

# Cognitive and Neuropathologic Correlates of Stroop Color–Word Test Performance in Alzheimer’s Disease

Mark W. Bondi

Veterans Affairs San Diego Healthcare System and  
University of California, San Diego

Agnes S. Chan

The Chinese University of Hong Kong

Dean C. Delis

Veterans Affairs San Diego Healthcare System, and  
University of California, San Diego

Adam B. Serody

University of California, San Diego

Sonja C. Ebersson-Shumate

University of California, San Diego

Lawrence A. Hansen and David P. Salmon

University of California, San Diego

The Stroop Color–Word Test (SCWT; C. Golden, 1978) was examined in 59 patients with probable Alzheimer’s disease (AD) and in 51 demographically comparable normal control (NC) participants. AD patients produced significantly larger Stroop interference effects than NC participants, and level of dementia severity significantly influenced SCWT performance. Principal-components analyses demonstrated a dissociation in the factor structure of the Stroop trials between NC participants and AD patients, suggesting that disruption of semantic knowledge and speeded verbal processing in AD may be a major contributor to impairment on the incongruent trial. Results of clinicopathologic correlations in an autopsy-confirmed AD subgroup further suggest the invocation of a broad network of integrated cortical regions and executive and language processes underlying successful SCWT performance.

The Stroop Color–Word Test (SCWT; Stroop, 1935) is one of the oldest and most widely used tests in psychology for examining attention and response inhibition (see Macleod, 1991, for review). In the original form of the task, participants are timed for how quickly they can (a) read words that are the names of colors (i.e., color words), (b) name the color of ink patches, and (c) name the color of the ink in which noncongruent color words are printed (i.e., say “red” when the word *green* is printed in red ink). The effects of response inhibition are indicated by participants’ slower response times when naming the color of the ink of noncongruent color words than when reading words that are the

names of colors or naming the color of ink patches. This response slowing presumably occurs because activation of the noncongruent color word interferes with the production of the correct name of the color and must be actively inhibited.

The SCWT has been used to examine susceptibility to interference and possible deficits in response inhibition in patients with Alzheimer’s disease (AD). A number of studies have shown a deficit in SCWT performance by patients with AD, and evidence suggests that this deficit is due to a breakdown in inhibitory processes that occurs early in the course of the disease (see Binetti et al., 1996; Fisher, Freed, & Corkin, 1990; Koss, Ober, Delis & Friedland, 1984; Spieler, Balota, & Faust, 1996). In the first of these studies, Koss et al. (1984) found that AD patients performed significantly worse than normal control (NC) participants in the interference condition of the task (i.e., naming the ink color in which color words are printed), even when a corrected interference score was used that statistically controlled for group differences in simple color-naming speed. In addition, these investigators found that patients with mild dementia exhibited a greater interference effect than patients with moderate dementia, a finding they attributed to an accuracy-for-speed trade-off allowed by the more efficient self-monitoring abilities of the patients with mild dementia.

Fisher and colleagues (1990) compared the performances of AD patients and NC participants on a standardized version of the SCWT that used a 45-s time limit for each trial (Golden, 1978). Rather than recording reaction time (RT) for each item, as in the study by Koss et al. (1984), this version of the task simply required recording the number of correct responses during the time interval on each trial. The

---

Mark W. Bondi and Dean C. Delis, Psychology Service, Veterans Affairs San Diego Healthcare System and Department of Psychiatry, University of California, San Diego; Adam B. Serody and Sonja C. Ebersson-Shumate, Department of Psychiatry, University of California, San Diego; Agnes S. Chan, Department of Psychology, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong; Lawrence A. Hansen, Department of Neurosciences and Department of Pathology, University of California, San Diego; David P. Salmon, Department of Neurosciences, University of California, San Diego.

This work was supported by National Institute on Aging Grants AG12674 and AG05131 and by funds from the Medical Research Service of the Department of Veterans Affairs. We thank staff and participants of the Alzheimer’s Disease Research Center at the University of California, San Diego.

Correspondence concerning this article should be addressed to Mark W. Bondi, Psychology Service (116B), Veterans Affairs San Diego Healthcare System, 3350 La Jolla Village Drive, San Diego, California 92161. E-mail: mbondi@ucsd.edu

number of correct responses could be compared across trials to evaluate reading efficiency, color-naming efficiency, and interference, respectively, and to minimize the potential confound of age-related slowing on the interference score (Birren, 1964). Using this procedure, Fisher et al. found that AD patients were significantly impaired on all three trials compared with NC participants. Fisher and colleagues also found a significant negative correlation between the SCWT interference scores and the scores on a measure of dementia severity, the Blessed Dementia Scale.

Spieler et al. (1996) used a modified, RT-based version of the SCWT to compare the performances of young NC participants, older NC participants, and AD patients with very mild to mild dementia. These investigators found that patients with AD had greater interference scores than the older NC participants; however, the increase in response time from the color-naming trial to the interference trial exhibited by the AD patients was proportional to that of the older NC participants after general RT was taken into account. Despite this similarity in the interference effect, patients with AD produced more intrusion errors than the NC participants during the interference trial, suggesting that they were more likely than the NC participants to let the irrelevant dimensions of the stimulus drive their responses. This propensity of AD patients to make intrusion errors increased with increasing dementia severity.

Taken together, the results of these studies are consistent in demonstrating that the SCWT performance of mildly demented patients with AD is impaired and characterized by deficient response inhibition. This conclusion is in agreement with the more general notion that deficits in executive functions are an early manifestation of AD (Albert, 1996; Binetti et al., 1996; Lafleche & Albert, 1995). It is less clear from these studies, however, whether SCWT performance is sensitive to dementia severity in AD. The studies provide conflicting results in this regard, perhaps because of relatively small sample sizes and differences in methodology. It is also unclear how SCWT performance relates to the pathology of AD, as none of these prior studies examined cognitive or neuropathologic correlates of performance in patients with AD.

