the Stroop in both normal older adults and AD patient populations. Subjects were screened for visual impairments and color blindness. In the first study, Stroop scores were examined for overall group differences and associations with dementia severity. Reaction time measures were not collected during administration of the Stroop. It was noted that AD patients made more blue/green color identification errors than control subjects, but that color confusion was not associated with dementia severity. These findings replicate Cohen et al. (1988).

Color confusion is both an interesting phenomenon and a potential impediment to continued use of the Stroop in patient populations. In order to separate various explanations for the observed color confusion, a second experiment was undertaken to determine if the patients who had difficulty discriminating between the two colors on one occasion were unable to discriminate the colors on a second test administration.

**METHOD**

The Stroop Color-Word test employs three cards, each containing five columns of twenty sets of symbols. Each column in the first card is composed of the words BLUE, GREEN, and RED printed in black ink (the Word subtest). The subjects are given 45 seconds to read as many words as possible. Each column on the second card is composed of XXXXs printed in blue, green or red ink (the Color subtest). The subject is to name as many colors as he can within 45 s. Each column in the third card (Color-Word subtest) is composed
Table 1. Experiment 1: Normal older adult and AD patient characteristics.

<table>
<thead>
<tr>
<th></th>
<th>Normal Older Adults (n = 36)</th>
<th>AD Patients (n = 36)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>72.9 (8.3)</td>
<td>70.1 (8.9)</td>
</tr>
<tr>
<td>Males/Females</td>
<td>13/23</td>
<td>14/22</td>
</tr>
<tr>
<td>Years of Education</td>
<td>14.6 (2.7)</td>
<td>13.9 (4.0)</td>
</tr>
<tr>
<td>Blessed Dementia Scale</td>
<td>1.5 (6.1)</td>
<td>15.5 (6.1)</td>
</tr>
</tbody>
</table>

of the words BLUE, GREEN, and RED printed in a contrasting ink (e.g., the word BLUE is printed in green ink). Each subject is given 45 s to name the color of the ink in which the words are printed (and not read the word itself) as quickly as possible (Golden, 1978).

The Stroop was administered to 36 patients with a clinical diagnosis of AD (Khachaturian, 1985; McKhann et al., 1984, for diagnostic criteria) and 36 normal older adult subjects. The normal controls were recruited through the University of Southern California Alzheimer's Disease Research Center (USC-ADRC) and usually were the spouse of an AD patient. The medical records of the patients and normal subjects were examined for notations of color blindness or other visual problems. All subjects included here do not have a history of color blindness, cataracts, or glaucoma. Table 1 lists subject characteristics.

The average Blessed Dementia Scale (BDS) score of the AD patients was 15.5 errors out of a possible 37 errors, indicating moderate levels of dementia for the AD subjects as a whole. The normal older adults has a significantly lower BDS score than the AD patients \(F(2.67) = 44.2, p < .00001\), as expected. The AD patients averaged 13.9 years of education, which was not significantly different from the control group \(F(2.67) = 1.2, p = .31\). Some subjects \(n = 24\) were tested at the Massachusetts Institute of Technology Clinical Research Center and the remainder were tested at the USC-ADRC.

RESULTS

Table 2 lists the mean scores for normal older adults and AD patients on the Stroop subtests, as well as the calculated Stroop Interference score. The Interference score was calculated by subtracting a predicted Color-Word score from the observed Color-Word score. The predicted score was calculated by multiplying the Word score by the Color score, and dividing the resultant number by the addition of the Word score and the Color score (Golden, 1978).

There was a significant Pearson correlation between the BDS and the Stroop Interference score \(r = -0.275, p = .013\). This indicates that those patients who
Table 2. Mean number of correct responses for each Stroop subtest, as well as calculated Interference scores, for the normal older adults and AD patients.

<table>
<thead>
<tr>
<th>Stroop Test Condition</th>
<th>Normal Older Adults (n = 36)</th>
<th>AD Patients (n = 36)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word (SD)</td>
<td>96.6 (15.8)</td>
<td>65.0 (18.9)</td>
</tr>
<tr>
<td>Color (SD)</td>
<td>64.9 (13.9)</td>
<td>32.7 (14.0)</td>
</tr>
<tr>
<td>Color-Word (SD)</td>
<td>33.4 (10.8)</td>
<td>9.7 (8.5)</td>
</tr>
<tr>
<td>Interference Scores (SD)</td>
<td>-5.2 (8.6)</td>
<td>-12.0 (7.0)</td>
</tr>
</tbody>
</table>

had high BDS scores (more demented) had large negative Interference score (little resistance to interference).

No normal adult had difficulty distinguishing between the colors used in the Stroop. However, 8 (5 female, 3 male) of the 36 AD patients displayed difficulty discriminating between blue and green on the Color subtest. For example, some subjects would look at the Color sheet and ask: "Which is green and which is blue?" Several subjects were able to utilize verbal feedback to improve their performance and subsequently were able to reliably distinguish green and blue.

