

# The Arabic version of Stroop test and its equivalency to the English version

Ali M Al-Ghatani<sup>1</sup>, Marc C Obonsawin<sup>2</sup>, Khalaf R Al Moutaery<sup>3</sup>

## Abstract

**Objectives:** The Stroop test is one of the most popular tests frequently used to assess the function of the frontal lobe in neurological and psychiatric patient populations. Performance on the Stroop test is very sensitive to lesions of the frontal lobes and is commonly used in clinical settings. In 1999, we decided to find out if the Stroop test will be as reliable after translation to the Arabic language as the original English version. We completed the work in Riyadh Military Hospital in 2000.

**Methods:** A sample of 10 Saudi adult healthy individuals participated in this study. Their mean age was 31.9 and their mean years of education were 17.3. All subjects performed the Arabic and the English versions of Stroop test.

**Results:** Performance showed no differences between the English and the Arabic versions of the Stroop test.

**Conclusion:** Equivalency data between the two versions of Stroop test are provided. The availability of this data will help us to provide a normative data for the Saudi Committee of Health Education, and in carrying out research on frontal lobe function. (p112-115)

## Introduction

The Stroop test is one of the most popular tests frequently used to assess the function of the frontal lobe in neurological and psychiatric patient populations. During the Stroop test the participant is presented with three cards. Two cards have columns of words (W) and each word is the name of a colour (C). On one card the words are printed in black ink and on the other card the words are printed in coloured ink, but never in the same colour as represented by the word. A third card has columns of XXXXs, printed in

coloured ink. The main point of the Stroop test is that it takes longer to say the colour of the ink that a word is printed in than it takes to read the word. This task is more difficult when the word is a colour that is different from the colour of the ink in which the word is printed.

In the earliest version of the Stroop test, there were 5 words printed on a sheet consisting of 10 columns and 10 rows and these words were red, green, brown, blue and purple on the first page. Each one of these words was printed in a different colour, for example the word blue was printed in green ink but never printed in blue ink. The task was to say the colour of the ink aloud, and not read the word. The second page contained rectangles printed in different colour ink, and the task was to say the colour of the rectangle. In the final page all the words blue, green, red, brown and purple were printed in black ink only, and the task was to read the words alone.

Stroop did the original version of the Stroop test.<sup>6</sup> Stroop used 5 colours in the test. Later investigators have used 3, 4 and 5 colours.<sup>2,7</sup>

Later, Golden redesigned the test.<sup>3</sup> He compared the results of the subject on the three versions of the Stroop test employing 3, 4 and 5 colours. He found no difference between the three versions of the test in producing the 'Interference Effect', so he standardized a test with 3 colours only. A variety of stimuli (e.g. solid rectangles,

<sup>1</sup>Consultant & Head of Neuropsychology Division  
Department of Clinical Neurosciences  
Riyadh Military Hospital  
Riyadh  
Saudi Arabia.

<sup>2</sup>Senior Lecturer, Department of Psychology  
University of Strathclyde  
Glasgow  
United Kingdom

<sup>3</sup>Professor Neurosurgeon  
Riyadh Military Hospital

## Correspondence:

Dr. Ali Al-Ghatani  
Head of Neuropsychology Division  
Department of Clinical Neurosciences  
Riyadh Military Hospital  
PO Box 45700  
Riyadh 11522  
Saudi Arabia  
Email: am1221@hotmail.com

swastikas, circles and nonsense syllables) were initially used to present the colour patches on the colour page only, and Golden decided to use the four XXXX's to make sure that the stimulus has no semantic value at all.<sup>3</sup> Four X's were used because the average number of letters in the words was four (red, green, and blue).

Golden's standardized version of the Stroop test consists of three white pages of the same size (8 1/5 × 11 inches).<sup>3</sup> Each page has 100 items presented in 5 columns and each column contains 20 items. On the first page, the words red, green and blue are printed in black ink and are randomly arranged. On the second page 5 columns of four XXXX's are printed in green, red, or blue ink. On the third page, the words from page one are printed with the same colours and sequence as the colours (XXXX) of page 2. All participants receive the same instructions to complete this test. Participants have 45 seconds to read the words on the first page as fast as they can, and 45 seconds to say the colour of each item on pages 2 and 3.

