International Journal of Gerontology 18 (2024) 80–84
https://doi.org/10.6890/IJGE.202404_18(2).0004

International Journal of Gerontology

journal homepage: http://www.sgecm.org.tw/ijge/

Original Article

Stroop Color-Word Test Performance of Chinese-Speaking Persons with Alzheimer's Dementia

Yu-Te Lin [a,b], Yi-Hsiu Lai [c *]

[a] *Center for Geriatrics and Gerontology, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan,* [b] *School of Medicine, College of Medicine, National Sun Yat-sen University, Kaohsiung, Taiwan,* [c] *National University of Kaohsiung, Kaohsiung, Taiwan*

ARTICLE INFO

Accepted 29 November 2023

Keywords:
Alzheimer disease,
neuropsychological tests,
cognition,
stroop test,
SI scores

SUMMARY

Background: Stroop Color Word Test (SCWT) is often adopted to evaluate processing of cognitive function or interference. This study addressed whether different indices of Stroop interference (SI) significantly distinguished Chinese-speaking Alzheimer's disease (AD) persons from control counterparts, and whether SCWT interference errors changed in AD persons of different severity.
Methods: Forty native speakers of Mandarin Chinese, including 20 AD persons and 20 controls, participated in this study. All of the participants took SCWT, and their performance was computed for eight SI indices. Different conditions, response types, and AD errors were also examined.
Results: Our findings revealed significant group differences in terms of raw scores, response types, and most SI scores. We additionally contributed to SCWT performance by AD persons of different severity. SCWT performance of very mild AD persons was slightly impaired in the incongruous condition, while the mildly-demented persons were more severely impaired in all of the conditions.
Conclusion: SI indices in the current study are proven to be helpful in distinguishing AD persons from controls. Moreover, dementia exerts a certain impact on AD SCWT performance. It is thus suggested that AD persons be given help with appropriate application of SCWT, and that SCWT interference errors be adopted as a diagnostic tool in future treatment.

Copyright © 2024, Taiwan Society of Geriatric Emergency & Critical Care Medicine.

1. Introduction

In the field of brain cognition, Stroop Color and Word Test (SCWT) has been widely used as an experimental and clinical neurolinguistic assessment. It was originally proposed[1] to examine participants' performance of naming colors in a limited period of time. There are three different conditions, two of which are congruent conditions (e.g. the word "red" is printed in red ink), and one is an incongruent condition (where color words are printed in different colors, e.g., the word "red" is printed in green ink). In the incongruent condition, participants have to name the color of the printed word rather than the word itself. The interference in the naming process caused by language outputs of colors or words is referred to as Stroop Effect,[1] in which the congruency effect is especially evident in less accurate responses in the incongruent than in the congruent conditions.

Later, an increasing number of researchers attempted to examine the Stroop Effect in the fields of attention, brain processing speed, cognitive flexibility, or working memory, and to compare performance by different populations, for example, senior citizens and those with such brain diseases as Alzheimer's disease (AD).[2–9] In doing SCWT, AD persons are frequently observed to suffer from cognitive impairments and to experience a certain degree of Stroop Effect. Greater Stroop Effect was often identified in AD persons than in controls.[4] AD persons were faced with more difficulty in doing SCWT, especially when processing color words printed in an inconsistent colored ink.

Compared with the western reports, relatively fewer SCWT studies have been conducted in the field of Chinese language and Chinese-speaking populations.[7,10] In order to collect naming information and specific situations in the operation of brain regions while doing SCWT, Event-Related Potentials (ERPs) were additionally used to investigate color-word naming in AD persons.[7] There are differences between Chinese and English words. Chinese language is made up of ideographic scripts, and recognition of Chinese characters involves a visual process from form (ideographic scripts) to meaning (semantic activation).[7,11] Due to different linguistic systems, naming response time also differs. Naming of color words in AD persons in both Chinese and English was studied. The results showed some electrophysiological evidence related to the Chinese-word Stroop Effect, with the impact of cross-linguistically lexical differences.[7] Recently, the Chinese SCWT (C-SCWT) and the Chinese Trail Making Test (C-TMT) were adopted to investigate language processing speed and brain flexibility in Chinese-speaking older adults.[10] Participants were forty healthy older adults, divided into two subgroups. It was observed that both C-TMT and C-SCWT demonstrated significant test-retest reliability in older adults. The findings highlighted that C-TMT and C-SCWT could be utilized to assess executive functions in Chinese-speaking older adults.