Given the association between SCWT performance and frontal activation that has been observed across a variety of brain imaging studies (see Brown et al., 1999, for review), one might expect that the magnitude of the performance deficit in AD patients will correlate with the severity of the frontal cortex pathology that occurs in the disease. It is likely, however, that the deficits in performing the relatively complex attentional and inhibitory processes engaged by the SCWT are mediated by more than a single region of pathology in patients with AD. As Albert and colleagues have suggested (Albert, 1996; Lafleche & Albert, 1995), the early manifestations of a general executive dysfunction in AD may be related to direct frontal neocortical damage as well as to neuropathologic changes either (a) in the basal forebrain, which modulates frontal function; (b) in long cortico-cortical projection systems, which can result in executive function deficits that are not anatomically based in the frontal lobes; or (c) in parietal structures, which are thought

to mediate the disengagement and switching components of complex attention (Filoteo et al., 1995; Posner, Walker, Friedrich, & Rafal, 1984). Recent functional neuroimaging evidence by Brown and colleagues (1999) supported the notion that the neural substrata involved in performing the SCWT do not encompass a single brain region. Rather, a network of brain regions that includes portions of both frontal and parietal cortex appears to be required.

Deficits in a variety of cognitive processes may also contribute to the poor SCWT performance of patients with AD. In addition to the obvious contribution of executive dysfunction, the disruption of semantic knowledge that occurs in AD may promote the SCWT deficit. This possibility is supported by a recent study by Luo (1999) that suggested semantic competition, rather than simply response inhibition, may be the major source of interference when performing the interference task. Luo simultaneously showed a colored bar (e.g., a red bar) and a colored word (e.g., *red* printed in red or blue) to participants and had them decide (a) whether the two items had the same meaning (meaning decisions) or (b) whether the two items had the same surface color (visual decisions). Interference was observed during the meaning-decision task when the word's visual information conflicted with the visual information in the bar (e.g., *red* printed in blue, against a red bar). In contrast, no interference was observed during the visual-decision task when the word's meaning conflicted with the meaning of the colored bar (e.g., *blue* printed in red, against a red bar). When an interval was imposed between presentation of the colored bar and the colored word, interference in meaning decisions was diminished but interference in visual decisions developed. According to Luo, these results occurred because the simultaneous presentation of the to-be-named color and the to-be-ignored word concurrently activated two similar semantic codes that gave rise to semantic competition. A similar explanation for the interference effect was provided by Klopfer (1996), who found that the semantic similarity between relevant and irrelevant dimensions of Stroop stimuli determined the amount of interference that was produced. If this explanation for the interference effect is correct, dysfunction in the semantic system could contribute to the poor SCWT performance of patients with AD. Unfortunately, the contributions of various cognitive deficits to the SCWT impairment exhibited by patients with AD remains unknown because previous studies have not systematically examined the broad-based cognitive correlates of SCWT performance.

Therefore, given the current state of knowledge regarding SCWT performance in patients with AD, the present study had five goals: (a) confirm previous studies that have shown that the SCWT is sensitive to inhibitory processing deficits early in the course of AD; (b) determine the degree to which the SCWT is sensitive to the progression of dementia in AD; (c) determine the degree to which the SCWT accurately classifies and differentiates AD patients from normally aging individuals; (d) identify the relationship between SCWT performance and performance on tests of a wide range of cognitive abilities in patients with AD; and (e) examine the contribution of neuritic plaque and neuro-

fibrillary tangle pathology in several neocortical regions and the hippocampal formation to SCWT performance in a sample of autopsy-confirmed AD patients.

## Method

### Participants

One hundred ten individuals participated in this study: 51 NC participants (20 men, 31 women) and 59 patients with the clinical diagnosis of probable AD (31 men, 28 women). All participants were part of a larger cohort participating in the Alzheimer's Disease Research Center of the University of California, San Diego (UCSD ADRC). Demographic characteristics of the two groups are shown in Table 1.

Participants were selected without regard to ethnicity or race. Written informed consent was obtained from all participants (or their conservators) after the protocol of the study had been fully explained. At the time of testing, two senior staff neurologists diagnosed AD in 57 of the 59 cases according to the criteria developed by the National Institute of Neurological and Communications Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA; McKhann et al., 1984); two of the 59 cases who were diagnosed at risk for AD by virtue of significant memory impairment without functional impairment converted to AD the following year, and later received pathologic confirmation of AD at autopsy. The NC participants either were spouses of the patients or were volunteers obtained through newspaper advertisements or community lectures. Volunteers with a history of alcoholism, drug abuse, learning disability, or neurologic or severe psychiatric illness were excluded. The NC participants and the AD patients did not differ in terms of age ( $t < 1$ ); education,  $t(108) = 1.77, p = .08$ ; or gender,  $\chi^2(1, N = 110) = 1.45, p = .23$ . As expected, AD patients significantly differed from NC participants on the Mattis Dementia Rating Scale (DRS; Mattis, 1988),  $t(108) = 16.22, p < .001$ . To examine the performance of AD subgroups on the SCWT by level of dementia severity, we divided AD patients on the basis of their DRS scores into very mild (119–139;  $n = 22$ ), mild (106–118;  $n = 25$ ), and moderate (90–105;  $n = 12$ ) subgroups, such that the resultant subgroups were separated roughly equally by their mean DRS total scores (very mild AD  $\approx 125$ , mild AD  $\approx 112$ , moderate AD  $\approx 99$ ). Thus,

given these DRS score ranges, we achieved roughly equal separation of very mild from mild and mild from moderate AD subgroups at approximately 13 points (with standard deviations within each group ranging from 4 to 6 points).