To administer this test, all three pages are given to the participant. These pages must be organized so that page one is presented first, the second page follows, and the third page is presented last. All three pages must be placed in front of the participant. The variable that is recorded is the number of completed items on all three pages. The purpose of the C and W page scores is to derive a pure interference score from the colour word page, which is not dependent on the participant word reading or colour naming speed. To determine the pure interference scores, the predicted (CW) scores must be subtracted from the raw (CW) scores, the result of page three. The formula calculates the predicted CW scores. CW means the predicted scores.

$$[(C \times W) \div (C + W) = CW']$$

The interference score is calculated from this formula:

$$CW - CW' = \text{interference score}$$

The mean interference mean score is 0.0. A score above zero indicates a high resistance to interference.

In reviewing most of the research done on the Stroop test, I could not find any modified version of the test to the Arabic population. Toufik Bahri and Ahmad Bendania1 in the United Arab Emirates University carried out one study on the Stroop test.<sup>1</sup> They did not modify the test for the Arabic speakers, but designed a new computerised English version. They used the same colours. Participants were instructed to press the letter key that corresponded to the colour of displayed stimulus.

Due to the difficulty in using the English version of the

Stroop test in the Arab population, and because of the sensitivity of this test to measure the frontal lobe function and improve patient care, we carried out this pilot study after translating the original version to ascertain if we could modify the original version of the Golden's Stroop test to the Arab population.<sup>3</sup>

The development of such a test for use in the Arabic language, and the availability of norms based on the performance of healthy people, would be useful to neuropsychologists working with individuals whose first language is Arabic.

## Methods

### Participants

There were 10 right-handed Saudi Arabian participants, 5 males and 5 females recruited from our institution. Their mean age was 31.9 and their mean of years of education was 17.3. Table 1 shows the demographic characteristics of the sample. All participants were almost as fluent in the English language as in the Arabic language. All participants reported no history of psychiatric illness, alcohol abuse, neurological disorder, or serious head injury. None of the participants were taking psychotropic medication at the time of testing.

**Table 1** - Participants demographic characteristics.

Subject No.	Age	Sex	Years of education
1	26	Male	17
2	28	Male	17
3	30	Male	17
4	34	Male	16
5	44	Male	20
6	24	Female	17
7	27	Female	16
8	31	Female	17
9	35	Female	16
10	40	Female	20

### Procedure

All the words on the first page (W) of Golden's version of the Stroop test were translated to the Arabic and the words written in black ink on a white sheet size 8 1/5 × 11 inches. Nothing was changed for the second page (C) because there is nothing to translate and it contains just the four XXXXs, and so was not modified. The third page (CW) was translated and fixed in the same colour order as in the Golden version. Each one of the three pages is containing 100 items each, presented in five columns and each column contained 20 items.

All participants received the English version of the Stroop

first and the Arabic version second. To administer the Stroop test, all three pages are given to the participant. These pages must be organized as page one (W) on the top, the second page (C) will follow, and third page (CW) will be the last. All three pages were placed in front of participants on a table or on a flat surface. The participant could rotate these pages up to 45 degrees only, and the participant was not allowed to lift any one of these pages from the flat surface. In addition, all three pages must be completed in the same way, for example, if the client rotated the second page, then the third page must be rotated in the same way. Every participant had 45 seconds to read as fast as he/she could the words on the first page, and the same time was allowed for pages 2 and 3. When the participant received the three pages, the instruction as in the original of Stroop test manual was translated to the Arabic and was given verbally to the participant (see Appendix C for the Arabic translation of the instructions).

To determine the pure interference scores, the predicted CW scores were subtracted from the raw CW scores (the result of page three). The formula  $([C \times W] \div [C + W] = CW')$  calculates the predicted CW scores, and  $CW'$  means the predicted scores. The interference is calculated from the formula:  $(CW - CW' = \text{interference score})$ . The interference mean value score is (0.0). A score above zero indicates a high resistance to interference.