Due to the larger number of patients displaying color confusion, for further analyses the AD patients were split into two groups: those who exhibited blue/green color confusion and those who did not. A repeated measures ANOVA was carried out, with Stroop subtest scores (Word, Color, Color-Word, Interference) as the within-subject factor and group (normal, AD-color confusion, AD-no color confusion) as the between-subject factor. The groups differed significantly across all Stroop scores \( F (2, 63) = 55.30, p < .001, MS = 310.03 \). Planned comparisons showed that this was due to the significantly higher scores in the control group. The patient groups did not differ significantly on the Word, Color, and Color-Word subtests (Figure 1). The Interference score (Figure 2) for the color confusion group were lower \( (M = -13.5, SD = 5.3) \) than for those patients who did not display color confusion \( (M = -9.5, SD = 8.9) \). This difference was not significant \( t (21) = 1.50, p = .15 \). It is important to note that the AD patients with color confusion did not differ with regard to severity of dementia from the AD patients without color confusion, as the mean BDS scores were nearly the same, with 14.9 mean errors \( (SD = 5.3) \) and 16.1 mean errors \( (SD = 7.0) \), respectively.
**STROOP PERFORMANCE IN AD**

**Fig. 1.** Experiment 1: Mean Stroop test condition scores for normal older adults, AD patients without color confusion, and AD patients with color confusion.

**DISCUSSION**

The AD patients were significantly impaired on all Stroop subtests, especially the Color-Word subtest, relative to normal older adults. The significant negative correlation between BDS scores and Stroop Interference scores suggests that Interference scores are sensitive to dementia severity. Twenty-two per cent of the AD patients administered the Stroop showed signs of difficulty discriminating between the colors blue and green. It should be emphasized that no normal older adult confused the colors blue and green. Similar observations have been made by the Laboratory of Neuroscience, National Institute of Aging (J. Haxby, personal communication, June 29, 1988), replicating the findings of Cohen et al. (1988). While performance on the Stroop by patients with blue/green color con-
nal impairment would result in transmission of degraded stimuli to visual processing centers.

One could also argue that the older subjects' lenses have become excessively yellowed or have opacities (e.g., cataracts). However, there was no such notation about excessive yellowing in the patients' medical records and there is no reason to suspect that a particular subgroup of AD patients had more yellowing than the control subjects.

In addition, other dysfunctions unrelated to dementia could contribute to color confusion, such as congenital and acquired color blindness. Red/green color blindness is the most frequent form of acquired color blindness, while yellow/blue dichromats make up only a small percentage of individuals with acquired color blindness (Hirrich, 1981). The most common form of congenital color blindness is protanopia, a dysfunction or lack of red cones in the retina. This genetic trait is sex-linked to the X chromosome and, therefore, males are more likely to inherit the abnormality than females (approximately 2% incidence in males; Mollon, 1982). Medical records for the 3 males with color confusion did not indicate any color vision problems. Thus, it is unlikely that congenital color blindness could account for the data. The other 5 AD patients exhibiting color confusion were female and therefore were less likely than the male patients to have congenital color blindness.

Blue/green color confusion could also be related to the cortical degeneration known to occur in AD (Khachaturian, 1985; McKhann et al., 1984). Perhaps the cortical degeneration may spread to the visual processing centers involved in color discrimination (Cronin-Golomb, Corkin, Rizzo, Cohen, Growden, & Banks, 1990). The following experiment was performed in order to evaluate these possible explanations for color confusion.

**EXPERIMENT 2**

In the light of alternative explanations for the cause of blue/green color confusion in AD patients, a second study was carried out to determine how stable this phenomenon is. If color confusion is the result of some peripheral impairment, it is expected to be permanent and irreversible. Therefore, subjects who displayed color confusion during the first administration of the Stroop should continue to display this confusion. If color confusion is related to impairments of higher order visual processing, color confusion is expected to be exhibited inconsistently. A subset of the original sample of subjects was given the Stroop on one other occasion.

**METHOD**

From the previous sample, 8 normal older adults and 15 AD patients were re-administered the Stroop 1 to 12 months following the first administration. Subjects' characteristics are given in Table 3. The Stroop was administered in the same manner as in Experiment 1.
Table 3. Experiment 2: Characteristics of normal older adults and AD patients.