To explore the relationship between SCWT performance and brain pathology, we examined separately the results from 25 of the 59 AD patients who subsequently died and came to autopsy. The average age of the autopsied AD cases was 74.16 years ( $SD = 6.84$  years), and the average interval between test administration and death was 4.80 years ( $SD = 1.72$  years; range = 1.2–8.5 years). Neuropathologic procedures used at the UCSD ADRC have been described in detail previously (Terry, Peck, DeTeresa, Schechter, & Horoupian, 1981). Briefly, the left hemisphere was fixed in 10% formalin for 5–7 days from which paraffin-embedded sections were made and stained with both hematoxylin-and-eosin and thioflavin S. Total neuritic plaque (NP) and neurofibrillary tangle (NFT) counts were then performed on tissue blocks taken from the midfrontal (MDFT), inferior parietal (IFPR), rostral superior temporal (SPTP), and hippocampal (HIPPO) regions (for more detailed procedures, refer to Terry et al., 1981).

### Materials

The standardized version of the SCWT was used (Golden, 1978). The SCWT consisted of three pages, each having 100 items, organized in five columns of 20. Page 1 consisted of the words *red*, *green*, and *blue* printed in black ink and randomly arranged on a white sheet of paper measuring  $8.5 \times 11$  in. No word followed itself within a column.

Page 2 also consisted of 100 items arranged in the same way as page 1, with the difference being that the items were all written as XXXX in either red, green, or blue ink. No color followed itself or matched the corresponding item on page 1.

Page 3 consisted of the words on page 1 (*red*, *green*, and *blue*) printed in the color of the items (XXXX) on page 2. This was done for all 100 items. None of the words matched the color in which they were printed.

### Procedure

All participants were administered the test individually as part of a larger neuropsychological test battery at the UCSD ADRC.

Table 1  
Means and Standard Deviations for Performance on the Stroop Color-Word Test

Variable	NC <sup>a</sup>		Very mild AD <sup>b</sup>		Mild AD <sup>c</sup>		Moderate AD <sup>d</sup>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age (years)	73.40	7.43	71.70	6.97	73.50	5.76	72.10	8.41
Education	13.90	2.97	13.90	3.24	12.40	3.81	11.80	2.45
DRS score	140.30	3.33	125.50	5.97	112.00	4.07	99.30	4.00
Stroop raw scores								
Word reading	79.00	16.75	70.30	22.86	65.20	16.05	54.80	26.56
Color naming (C)	53.70	13.90	44.20	14.39	34.40	13.60	32.90	17.33
Color-Word (CW)	28.80	10.27	18.10	6.83	12.10	7.65	8.60	6.75
(C - CW)/C	0.45	0.19	0.59	0.21	0.61	0.23	0.72	0.20
Stroop errors								
Word reading	0.01	0.14	0.73	1.75	0.56	0.92	1.67	3.68
Color naming	0.80	1.18	1.77	2.64	1.60	1.35	1.58	1.68
Color-Word	0.94	1.35	2.09	2.02	2.64	1.96	2.92	3.03

Note. NC = normal control participants; AD = Alzheimer's disease; DRS = Dementia Rating Scale (Mattis, 1988; 144 points possible); (C - CW)/C = the ratio that corrects for color naming speed.

<sup>a</sup>  $n = 51$ . <sup>b</sup>  $n = 22$ . <sup>c</sup>  $n = 25$ . <sup>d</sup>  $n = 12$ .

All three pages (stacked in consecutive order) were placed in front of each participant on a flat surface. Instructions were then read to each participant, and repeated or paraphrased when necessary both before and during the test.

For the word-reading trial (page 1), participants were asked to read down the columns starting with the top word in the leftmost column; if they finished all five columns they were instructed to return to the first column and begin again. In the color-naming trial (page 2) participants were instructed to name the color of the item with all other instructions being identical to those of the previous trial. During the color-word (interference) trial (page 3) the participants were asked to read down the columns, naming the color of the words, ignoring the word that was printed in each item. All other instructions were also identical to those of the previous trials.

For all trials, the participants were asked to complete the trials as quickly as possible. If an error in reading or naming was made, the experimenter said "no," which cued the participant to correct the error and then continue. After 45 s, the experimenter said "stop" and the experimenter noted the item they last named.

### Scoring

The SCWT produces three scores: The word-reading (W) score is made up of the number of items completed in 45 s on page 1; the color-naming (C) score is made up of the number of items completed in 45 s on page 2; and the color-word (CW) or interference score is made up of the number of items completed in 45 s on page 3. Many investigators have also computed other scores based on various mathematical manipulations of the C, W, and CW scores (Golden, 1978). In this present study, a ratio that corrects for color-naming speed ( $[C - CW]/C$ ) was used as another measure of the interference effect. Examining interference effects in the context of one's color-naming ability is thought to control for performance associated with processing a color stimulus, for lexical access of color words, and for the final motor pathways activated by the articulation of color-word responses (Taylor, Kornblum, Lauber, Minoshima, & Koeppel, 1997). In addition, the number of errors committed was recorded for each of the three trials.

## Results

### Group Differences

One-way analyses of variances (ANOVAs) were used to compare SCWT scores across levels of dementia severity (see Table 1). Analyses revealed a significant main effect for group on the word reading trial,  $F(3, 106) = 6.61, p < .001$ ; color naming trial,  $F(3, 106) = 13.68, p < .001$ ; color-word (interference) trial,  $F(3, 106) = 24.63, p < .001$ ; and the corrected interference score ( $[C - CW]/C$ ),  $F(3, 106) = 7.57, p < .001$ . Scheffé's post hoc comparisons revealed that NC participants scored significantly better ( $p < .05$ ) than mild and moderate AD patients on the word reading, color naming, and the corrected interference score ( $[C - CW]/C$ ). On the color-word (interference) trial, NC participants performed significantly better than all three of the AD subgroups. However, there were no significant differences between the very mild, mild, and moderate AD subgroups on any of the three trials or on the corrected interference score. By contrast, Pearson product-moment correlations revealed significant relationships between each of the AD patients' SCWT scores and DRS total score, W:

$r(59) = .41, p = .001$ ; C:  $r(59) = .41, p = .001$ ; CW:  $r(59) = .56, p < .001$ ;  $([C - CW]/C)$ :  $r(59) = -.34, p = .009$ , indicating at least a mild degree of association between dementia severity and SCWT performance.