In sum, SCWT is commonly applied in western countries to evaluate the processing of cognitive functions in AD persons, re-

* Corresponding author. 700, Kaohsiung University Rd., Nanzih District, Kaohsiung 811, Taiwan.
E-mail address: yhlai@nuk.edu.tw (Y.-H. Lai)

vealing their significant differences and Stroop Effect from healthy controls.[3,8,12] So far, this issue has been comparatively less addressed in Chinese-speaking AD population. Limitations, however, remain in the preceding literature. Firstly, different indices of Stroop interference were adopted in different research reports,[9,10,12] making it difficult to have a broad overview of SCWT performance for a specific population. Secondly, it remains unclear whether SCWT performance is sensitive to dementia severity in AD persons.[4] No universal agreement has been reached concerning whether SCWT performance changes significantly in relation to the severity of dementia.[3] Also, few studies have recorded and analyzed errors made by AD persons to illustrate the process or difficulty they might encounter while doing SCWT. Hence, in the present study, we aim to examine whether different indices of Stroop interference (SI) significantly distinguish Chinese-speaking AD persons from control counterparts and to investigate whether SCWT performance changes differently in these AD persons of different severity via error analysis.

2. Method

Participants in this study were forty native speakers of Mandarin Chinese in Taiwan, including 20 AD persons (14 females and 6 males) and 20 healthy controls (12 females and 8 males). Their ages ranged from 61 to 90 years old. The AD persons were clinically-diagnosed as having AD, not other types of dementia, according to the NINCDS-ADRDA criteria.[13] Their clinical dementia rating (CDR)[14] scores were 0.5 (N = 3) and 1 (N = 17) respectively. A CDR of 0.5 is indicative of very mild dementia, while that of 1 indicates mild dementia. The controls were made up of healthy senior citizens, who scored less than 1 on the AD-8 test.[15] The Ethics Committee, the Institutional Review Board (IRB) at Kaohsiung Veterans General Hospital (KVGH), gave ethical approval for the study.

Participants were first screened by a test of color blindness.[16] Those who passed the test of color blindness proceeded to the second test, also SCWT.[10] Following the widely-adopted framework,[1,10,12] we examined three conditions, namely (a) color-words printed in black ink (W), (b) different color patches (C), and (c) the color-word (CW) condition: color-words printed in an inconsistent colored ink. To minimize the confusion of individual perception or understanding, we adopted common basic colors, for example, red, yellow, green, blue, and black. Clear direction and two examples of naming colors were given before the formal testing. The test was conducted individually and in a quiet room with sufficient lighting.

After data collection, the recorded performance was analyzed in three different conditions, as introduced earlier: W, C, and CW. The first two are considered a congruous condition, while the third one an incongruous condition. Based on previously-proposed formula,[9,12] eight types of SI scores were calculated and compared. They are listed below:

SI-1:[9] SI-1 is the raw score itself of the CW page of the SCWT (i.e., the CW score).
SI-2:[9] SI-2 is the difference between the C and CW scores (C – CW).
SI-3:[9] SI-3 is the ratio of (C – CW)/C.
SI-4:[9] SI-4 is the difference between the predicted CW score of (W C)/(W + C) and the CW score ((W C)/(W + C) – CW).
SI-5:[9] SI-5 is computed by the equation CW – (W + C)/2.
SI-6:[9] SI-6 is calculated using the equation (W + CW) – C.
SI-7:[12] **Pcw** = (W × C)/(W + C) [Pcw: apredicted CW score].
SI-8:[12] **IG** = CW – Pcw [IG: an interference score].

Additionally, three response types, namely missing answers (MA), wrong answers (WA), and correct answers (CA), were evaluated and discussed. Examples of missing answers (MA) were no response or such replies as “No idea” and “I don’t know.” To further investigate the errors AD persons made in SCWT, their wrong answers were further examined in an error analysis of three different conditions, that is, C, W, and CW, respectively.

