

EXAMINATION OF SOME ASPECTS OF THE STROOP COLOR-WORD TEST¹

E. C. DALRYMPLE-ALFORD AND BUSHRA BUDAYR

American University of Beirut, Lebanon

Summary.—In the Stroop Test color naming is found to be slower when the colors are those in which non-corresponding color names are written. A similar result was found when the incongruent word-color combinations were presented individually. When presented in lists, the structure of the lists was found to contribute to impairment of color naming.

The usual procedure in the Stroop Color-Word Test is to present the subject with a list of color names written in non-corresponding colors (e.g., the word 'red' written in blue ink), with instructions to name the colors in which the words are written as quickly as possible while ignoring the words themselves. Performance under these conditions is decidedly inferior to that under control conditions in which colors appear as patches or Xs. While research with this test has been increasing recently (for references see Rand, Wapner, Werner, & McFarland, 1963; Klein, 1964), scarcely any work has been reported on procedural aspects of the task that could throw light on the processes involved. Two such studies are briefly reported here.

EXPERIMENT I

This study had two aims, only one of which is relevant here. This was to examine whether interference with color naming depends on the serial nature of the task. Would one still observe impaired color naming when perceptual confusion between the items is removed by presenting the items individually?

Method

Ten Arabic-English bilingual university students, tested individually, were presented with stimulus cards on each of which was stamped either an English or an Arabic color name or a nonsense squiggle (composed of segments of English and Arabic letters) in red, blue, green, or black ink. Each card was exposed in a tachistoscope until the subject called out the color of the ink in English or Arabic as specified before presentation. The interval between the onset of the display and the onset of the subject's response, picked up by a voice key, was measured to the nearest 2 msec. There were 40 each of the English, Arabic, and Control cards; for half of each the response was in English, the rest in Arabic. The order of the conditions was randomised.

Results

Mean latencies of correct responses (approximately 90% of all the inter-

¹This work was supported in part by a grant from the Arts and Sciences Research Committee of the American University of Beirut.

ference condition responses and 98% of the control), are set out in Table 1 with the outcomes of *t* tests comparing interference-control differences indicated. Implications of the manifest interlingual interference concern the other aim of the study and will be discussed elsewhere. For the present we conclude that impaired color naming observed in the serial-responding version of the Stroop Test cannot be attributed solely to perceptual confusion between the items.

TABLE 1
MEAN LATENCIES (IN MSEC.) OF CORRECT RESPONSES

	Interference in		Control
	Arabic	English	
Arabic response	366‡	358‡	336
English response	418‡	404‡	381

†Interference-control difference significant at 5% level.

‡Interference-control difference significant at 1% level.

EXPERIMENT II

This study was concerned with the effect of list structure on performance on the serial version of the test.

Method

There were 7 conditions: RB—*reading* color names written in black; RC—*reading* color names written in incongruent colors; CC—*naming* the colors of Xs; S—*naming* the colors of incongruent word-color combinations, the correct response never being the same as the preceding or following correct response or word to be suppressed; SX—similar to S except that the correct response at any point in the list was identical to the word interfering with the next correct response; XS—the 'reverse' of SX, the correct response being the same as the interfering word of the preceding correct response; and HS—similar to S except that for half the items, randomly selected, the word matched the color of the ink.

We expected the word reading conditions to be the easiest, with RC being somewhat inferior to RB because of the greater clarity of black print. We expected the control, CC, for the color naming conditions to be next best. Apart from the general prediction that the S, SX, XS, and HS conditions would be inferior to the others, we had no clear prediction of differences within this set of interference conditions. It was, however, considered likely that HS ('half-Stroop') would be more difficult than S since, in cases of congruent word-color combinations of the HS condition, a general set to suppress the word could also lead to temporary suppression of the correct response, this being the same. It was also considered likely that the SX ('say-suppress') and XS ('suppress-say') conditions would be worse than the S condition since they could result in the subject's being confused by the successive emission and suppression (and vice versa) of the same response.

Four sets of 20-item lists were prepared for the 7 conditions. The series of correct responses were identical in each set, with all words and required responses in English. The 28 lists, randomly ordered, were presented one at a time to the subject who was instructed to respond as fast as possible without error. Apart from preliminary explanation of the tasks and an indication before each list whether word reading or color naming was required, the subjects were told nothing of the structure of the lists. Tape-recordings of their performance were examined and timed later. The subjects were 16 university students.

Results and Discussion

Mean reading times and mean number of errors are shown in Table 2. After analysis of variance of reading times which showed differences due to conditions to be significant ($P < 0.001$), the Newman-Keuls procedure (Winer, 1962) was carried out to compare the means. Broken lines in the table partition the conditions into 4 sets so that those within the same set had means that did not differ significantly ($P > 0.05$), while means in different sets differed beyond the 0.01 level.

TABLE 2
MEAN READING TIMES AND MEAN ERRORS (CORRECTED AND UNCORRECTED)

Condition	Time (sec.)	Errors
RB	7.48	0.032
RC	8.40	0.047
CC	12.38	0.172
S	16.52	0.500
SX	16.66	0.246
HS	16.92	0.313
XS	19.72	0.860

The clear difference between XS ('suppress-say') and the other interference conditions (S, SX, and HS) was somewhat unexpected. It seems that an explanation in terms only of confusion arising from successive suppression and emission of the same response is inadequate since SX, which requires successive emission and suppression of the same response and therefore should also give rise to confusion, is not significantly different from S or HS. We conjecture that the suppression of a response results in temporary unavailability of that response. This would be an advantage in the S, SX, and HS conditions but not in XS.

We conclude from these experiments that, while performance on the Stroop

Test is not simply a consequence of the serial nature of the task, the structure of the list does have an effect.

REFERENCES

- KLEIN, G. S. Semantic power measured through the interference of words with color-naming. *Amer. J. Psychol.*, 1964, 77, 576-588.
- RAND, G., WAPNER, S., WERNER, H., & MCFARLAND, R. Age differences in performance on the Stroop Color-Word Test. *J. Pers.*, 1963, 31, 534-558.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

Accepted November 10, 1966.