### Error Analyses

One-way ANOVAs were used to examine differences between NC participants and AD patients on the number of errors committed during the three Stroop trials. Significant main effects emerged on the word-reading trial,  $F(3, 106) = 4.44, p < .001$ , and the color-word trial,  $F(3, 106) = 6.89, p < .001$ , but not on the color-naming trial,  $F(3, 106) = 2.50, p > .05$ . Scheffé's post hoc analyses revealed that NC participants committed significantly fewer errors than moderate AD patients on the word-reading trial ( $p < .05$ ) and significantly fewer errors than mild and moderate AD patients during the color-word trial ( $p < .05$ ). The three AD subgroups did not significantly differ on any of the trials (see Table 1).

### Logistic Regression

A stepwise logistic regression (with forward selection) was performed to examine the use of the SCWT in differentiating between AD patients and NC participants. The group variable was dementia status (either NC or AD) and the predictors were gender, age, years of education, and the three SCWT trials (word-reading, color-naming, and color-word trials). The final model was significant,  $\chi^2(5, N = 110) = 46.91, p < .001$ , and included only the Stroop color-word (interference) score,  $\beta = -.14, R = -.42, p < .001$ ; none of the other variables were retained in the final model. Overall correct classification rate was 81%, with 13 out of 51 NC participants and 8 out of 59 AD patients misclassified.

### Principal-Components Analyses (PCAs)

The cognitive abilities contributing to SCWT performance were explored with PCA that included scores from the larger battery of neuropsychological tests administered to participants at the UCSD ADRC. PCAs with varimax rotation were performed separately for the 51 older NC participants and the 59 AD patients. Four factors were extracted in each PCA, and their factor loadings, communalities, and the percent of variance accounted for by each factor are shown in Table 2. Briefly, results revealed a different factor structure for NC participants and AD patients. For AD patients, the three trials of the SCWT loaded together on a factor that also included the Trail Making Test (Part A), the Wechsler Adult Intelligence Scale—Revised (WAIS-R) Digit Symbol, and the Wechsler Intelligence Scale for Children (WISC) Block Design. This factor appears to represent speeded visual processing and visuomotor sequencing and is distinct from factors that appear to reflect (a) executive functions (Wisconsin Card Sorting Test [WCST] Categories, WCST Perseverative Errors, Trail Making Test [Part B]), (b) semantic knowledge and verbal

Table 2

*Normal Control Participants' and Alzheimer's Disease Patients' Factor Loadings, Communalities ( $h^2$ ), and Percentage of Variance for Principal-Components Analysis With Varimax Rotation on a Neuropsychological Test Battery*

Test battery	Normal control participants					Alzheimer's disease patients				
	Factor 1	Factor 2	Factor 3	Factor 4	$h^2$	Factor 1	Factor 2	Factor 3	Factor 4	$h^2$
Boston Naming Test	.44	<b>.60</b>	-.10	.34	.67	.21	-.03	<b>.82</b>	-.05	.72
Letter fluency (FAS)	.18	<b>.68</b>	.29	.10	.59	.03	.27	<b>.69</b>	.29	.63
Stroop word reading (W)	<b>.70</b>	.18	-.09	.16	.57	<b>.74</b>	-.03	.40	-.08	.70
Stroop color naming (C)	<b>.70</b>	.37	.19	.19	.67	<b>.67</b>	-.09	.24	.30	.60
Stroop CW interference	.37	<b>.55</b>	.25	.25	.54	<b>.57</b>	.43	.17	.32	.64
Trails A (s)	<b>-.82</b>	.03	.19	-.05	.71	<b>-.85</b>	-.03	-.05	.08	.73
Trails B (s)	<b>-.75</b>	.08	.40	-.20	.77	<b>-.54</b>	<b>-.62</b>	-.02	-.03	.67
WAIS-R digit span	.22	<b>.65</b>	.32	.04	.57	.27	.36	.48	.45	.64
WAIS-R digit symbol	<b>.81</b>	.13	.30	.10	.78	<b>.64</b>	.45	.16	.04	.64
WAIS-R vocabulary	-.09	<b>.84</b>	.03	.18	.73	.11	.12	<b>.86</b>	.19	.81
WCST categories	.28	.27	<b>.87</b>	.11	.91	.31	<b>.83</b>	.13	.17	.82
WCST perseverative errors	-.24	-.18	<b>-.87</b>	-.07	.85	.13	<b>-.82</b>	-.12	.03	.70
WISC block design	<b>.76</b>	.27	.22	-.01	.71	<b>.77</b>	.42	-.09	-.06	.78
WMS-R delayed recall	.06	.20	.08	<b>.94</b>	.94	-.09	.14	.02	<b>.94</b>	.90
WMS-R immediate recall	.17	.16	.10	<b>.94</b>	.93	.12	-.01	.29	<b>.88</b>	.90
Eigenvalues	6.51	1.92	1.40	1.14		5.61	2.35	1.64	1.27	
Percent of variance	43.40	12.80	9.20	7.60		37.40	15.70	10.90	8.50	

*Note.* Bold type represents the factor loadings of the Stroop Color-Word Test components for the word-reading, color-naming, and color-word (interference) trials. FAS = the letters *F*, *A*, and *S*; W = word-reading score; C = color-naming score; CW = color-word interference score; Trails A = Trail Making Test (Part A); Trails B = Trail Making Test (Part B); s = seconds; WAIS-R = Wechsler Adult Intelligence Scale—Revised; WCST = Wisconsin Card Sorting Test; WISC = Wechsler Intelligence Scale for Children; WMS-R = Wechsler Memory Scale—Revised.

processing speed (Boston Naming Test, WAIS-R Vocabulary, Letter Fluency), and (c) episodic memory processes (Wechsler Memory Scale—Revised [WMS-R] Logical Memory Immediate and Delayed Recall), respectively. For NC participants, the word-reading and color-naming trials loaded with the Trail Making Test (Parts A and B), WAIS-R Digit Symbol, and WISC Block Design on a speeded visual processing/visuomotor sequencing factor; however, the color-word (interference) trial loaded with the Boston Naming Test, Letter Fluency, WAIS-R Digit Span, and WAIS-R Vocabulary on a factor that appears to represent semantic knowledge and verbal processing speed. SCWT components did not load on either executive function or episodic memory factors for the NC participants.