<table>
<thead>
<tr>
<th></th>
<th>Normal Adults (n = 8)</th>
<th>AD-color confusion (n = 9)</th>
<th>AD-no color confusion (n = 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (SD)</td>
<td>68.0 (10.3)</td>
<td>70.9 (11.7)</td>
<td>70.8 (11.8)</td>
</tr>
<tr>
<td>Males/Females</td>
<td>3/5</td>
<td>3/6</td>
<td>3/3</td>
</tr>
<tr>
<td>Years of Education (SD)</td>
<td>15.8 (3.2)</td>
<td>12.9 (3.8)</td>
<td>12.7 (3.8)</td>
</tr>
<tr>
<td>Blessed Dementia Scale (SD)</td>
<td>0.8 (0.8)</td>
<td>13.1 (4.5)</td>
<td>16.8 (8.0)</td>
</tr>
</tbody>
</table>

RESULTS

None of the 8 normal older adults displayed color confusion on either test occasion. Of the 15 AD patients, 6 never exhibited color confusion and 7 exhibited color confusion on just one of the test administrations. Only 2 AD patients exhibited color confusion during both test administrations.

We conducted repeated measures ANOVA with test occasion (Time 1, Time 2) and Stroop subtest scores (Word, Color, Color-Word, Interference) as the within-subject factors and group (control, AD-color confusion, and AD-no color confusion) as the between-subject factor. For all subjects, the Stroop subtest and interference scores were not significantly different from Time 1 to Time 2 [F (1,16) < 1]. The groups were significantly different [F (2,16) = 237.89, p < .0001] across all Stroop subtests (Figure 3) and Interference scores (Figure 4). Planned comparisons showed that the normal older adults had significantly higher scores on all subtests than the 2 patient groups. In addition, the normal older adults experienced less interference than the patient groups. The 2 AD patient groups were not significantly different from each other on Stroop subtest and interference scores.

The Stroop Interference score on the second test occasion correlated (r = -0.517, p = .02) with the BDS score from time 2. While higher BDS scores were associated with larger negative interference scores (i.e., little resistance to interference), the AD patient groups did not differ in mean BDS score.

DISCUSSION

Repeated testing of normal older adults and AD patients demonstrated that only 2 of 15 AD patients exhibited color confusion on both test occasions. Further-
more, 7 AD patients exhibited color confusion on only one test occasion. This indicates that color confusion is inconsistently displayed, tentatively supporting the notion that color confusion is not solely due to impairment in peripheral processes. Stroop subtest, interference scores, and BDS scores did not reveal any differences between the 2 AD patient groups, suggesting that color confusion may be unrelated to the processes which underlie those tasks.

GENERAL DISCUSSION

The Stroop was administered to normal older adults and AD patients. At the time of the first testing, 8 AD patients were unable to distinguish between the colors
Fig. 4. Experiment 2: Second test administration mean Stroop Interference scores for normal older adults, AD patients who never displayed color confusion, and AD patients who displayed color confusion.

blue and green. Analyses did not reveal any significant differences among the patient groups, although the AD patients with color confusion tended to have larger Interference scores.

The Stroop was administered a second time to a subset of the experimental subjects. Only 2 AD patients displayed color confusion at both test sessions. Eight normal subjects and 6 AD patients never had difficulty discriminating among the colors. In addition, 7 of the 15 AD patients exhibited color confusion in only one of the test sessions. This lends support to the hypothesis that color confusion is not due to peripheral degeneration in the visual system. While it is too early to implicate degeneration in the cortical visual centers, further studies are needed to investigate color discrimination in patients with AD and possible relationships to cortical pathology.
It is not surprising that Stroop subtest and Interference scores did not differentiate the patient groups. Blue/green color confusion appears to be a phenomenon unrelated to the cognitive processes measured by the Stroop. Further, color confusion appears to be unrelated to dementia severity. Additional studies are needed to clarify the nature of this deficit in AD patients.

It is interesting to note that an impairment in color discrimination (dyschromatopsia - from blue/yellow loss to more complex color vision loss) has been observed in people exposed to solvents, such as printers (Braun, Daigleault, & Gilbert, 1989). The authors tested 29 printers and 29 control subjects on color discrimination and a wide variety of neuropsychological tests. The authors found no difference among the groups on the neuropsychological tests, but the printers had a higher prevalence of dyschromatopsia than the controls. Braun et al. (1989) speculated that color vision loss may be an early sign of neurotoxicity. These data would lead one to speculate that the color confusion observed in a subpopulation of AD patients may be an indication of the environmental insults thought to contribute to the cognitive deficits seen in AD. While the AD patients in our sample did not report extensive exposure to known toxic substances, exposure is often inferred from the major occupation of the patient. This may be misleading, in that early exposure to toxic substances may go unreported. Perhaps a detailed life and/or occupational history of the patients may have revealed some with significant levels of exposure to toxic substances. The relationship between occupational history and the deficits found in AD is important and requires further research.

In the light of the present study, the authors recommend exercising caution in interpreting Stroop Color-Word Test scores in demented older adults. Further investigation is needed to determine the extent of the color vision deficits in AD patients and what the source of the deficit is (peripheral or cortical level). Regardless of the etiology of the impairment, it is possible that color confusion could provide a means of differentiating a subgroup of patients with AD.

REFERENCES