3. Results

This section presents statistical results of SCWT by Chinese-speaking AD persons and their healthy counterparts. Their SCWT performance was analyzed in terms of raw scores, response types, SI scores, and AD errors.

Table 1 reports the two groups’ raw scores on SCWT; there are a number of findings. Firstly, as shown by the mean scores, the control group excelled in all three conditions, including W, C, and CW. Secondly, significant differences were observed in two conditions, W and CW. In the W condition, the control group (Mean = 8.80) scored significantly higher than the AD group (Mean = 7.45). This pattern was identified in the CW condition, in which the control group (Mean = 15.45) got significantly higher scores than the AD group (Mean = 10.45). In brief, the control group performed better than the AD group on SCWT, specifically in the W and CW conditions.

Table 2 shows response types made by these two groups in SCWT. To start with, significant group differences were identified in all of these response types. The AD group (Mean = 2.85) produced significantly more missing answers (MA) than did the controls (Mean = 0), who had no missing answers. Also, the AD group (Mean = 9.00) made a significantly larger number of wrong answers (WA) than the controls (Mean = 1.10). By contrast, the AD group (Mean = 28.15) uttered significantly fewer correct answers (CA) than did their control counterparts (Mean = 38.90). To sum up, it can be argued that AD participants performed significantly worse than their healthy counterparts in these three response types.

Table 3 summarizes eight types of SI scores obtained by these two groups in SCWT. Significant differences were observed in seven types of SI scores (i.e., SI-1, SI-2, SI-4, SI-5, SI-6, SI-7, SI-8) except for SI-3. Moreover, the AD group performed significantly worse in these types of SI scores. To illustrate, in naming color words printed in an inconsistent color ink (CW), also shown in SI-1 scores, the control group (Mean = 15.45) scored approximately 15 times higher than the

Table 1
Raw scores of SCWT by AD and control groups.

Condition type	Group	*N*	Mean	*SD*	*t*
W	AD	20	7.45	2.32	2.53*
	Control	20	8.80	0.52	
C	AD	20	13.35	2.96	1.87
	Control	20	14.65	0.93	
CW	AD	20	10.45	5.27	4.14***
	Control	20	15.45	1.14	

Note: * $p < .05$; *** $p < .001$.

Table 2
Response types in SCWT by the AD and control groups.

Response type	Group	*N*	Mean	*SD*	*t*
MA	AD	20	2.85	5.91	-2.16*
	Control	20	0	0	
WA	AD	20	9.00	8.22	-4.20***
	Control	20	1.10	1.77	
CA	AD	20	28.15	12.77	3.73**
	Control	20	38.90	1.77	

Note: * $p < .05$; ** $p < .01$; *** $p < .001$.

AD group (Mean = 1.15). As for the SI-2 index, the lower the SI-2 score is, the less interference from incongruent words in the CW condition.[9] The control group (Mean = -0.80) obtained significantly lower scores than the AD group (Mean = 1.54), demonstrating that the former group experienced significantly less Stroop interference than the latter group. These results clearly indicate that the control group performed better than the AD group in the incongruous condition. Additionally, as revealed in the SI-8 (IG) scores, the control group (Mean = 9.96) significantly outperformed the AD group (Mean = 0.99). Looking into the individual reports of the SI-8 (IG) scores, we found three cases of negative IG values, all of which were made by mild AD participants (CDR 1). These negative IG scores were respectively -1.44, -2.53, and -3.2. In contrast, all participants in the control group got positive IG scores. It can be argued that this finding truly reflected one notable group distinction from the perspective of SCWT SI scores.

AD participants made 120 errors while doing SCWT; the results are summarized in relation to the three conditions (C, W, and CW) in Table 4. Firstly, there were 13 errors in the W condition, and all of them were made by mild AD persons (CDR 1). Instead of offering the correct answer "black," 10 errors were attributed to Chinese characters; for example, they produced *hóng* "紅" ("red") when given the character "紅" printed in black ink. Secondly, there were 16 errors in the C condition, all of which were made by mild AD persons (CDR 1). Instead of making character-related errors, these AD participants made some other responses not related to the target characters. For example, the response *huáng* "黃" ("yellow") was made when presented with the character "紅" printed in red ink. The results might be attributed to their occasional color confusion. Finally, approximately three-fourths of AD errors were found in the CW condition (Frequency = 91). In Error Type 1, significantly more errors were observed in mild AD persons (CDR 1) (Frequency = 69) than very mild ones (CDR 0.5) (Frequency = 5). In a similar vein, mild AD persons (CDR 1) made significantly more errors (Frequency = 15) than very mild ones (CDR 0.5) (Frequency = 2) in Error Type 2. To sum up, these findings clearly manifested that the severity of Alzheimer's dementia significantly influenced AD errors in doing SCWT, and these error patterns truly reflected the inhibition process for AD persons.