### *Clinicopathologic Comparisons*

The correlation between scores on the SCWT and the NP and NFT counts in three neocortical regions and the hippocampus were carried out for 25 AD patients who came to autopsy. Because plaque and tangle counts for each of the four tissue blocks were capped at 50 per field, we performed analyses using Spearman's rank-order correlation coefficients. Results revealed no significant correlations between SCWT scores and NP counts in any of the four regions sampled. NFT counts, however, were significantly correlated with Stroop W and CW scores in each of the four cortical regions, although no significant correlations were demonstrated with C scores (see Table 3). It should also be noted that no significant correlations were demonstrated between DRS total score and either NP or NFT counts in any of the four regions sampled (Table 3).

A second analysis involved dividing the AD patients into low- and high-performance groups on the basis of a median split for the word-reading, color-naming, and color-word (interference) trials and comparing the groups on plaque and tangle counts in each sampled brain region. Possible differences in the number of days that elapsed between SCWT administration and patient death were examined with independent samples *t* tests comparing low and high performers on each of the SCWT trials. It is important to note that no significant differences in time interval between testing and death emerged between groups for the word-reading, color-naming, and color-word (interference) trials, respectively,  $t(23) = 1.16, p = .26$ ;  $t(23) = 1.05, p = .31$ ;  $t(23) = 1.93, p = .06$ . Because plaque and tangle counts for each of the four tissue blocks were capped at 50 per field, group comparisons for these measures were performed using nonparametric Mann-Whitney *t* tests.

As shown in Table 4, when groups were formed on the basis of the word reading trial, no significant differences emerged between high and low performers on total plaque counts taken from the MDFT, IFPR, SPTP, and HIPP regions. However, low performers had significantly more tangles than high performers in each of the four regions examined. When groups were formed on the basis of the color naming trial, no significant differences were found between high and low performers on total plaque counts or total tangle counts taken from the MDFT, IFPR, SPTP, and HIPP regions. Finally, when groups were formed on the basis of the color-word (interference) trial, no significant differences emerged between high and low performers on total plaque counts taken from the MDFT, IFPR, SPTP, and

Table 3  
*Spearman Rank-Order Correlation Coefficients ( $\rho$ ) of Total Plaque and Neurofibrillary Tangle Counts on Dementia Rating Scale (DRS) and Stroop Word Reading, Color Naming, and Color-Word Interference Trials in 25 Autopsy-Confirmed Alzheimer's Disease Patients*

Variable	Hippocampal		Inferior parietal		Midfrontal		Superior temporal	
	$\rho$	$p$	$\rho$	$p$	$\rho$	$p$	$\rho$	$p$
Plaque counts								
Word reading	-.02	.93	-.34	.10	-.31	.15	-.25	.24
Color naming (C)	-.18	.38	-.27	.19	-.17	.44	-.17	.44
Color-Word (CW)	-.23	.28	-.35	.10	-.27	.20	-.32	.12
(C - CW)/C	.10	.63	.08	.72	.24	.27	.26	.22
DRS total score	-.20	.34	-.08	.70	-.15	.49	-.01	.95
Neurofibrillary counts								
Word reading	-.53	<b>.007</b>	-.50	<b>.01</b>	-.55	<b>.005</b>	-.57	<b>.004</b>
Color naming	-.02	.93	-.37	.07	-.22	.31	-.14	.50
Color-Word	-.43	<b>.03</b>	-.51	<b>.01</b>	-.45	<b>.03</b>	-.43	<b>.03</b>
(C - CW)/C	.52	<b>.007</b>	.21	.32	.29	.18	.40	<b>.05</b>
DRS total score	-.23	.26	-.28	.18	-.26	.22	-.20	.36

*Note.* Bold type indicates significant correlations. (C - CW)/C = the ratio that corrects for color-naming speed.

HIPP regions. However, low performers had significantly more tangles than high performers in the IFPR, SPTP, and HIPP regions, and increased tangle counts approached statistical significance in the MDFT region ( $p = .10$ ).

### Discussion

Results of this study indicate that AD patients perform worse than healthy older adults on the congruent (word reading and color naming) and incongruent (color-word [interference]) trials of the SCWT. Moreover, the impaired performance of the AD patients on the incongruent trial remains evident even after baseline color-naming ability is taken into account. These results are consistent with a number of previous studies (Binetti et al., 1996; Fisher et al. 1990; Koss et al., 1984; Spieler et al., 1996) and suggest that even mildly demented AD patients have some degree of psychomotor slowing and a deficit in inhibitory processes that impairs their ability to disregard the irrelevant dimension of the incongruent stimulus. The results also support previous contentions that this and other forms of executive dysfunction may be a particularly sensitive marker of early AD (see Albert, 1996; Binetti et al., 1996; Lafleche & Albert, 1995).

Whereas the SCWT was found to be rather insensitive to dementia severity in our blocked AD subgroup analyses, simple bivariate correlational analyses showed significant, though relatively modest, associations between decline on the SCWT and increasing dementia severity. These somewhat discrepant results may be related to the availability of less statistical power in the blocked analyses than in the bivariate correlational analyses. Despite the significant association, however, the modest size of the correlations between SCWT performance and dementia severity suggests

that, like rigorous tests of episodic memory and executive functions (see Albert, Moss, Tanzi, & Jones, 2001; Bondi, Salmon, Galasko, Thomas, & Thal, 1999), the test may be less useful in tracking disease progression into the more advanced stages of AD than for early detection of the disorder.