**Results**

Table 2 shows the number of correct items named on the English and Arabic word naming page, the English and Arabic colour naming page, the English and Arabic colour-word naming page, and the English and Arabic interference scores. Paired sample t-tests (Table 3) shows no differences between the English and the Arabic versions of the Stroop

**Table 2** - The scores of participants on the English and Arabic Stroop versions.

Subj. No.	Eng (W)	Arb (W)	Eng (C)	Arb (C)	Eng (CW)	Arb (CW)	Eng (Int)	Arb (Int)
1	127	123	99	111	62	58	6.35	-.35
2	131	127	101	109	69	69	11.97	10.34
3	120	118	93	91	60	55	7.61	3.62
4	133	131	102	102	70	71	12.27	13.65
5	147	141	105	94	68	78	6.75	21.60
6	133	139	106	107	67	68	8.01	7.54
7	117	119	97	90	61	61	7.97	9.76
8	124	130	98	99	64	63	9.26	6.80
9	131	138	101	106	68	62	10.97	2.05
10	144	142	109	101	69	73	6.96	13.98
Mean ± SD	130.70 ± 9.51	130.80 ± 8.99	101.10 ± 4.70	101.00 ± 7.45	65.80 ± 3.71	65.80 ± 7.19	8.81 ± 2.19	8.90 ± 6.49

test on all the three pages and on the interference scores. The results with the Arabic Stroop can also be compared to results obtained by other investigators using the English version of the Stroop with people whose first language is English. In Table 4 the results from the Arabic Stroop test is displayed along with results obtained by previous investigators.<sup>4,5,6</sup> Their results have been taken from the Stroop test Manual.<sup>3</sup> We can see from Table 4 that there are only very small differences between the mean scores of the three pages in three previous studies with the mean of the Arabic study.<sup>4,5,6</sup> The values for all three variables for the Arabic version of the Stroop test are slightly higher than the values for the English version of the test. It is more important to compare the relationship between each variable. Table 5 shows one way of comparing this relationship between each study. By expressing the colour score and the colour-word score as a proportion of the word score we can obtain an index of the relationship between the variable. Table 5 shows that the difference between the colour scores divided by the word scores and the colour-word scores divided by the word scores in all the four studies are similar.

**Table 3** - Paired sample t-test between the English and the Arabic versions.

Paired samples	Df	T	Sig.
English word (vs) Arabic word	9	.07	.949
English colour (vs) Arabic colour	9	-.04	.966
English colour-word (vs) Arabic colour-word	9	.000	1.000
English interference (vs) Arabic interference	9	.04	.969

**Table 4** - Mean scores for three previous studies and the Arabic mean scores.

Sample	Word raw scores	Colour raw scores	Colour-word raw scores	Interference scores
Stroop (1935) (N = 100)	110	71	41	- 2.15
Jensen (1965) (N = 436)	118	78	45	- 1.96
Jensen (1966) (N = 436)	121	81	51	2.48
Arabic (N = 10)	130	101	65	4.54

**Table 5** - The difference between word, colour, and colour-word.

Sample	Colour (divided by) word	Colour-word (divided by) word
Stroop (N = 100)	0.64	0.37
Jensen (N = 436)	0.66	0.38
Jensen (N = 436)	0.67	0.42
Arabic (N = 10)	0.78	0.50

## Discussion

The aim of this pilot study was to determine whether the Arabic Stroop test is as useful with Arabic native speakers as the English version is with English speakers. All results showed that the English Stroop version and the Arabic Stroop version are almost the same. There is some indication that the number of items completed in 45 seconds is greater in the Arabic version than in the English version. This result came as a surprise because the Arabic words for “red”, “green” and “blue” consists of two syllables each, and it was anticipated that fewer Arabic items would be completed during the 45 second period. Perhaps more Arabic items were completed because of the phonetic similarities between the Arabic words for these colours.

Of more importance, it appears that the proportion of items completed on the Word, Colour and Colour Word pages is similar for the English and Arabic versions of the Stroop test.

The data presented in this pilot study only allow to be sure that the Arabic translation would be appropriate for the Stroop test.

## References

1. Bahri T, Bendania A: Effect of language dominance on cognitive processes in a Stroop task. *Percept Mot Skills* 1997, 84: 979-988
2. Broverman DM: Cognitive style and intra-individual variations in abilities. *J Pers* 1960, 28: 240-256
3. Golden CJ: A group version of the Stroop colour and word test. *J Pers Assess* 1975, 39: 386-388
4. Jensen AR: Scoring the Stroop test. *Acta Psychol* 1965, 24(5): 398-408
5. Jensen AR, Rohwer WD Jr: The Stroop colour-word test: a review. *Acta Psychol* 1966, 25(1): 36-93
6. Stroop JR: The basis of Ligon's theory. *Am J Psychol* 1935, 47: 499-504
7. Thurstone LL, Mellinger JJ: The Stroop test. *The Psychometric Laboratory, University of North Caroline* 1953, No. 3