Table 3
SI scores in SCWT by AD and control groups.

SI-scores	Group	*N*	Mean	*SD*	*t*
SI-1	AD	20	1.15	5.28	4.14***
	Control	20	15.45	10.45	
SI-2	AD	20	1.54	5.37	-2.96**
	Control	20	-0.80	2.90	
SI-3	AD	20	0.12	0.81	-0.91
	Control	20	-0.06	0.11	
SI-4	AD	20	1.11	5.07	-4.29***
	Control	20	-14.79	-9.82	
SI-5	AD	20	1.07	4.89	3.28**
	Control	20	3.73	0.05	
SI-6	AD	20	1.96	6.05	3.55**
	Control	20	9.60	4.55	
SI-7 (Pcw)	AD	20	0.25	1.33	2.56*
	Control	20	5.49	4.71	
SI-8 (IG)	AD	20	0.99	4.82	3.84***
	Control	20	9.96	5.74	

Note: * $p < .05$; ** $p < .01$; *** $p < .001$.

4. Discussion

Chinese-speaking AD persons' and healthy controls' performance in SCWT is discussed in this section. Results outlined previously are discussed with those in the relevant literature.

First of all, our findings reveal significant group differences in terms of raw scores and response types. Overall, the AD group performed significantly worse than the control group in doing SCWT, specifically in the W and CW conditions. From the perspective of response types, AD participants gave significantly more missing answers, more wrong answers, but fewer correct answers. The result that AD persons performed significantly worse than their control counterparts in doing SCWT is consistent with previous studies.[5,8]

Based on different types of SI scores, we offer supporting evidence to elaborate greater Stroop interference on Chinese-speaking AD persons than on healthy controls. Significant group distinction was identified in seven out of eight SI scores, namely SI-1, SI-2, SI-4, SI-5, SI-6, SI-7, and SI-8. No significant group difference was observed in the index of SI-3, which refers to the ratio of C-CW score difference. Probably, the numerical difference was so small that it led to the insignificant result. Our major findings regarding significant group-distinction in most of the SI scores also helped account for previously-made assumptions.[9,12] A lower SI-2 score, for example, represents less interference from incongruent words in color naming on the CW condition.[9] Our controls scored significantly lower than AD persons, manifesting that they experienced significantly less Stroop interference. As for the SI-8 (IG) score, a negative IG value indicates a pathological ability to inhibit interference, and the lower score a person obtains, the greater difficulty they have inhibiting interference.[12] In the current investigation, we found three individual cases of negative SI-8 (IG) value, all of which were for AD persons with CDR 1. These three cases of negative SI-8 (IG) are highly

Table 4
AD errors in three conditions of SCWT.

	CDR 0.5	CDR 1	*F*	*P*	*t*
W condition					
Type 1: Character-related responses	0	10	10	77%	1.16
Type 2: Responses not related to characters	0	3	3	23%	
Subtotal	0	13	13	100%	
C condition					
Type 1: Character-related responses	0	0	0	0%	-4**
Type 2: Responses not related to characters	0	16	16	100%	
Subtotal	0	16	16	100%	
CW condition					
Type 1: Character-related responses	5	69	74	81%	5.93***
Type 2: Responses not related to characters	2	15	17	19%	
Subtotal	7	84	91	100%	
Total		120			

Note: F, frequency; P, percentage. ** $p < .01$; *** $p < .001$.

associated with dementia severity, as shown in the larger CDR scores. It clearly demonstrated the pathological interference in processing information and different degrees of Stroop Effect. The more severely a person suffers from Alzheimer's dementia, the greater Stroop interference they might experience.