The results of the PCAs demonstrated a great deal of similarity in the factor structures of the AD and NC groups. Both PCAs demonstrated a four-factor solution with all but a few variables loading on the same or analogous factors. Nevertheless, an interesting dissociation in the loading of the three SCWT variables was observed between AD and NC groups. Although all three SCWT components loaded on the same factor for AD patients, NC participants revealed a dissociation among the SCWT components, with scores on the word-reading and color-naming trials loading on a factor with tasks requiring organization and information processing speed and scores on the color-word (interference) trial loading on a factor primarily representing semantic knowledge and verbal-processing speed. This result is consistent with the notion that the interference effect may be largely driven by semantic competition (see Luo, 1999, for a discussion of this issue). Effective performance on the interference trial may require the activation of the semantic and lexical representation of the color name and the inhibition of overlearned reading skills that promote reading the incongruent word.

Deficits in both of these processes may underlie the poorer speed and accuracy of the patients with AD on the interference trial. That is, AD patients may not activate the representation of the color name to a normal level because of their deficits in the semantic system (Salmon & Bondi, 1999), and they may not be able to inhibit generation of the

Table 4  
*Descriptive Statistics and p Values on Total Plaque and Total Tangle Counts for Low Performers and High Performers on Word-Reading, Color-Naming, and Color-Word Interference Trials*

Variable	Low Stroop performers					High Stroop performers					<i>p</i>
	<i>n</i>	Min.	Max.	<i>M</i>	<i>SD</i>	<i>n</i>	Min.	Max.	<i>M</i>	<i>SD</i>	
Word-reading trial											
Total plaques MDFT	13	29	50	47.23	6.18	11	35	50	46.91	5.84	<i>ns</i>
Total plaques IFPR	12	28	50	46.67	6.83	12	30	50	44.50	6.29	<i>ns</i>
Total plaques SPTP	13	28	50	44.38	8.01	11	22	50	42.64	9.73	<i>ns</i>
Total plaques HIPP	13	4	33	14.69	9.26	12	8	25	14.08	5.02	<i>ns</i>
Total tangles MDFT	13	0	11	4.15	3.24	11	0	6	1.64	2.38	.04
Total tangles IFPR	12	0	14	6.00	4.59	12	0	5	2.50	1.93	.05
Total tangles SPTP	13	0	18	6.92	4.72	11	0	8	3.27	2.65	.03
Total tangles HIPP	13	2	49	24.77	15.09	12	1	32	10.25	8.16	.01
Color-naming trial											
Total plaques MDFT	13	29	50	47.23	6.18	11	35	50	46.91	5.84	<i>ns</i>
Total plaques IFPR	13	28	50	46.62	6.49	11	30	50	44.36	6.64	<i>ns</i>
Total plaques SPTP	13	28	50	44.77	8.05	11	22	50	42.18	9.57	<i>ns</i>
Total plaques HIPP	13	4	33	15.08	8.94	12	8	25	13.67	5.53	<i>ns</i>
Total tangles MDFT	13	0	11	3.54	3.36	11	0	6	2.36	2.77	<i>ns</i>
Total tangles IFPR	13	0	14	5.54	4.35	11	0	9	2.73	2.69	<i>ns</i>
Total tangles SPTP	13	0	13	5.38	3.38	11	0	18	5.09	5.28	<i>ns</i>
Total tangles HIPP	13	2	39	17.85	12.95	12	1	49	17.75	15.86	<i>ns</i>
Color-word (interference) trial											
Total plaques MDFT	15	41	50	49.00	2.70	9	29	50	43.89	8.30	<i>ns</i>
Total plaques IFPR	14	40	50	47.71	3.89	10	28	50	42.60	8.34	<i>ns</i>
Total plaques SPTP	14	30	50	46.21	6.19	10	22	50	39.90	10.56	<i>ns</i>
Total plaques HIPP	15	5	33	14.93	8.21	10	4	25	13.60	6.28	<i>ns</i>
Total tangles MDFT	15	0	11	3.80	3.19	9	0	6	1.67	2.55	<i>ns</i>
Total tangles IFPR	14	0	14	5.71	4.30	10	0	5	2.20	1.93	.03
Total tangles SPTP	14	2	18	6.93	4.32	10	0	8	2.90	2.96	.01
Total tangles HIPP	15	3	49	22.60	15.04	10	1	32	10.60	9.12	.03

*Note.* Min. = minimum; Max. = maximum; MDFT = midfrontal; IFPR = inferior parietal; SPTP = superior temporal; HIPP = hippocampal; *ns* = nonsignificant.

incongruent word because of frontally mediated executive dysfunction (Binetti et al., 1996; Bondi, Monsch, Butters, Salmon, & Paulsen, 1993). The finding that all three components of the SCWT performance of AD patients loaded on a single factor that appears to reflect information-processing speed (see also Boone, Ponton, Gorsuch, Gonzalez, & Miller, 1998) and not on a language factor that included measures of verbal fluency, confrontation naming, and word definitions, or on an executive functioning factor that included measures from the Trail Making Test and the WCST, is consistent with this possibility. In other words, for normally aging individuals, it appears that word reading and color naming rely more heavily on speeded visual processing, whereas successful completion of the interference trial necessitates reliance on the semantic system. These findings agree with Luo's (1999) supposition that word reading and color naming place demands on the verbal-lexical system (i.e., a knowledge base consisting of units representing the orthographic and phonological features of words—not their meanings), whereas the interference trial invokes the semantic system (i.e., a knowledge base consisting of units representing meanings or concepts of words and objects). In

AD, however, disruption of the semantic system and decrements in verbal-processing speed force greater reliance on visual processing in order to complete the interference trial.

On the basis of findings from recent functional imaging studies (Bench et al., 1993; Brown et al., 1999; Pardo, Pardo, Janer & Raichle, 1990; Taylor et al., 1997; Taylor, Kornblum, Minoshima, Oliver & Koeppel, 1994), including their own, Brown et al. (1999) proposed an integrated model of brain activation that underlies SCWT performance in normal individuals. In this model they proposed that, because mature readers such as adults must learn to attend to the lexical features of words across different types of scripts, individuals are biased to attend to the lexical features of the incongruent stimuli rather than to the color features. Shifting attention to color features is thought to activate neuronal pools in the right parietal lobe involved in selective visual attention. However, because the lexical features of the incongruent Stroop stimuli are automatically activated, two codes, one representing lexical features and the other representing color features, are activated for further processing.

Luo (1999) suggested that color naming involves the activation of the semantic system and the mapping of semantic information into its corresponding verbal information in the verbal-lexical system, whereas word reading requires only the activation of the lexical-verbal system, and the involvement of the semantic system is optional. Brown et al. (1999) indicated that the premotor anterior cingulate system biases the network toward selecting the motor program associated with stimulus color. Despite this biasing, the tendency for the lexical features to activate the motor representation of the word name is not entirely suppressed. Thus, another aspect in resolving the conflict is thought to involve programming the sequence of articulation associated with the code that exceeds its threshold first, and Brown et al. suggested that the competitive activation of speech codes increases neuronal activity in the pars opercularis (Brodmann's Area 44). Within this system, there is no single location where inhibition of the Stroop color-word response occurs. Rather, they suggested the system activates the color-name code and progressively promotes the likelihood of its motor expression. Thus, in such a scheme, diffuse neuropathologic changes in a number of brain regions would be expected to be associated with poor SCWT performance, as was demonstrated in our analyses with the subsample of autopsy-confirmed AD patients.

Despite a rather lengthy interval between the time of SCWT performance and autopsy, performance on the word-reading and interference trials of the test was significantly associated with AD pathology in the hippocampus and a number of neocortical regions. The significant clinicopathologic relationships are all the more impressive if one considers that SCWT performances were also influenced by the level of dementia severity, which one could argue might serve to diminish possible associations given the advanced stage of the disease by the time the patients came to autopsy. These associations were limited to NFT pathology, consistent with several previous studies that have shown that NFT pathology is a stronger correlate of cognitive dysfunction in AD patients than is NP pathology (Terry et al., 1991; see also Terry, Masliah, & Hansen, 1999, for review), although neither NFTs nor NPs appear to be as strong a correlate as does synapse density. For example, Terry et al. (1991) found a particularly strong correlation in AD between the DRS total score and synapse density in the MDFT region ( $r = .65$ ), and Samuel, Terry, DeTeresa, Butters, and Masliah (1994) have shown that 50% of the variance in Mini-Mental State Examination (MMSE) score is related to MDFT synapse density, and 24% of the variance is due to NFTs in the basal nucleus of Meynert.

The finding that SCWT performance is associated with AD pathology in a wide array of neocortical sites is consistent with current models (e.g., Brown et al., 1999) that propose that the task is mediated by a distributed cortical network that involves activation and inhibition of semantic, lexical, and motor (i.e., speech) processes. Although broadly based, these associations appear to be somewhat selective given that measures of AD pathology were not

significantly associated with a broad index of dementia severity (i.e., the DRS total score) in this study. The failure to find significant relationships between AD pathology and DRS performance may have occurred because the test has some subcomponents (e.g., basic attentional skills, basic construction skills, basic conceptualization skills) that are not particularly sensitive to the early cognitive dysfunction associated with AD (Monsch et al., 1995).

There has been some concern that the standardized version (Golden, 1978) of the SCWT imposes restrictions on the manner in which stimuli are presented during the test, particularly with regard to conditions being blocked (i.e., congruent vs. incongruent) rather than randomized (i.e., alternating congruent and incongruent trials; Spieler et al. 1996; Tzelgov, Henik, & Berger, 1992). Spieler et al., for example, argued that blocked trials might allow participants to develop strategies to focus on the relevant dimension of the stimuli. It could be argued, however, that this strengthens the test's sensitivity to dementia because such strategies used by controls are less likely to be generated by patients. A difference in the ability to use internally generated strategies to facilitate performance is supported by the different factor structures obtained for the NC and AD groups in this study. Another potential advantage of the blocked procedure is that it minimizes memory demands that are inherent in a randomized trial, thereby reducing potential confounds with the memory deficit of AD. It may be the case, however, that the blocked version of the task is easier than the randomized computer-based version and that this difference accounts for the diminution of an effect of dementia severity whereas a stronger effect is evident on the randomized version (cf. Fisher et al., 1990).

Further efforts might concentrate not only on the differences seen in performance between nondemented older adults and AD patients, but also on possible differences in inhibitory processing between cortical dementia (e.g., Alzheimer's disease) and subcortically based dementias (e.g., Parkinson's and Huntington's disease). For example, one functional magnetic resonance imaging study, using a conceptual reasoning task as a measure of executive functioning, has demonstrated not only cortical activation of the anterior cingulate and frontal regions, but also subcortical activation of sites primarily involving the right basal ganglia, right thalamus and left lateral cerebellum (Rao et al., 1997). Clues from differences in inhibitory processing among these groups may aid in a better understanding of the underlying neuroanatomic correlates to executive dysfunction in various forms of dementia.

## References

- Albert, M. S. (1996). Cognitive and neurobiologic markers of early Alzheimer's disease. *Proceedings of the National Academy of Science, USA*, 93, 13547-13551.
- Albert, M. S., Moss, M. B., Tanzi, R., & Jones, K. (2001). Preclinical prediction of AD using neuropsychological tests. *Journal of the International Neuropsychological Society*, 7, 631-639.