According to the error analysis, this report is additionally contributive to SCWT performance by Chinese-speaking AD persons of different severity. Making analysis of AD errors helps unveil how they process the given information and what their inhibitory deficits are attributed to. Judging from error frequency, most of AD errors were made by demented persons with greater severity (CDR 1). Among 120 AD errors, only seven were produced by very mild AD persons (CDR 0.5). Besides, examining the nature of the errors, we observed an overwhelming influence of Chinese characters on AD processing in two conditions: W and CW. In the W condition, all color-words were printed in black ink. Instead of offering the correct answer "black," mild AD participants (CDR 1) were erroneous in making more character-related responses, demonstrating the influence of Chinese characters on their SCWT performance. In the CW condition, AD persons made significantly more character-related errors, and significantly more errors were observed in mildly-demented persons (CDR 1) than in very mild ones (CDR 0.5). To sum up, the findings of error analysis manifested that AD participants tended to be visually influenced by Chinese characters, and that severity of the disease played an essential role in their inhibition process while doing SCWT.

Moreover, significantly more Chinese character-related responses in AD errors add further support to previously-claimed theories, for example, relative speed-of-processing[17] and automaticity theory.[18,19] Relative speed-of-processing[17] argues that people practice naming words more often than naming colors, thus being more familiar with naming words. Also, the speed of processing colors, which might stimulate multiple responses, is relatively slower than that of words. According to automaticity theory,[18,19] the encoding of words processes automatically in SCWT. It is generally agreed that making responses to words is faster than to colors, and that processing of words and of colors are respectively undertaken via different routes.[20] In other words, processing of words or naming characters is indeed a response of automation. These claims help account for the deficient response inhibition for Chinese-speaking AD persons who tend to make Chinese character-related errors. Additionally, these errors might occur when AD persons got confused in the test condition. Future study is thus called for regarding to what extent they fully understand what they are supposed to name.

Taking dementia severity into consideration, we make further contributions regarding notable different SCWT performance between very mild AD persons (CDR 0.5) and mildly-demented persons (CDR 1) in relation to symbolic compatibility.[21] Referring to symbolic compatibility, high compatibility occurs when stimulus and response require identical encoding, as demonstrated in the congruous condition (also C and W). Low compatibility, by contrast, is evident in naming color of words in the incongruous condition (also CW). Processing tasks of low compatibility poses relatively greater challenges than processing those of high compatibility. In the current investigation, the SCWT performance of very mild AD persons (CDR 0.5) is slightly impaired in the incongruous condition (CW), as is the task of low compatibility. By contrast, those mildly-demented persons (CDR 1) are more severely impaired in all of the conditions, including tasks of high compatibility (also W and C) and tasks of low compatibility (also CW). Those mildly-demented persons are characterized by more deficient response inhibition, as manifested by a larger number of errors. In sum, the severity of the disease influenced AD performance in SCWT, and greater challenges were observed in the incongruent color-word trials. It can be argued that Stroop interference errors, particularly errors on the incongruous trials, are sensitive to discriminate persons in the early stage of AD.

5. Conclusion

This study addressed the application of SCWT in a Chinese-speaking AD population, as compared to their healthy counterparts. Their test scores were calculated and compared in a number of indexes. Significant group differences were evident in most of SI scores. A notable Stroop interference was observed in AD persons with cognitive impairments. Errors made by AD persons were further analyzed. Compared with very mild AD persons, those mildly-demented persons were more seriously deteriorated with a larger number of Chinese character-related errors. It can be argued that dementia has exerted a certain impact on AD performance of SCWT. These in turn imply that AD persons should be given assistance with appropriate use of SCWT, and that errors in SCWT performance can also be adopted as a diagnostic instrument in future treatment. Finally, due to the limited number of participants, it is suggested that more participants be recruited in further research to generalize the pattern identified in the present study.

Acknowledgments

This study was supported by a national grant from the National Science and Technology Council (NSTC) (110-2410-H-390-006), the Republic of China (ROC) Taiwan. The researchers thanked the participants in the study.