- Bench, C. J., Frith, C. D., Grasby, P. M., Friston, K. J., Paulesu, E., Frackowiak, R. S. J., & Dolan, R. J. (1993). Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia*, *31*, 907–922.
- Binetti, G., Magni, E., Padovani, A., Cappa, S. F., Bianchetti, A., & Trabucchi, M. (1996). Executive dysfunction in early Alzheimer's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *60*, 91–93.
- Birren, J. (1964). *The psychology of aging*. Englewood Cliffs, NJ: Prentice Hall.
- Bondi, M. W., Monsch, A. U., Butters, N., Salmon, D. P., & Paulsen, J. S. (1993). Utility of a modified version of the Wisconsin Card Sorting Test in the detection of dementia of the Alzheimer type. *The Clinical Neuropsychologist*, *7*, 161–170.
- Bondi, M. W., Salmon, D. P., Galasko, D., Thomas, R. G., & Thal, L. J. (1999). Neuropsychological function and Apolipoprotein E genotype in the preclinical detection of Alzheimer's disease. *Psychology and Aging*, *14*, 295–303.
- Boone, K. B., Ponton, M. O., Gorsuch, R. L., Gonzalez, J. J., & Miller, B. L. (1998). Factor analysis of four measures of prefrontal lobe functioning. *Archives of Clinical Neuropsychology*, *13*, 585–595.
- Brown, G. G., Kindermann, S. S., Siegal, G., Granholm, E., Wong, E. C., & Buxton, R. B. (1999). Brain activation and pupil response during covert performance of the Stroop Color Word Task. *Journal of the International Neuropsychological Society*, *5*, 308–319.
- Filoteo, J. V., Delis, D. C., Roman, M. J., Demadura, T. L., Ford, E., Butters, N., et al. (1995). Visual attention and perception in patients with Huntington's disease: Comparisons with other subcortical and cortical dementias. *Journal of Clinical and Experimental Neuropsychology*, *17*, 654–667.
- Fisher, L. M., Freed, D. M., & Corkin, S. (1990). Stroop Color-Word Test performance in patients with Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, *12*, 745–758.
- Golden, C. (1978). *Stroop color and word test manual* (Cat. 30150M). Chicago: Stoelting Co.
- Klopfer, D. S. (1996). Stroop interference and color-word similarity. *Psychological Science*, *7*, 150–157.
- Koss, E., Ober, B., Delis, D., & Friedland, R. (1984). The Stroop Color-Word Test: Indicator of dementia severity. *International Journal of Neuroscience*, *24*, 53–61.
- Lafleche, G., & Albert, M. S. (1995). Executive function deficits in mild Alzheimer's disease. *Neuropsychology*, *9*, 313–320.
- Luo, C. R. (1999). Semantic competition as the basis of Stroop interference: Evidence from color-word matching tasks. *Psychological Science*, *10*, 35–40.
- Macleod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203.
- Mattis, S. (1988). *Dementia Rating Scale: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology*, *34*, 939–944.
- Monsch, A. U., Bondi, M. W., Salmon, D. P., Butters, N., Thal, L. J., Hansen, L. A., et al. (1995). Clinical validity of the Mattis Dementia Rating Scale in detecting dementia of the Alzheimer type: A double cross-validation and application to a community-dwelling sample. *Archives of Neurology*, *52*, 899–904.
- Pardo, J. V., Pardo, P. J., Janer, K. W., & Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Science, USA*, *87*, 256–259.
- Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. (1984). The effects of parietal injury on covert orienting of attention. *Journal of Neuroscience*, *4*, 1863–1874.
- Rao, S. M., Bobholz, J. A., Hammeke, T. A., Rosen, A. C., Woodley, S. J., Cunningham, J. M., et al. (1997). Functional MRI evidence for subcortical participation in conceptual reasoning skills. *Neuroreport*, *8*, 1987–1993.
- Salmon, D. P., & Bondi, M. W. (1999). Neuropsychology of Alzheimer disease. In R. D. Terry, R. Katzman, K. L. Bick, & S. S. Sisodia (Eds.), *Alzheimer disease* (2nd ed, pp. 39–56). New York: Lippincott, Williams & Wilkins.
- Samuel, W., Terry, R. D., DeTeresa, R., Butters, N., & Masliah, E. (1994). Clinical correlates of cortical and nucleus basalis pathology in Alzheimer dementia. *Archives of Neurology*, *51*, 772–778.
- Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 461–479.
- Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–661.
- Taylor, S. F., Kornblum, S., Lauber, E. J., Minoshima, S., & Koeppe, R. A. (1997). Isolation of specific interference processing in the Stroop task: PET activation studies. *Neuroimage*, *6*, 81–92.
- Taylor, S. F., Kornblum, S., Minoshima, S., Oliver, L. M., & Koeppe, R. (1994). Changes in medical cortical blood flow with a stimulus-response compatibility task. *Neuropsychologia*, *32*, 249–255.
- Terry, R. D., Masliah, E., & Hansen, L. A. (1999). The neuropathology of Alzheimer's disease and the structural basis of its cognitive alterations. In R. D. Terry, R. Katzman, K. L. Bick, & S. S. Sisodia (Eds.), *Alzheimer's disease* (2nd ed, pp. 187–206). New York: Lippincott, Williams & Wilkins.
- Terry, R. D., Masliah, E., Salmon, D. P., Butters, N., DeTeresa, R., Hill, R., et al. (1991). Physical basis of cognitive alterations in Alzheimer's disease: Synapse loss is the major correlate of cognitive impairment. *Annals of Neurology*, *30*, 572–580.
- Terry, R. D., Peck, A., DeTeresa, R., Schechter, R., & Horoupian, D. S. (1981). Some morphometric aspects of the brain in senile dementia of the Alzheimer type. *Annals of Neurology*, *10*, 184–192.
- Tzelgov, J., Henik, A., & Berger, J. (1992). Controlling Stroop effects by manipulating expectations for color words. *Memory & Cognition*, *20*, 727–735.

Received July 5, 2001

Revision received December 10, 2001

Accepted January 22, 2002 ■