References

1. Stroop JR. Studies of interference in serial verbal reactions. *J Exp Psychol*. 1935;18(6):643–662. doi:10.1037/h0054651
2. Amieva H, Lafont S, Rouch-Leroyer I, et al. Evidencing inhibitory deficits in Alzheimer's disease through interference effects and shifting disabilities in the Stroop test. *Arch Clin Neuropsychol*. 2004;19(6):791–803. doi:10.1016/j.acn.2003.09.006
3. Ben-David BM, Tewari A, Shakuf V, Van Lieshout PH. Stroop effects in Alzheimer's disease: selective attention speed of processing, or color-naming? A meta-analysis. *J Alzheimers Dis*. 2014;38(4):923–938. doi:10.3233/JAD-131244
4. Bondi MW, Serody AB, Chan AS, et al. Cognitive and neuropathologic correlates of Stroop Color-Word Test performance in Alzheimer's disease. *Neuropsychology*. 2002;16(3):335–343. doi:10.1037//0894-4105.16.3.335
5. Fisher LM, Freed DM, Corkin S. Stroop Color-Word Test performance in patients with Alzheimer's disease. *J Clin Exp Neuropsychol*. 1990;12(5):745–758. doi:10.1080/01688639008401016
6. Hutchison KA, Balota DA, Duchek JM. The utility of Stroop task switching as a marker for early-stage Alzheimer's disease [published correction appears in Psychol Aging. 2010 Dec;25(4):778. Ducheck, Janet M [corrected to Duchek, Janet M]]. *Psychol Aging*. 2010;25(3):545–559. doi:10.1037/a0018498
7. Qiu J, Luo Y, Wang Q, Zhang F, Zhang Q. Brain mechanism of Stroop interference effect in Chinese characters. *Brain Res*. 2006;1072(1):186–193. doi:10.1016/j.brainres.2005.12.029
8. Spieler DH, Balota DA, Faust ME. Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *J Exp Psychol Hum Percept Perform*. 1996;22(2):461–479. doi:10.1037//0096-1523.22.2.461
9. Yun JY, Lee DY, Seo EH, et al. Neural correlates of Stroop performance in Alzheimer's disease: a FDG-PET study. *Dement Geriatr Cogn Dis Extra*. 2011;1(1):190–201.doi:10.1159/000329517
10. Wang RY, Zhou JH, Huang YC, Yang YR. Reliability of the Chinese version of the trail making test and Stroop color and word test among older adults. *Int J Gerontol*. 2018;12(4):336–339. doi:10.1016/j.ijge.2018.06.003

11. Zhou X, Marslen-Wilson W. Direct visual access is the only way to access the Chinese mental lexicon. In: Cottrell GW, ed. *Proceedings of the 18th annual conference of the Cognitive Science Society*. 1st ed. Routledge; 1996:714–719
12. Scarpina F, Tagini S. The Stroop color and word test. *Front Psychol*. 2017; 8:557. doi:10.3389/fpsyg.2017.00557
13. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the NINCDSADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology*. 1984;34(7):939–944. doi:10.1212/wnl.34.7.939
14. Morris JC. The clinical dementia rating (CDR): current version and scoring rules. *Neurology*. 1993;43(11):2412–2414. doi:10.1212/wnl.43.11.2412-a
15. Yang YH, Lee MB, Liu JK. Screening of vert early Alzheimer's dementia. Article in Chinese. *Taiwan Med J*. 2009;52(9):8–10. doi:10.30044/TMJ.200909.0002
16. Ishihara S. *Tests for Color-blindness*. Handaya Hongo Harukich; 1917.
17. Cattell JM. The time it takes to see and name objects. *Mind*. 1886;11(41): 63–65.
18. Posner MI, Snyder CRR. Attention and cognitive control. In Solso RL, ed. *Information processing and cognition: The Loyola symposium*. Lawrence Erlbaum; 1975:55–85.
19. Logan GD. Attention and automaticity in Stroop and priming tasks: theory and data. *Cogn Psychol*. 1980;12(4):523–553. doi:10.1016/0010-0285(80)90019-5
20. Chen MS, Chen JY. Scheduling of mental processes in the Stroop task with critical path method. *Chin J Psychology*. 2003;45(4):379–400. doi:10.6129/CJP.2003.4504.06
21. Simon JR, Sudalaimuthu P. Effects of S-R mapping and response modality on performance in a Stroop task. *J Exp Psychol Hum Percept Perform*. 1979;5(1):176–187. doi:10.1037//0096-1523.5.1.176