A reversal of the Stroop interference effect, through scanning*

JAMES S. ULEMAN and JEANINE REEVES
Michigan State University, East Lansing, Michigan 48823

Interference on the Stroop test has been explained on the basis of a stronger habit (word reading) interfering with a weaker one (color naming) on an incongruous color-word (CW) card. Since scanning for words is slower (weaker) than scanning for colors, it was predicted that making the Stroop a scanning task would produce more interference with scanning for words rather than colors on the CW card. This prediction was confirmed. In addition, analyses of individual differences offered some support for differential habit-strength theories of Stroop interference. However, differences between the standard and scanning Stroop tests suggested that some kind of information-processing analysis would be more fruitful.

The Stroop color-word test consists of three cards: a color card (C) with rows of color patches, a word card (W) with the names of colors printed in black ink, and an incongruous color-word card (CW) with each color name printed in ink of a conflicting color. That is, RED would be printed in all colors except red. In naming the ink colors on CW, Ss take longer and commit more errors than in either reading CW words or in naming C colors. Since W reading is faster than C color naming, this is usually discussed in terms of the stronger habit (word reading) interfering with the weaker (color naming). The basis for this differential habit strength is unclear. Most investigators favor a differential practice explanation [see Jensen & Rohwer's (1966) comprehensive review of the Stroop literature].

Lund (1927, cited by Jensen & Rohwer, 1966) found that *scanning* times for colors are shorter than for words. Thus, if Ss scan for red on C, they should do this faster than scanning for RED on W. This is a reversal of the usual relation between times on C and W, when Ss name or read each instance. If differential habit strength is the basis for interference on CW, and if the Stroop task is changed to scanning, one would expect a reversal of the CW interference effect. Most interference should occur in scanning for words rather than for ink colors; and there should be a positive correlation between the amount of CW interference and the difference in C and W times (habit strengths).

METHOD

Apparatus

Design of the three Stroop cards was based on recommendations by Jensen and Rohwer (1966). Each was done on white

cardboard, 22 in. high x 24 in. wide, and presented to Ss on an easel at eye level.

The color card (C) had 12 rows of seven rectangular color patches. Colors were chosen randomly from among brown, green, blue, and red, with the restrictions that no adjacent colors be the same and that each appear 21 times. Each rectangle was the same size as the color word in the corresponding position on the word card (W): $\frac{1}{4} \times 3$ in. for BROWN, $\frac{3}{4} \times 2\frac{7}{8}$ in. for GREEN, $\frac{1}{4} \times 2\frac{1}{4}$ in. for BLUE, and $\frac{1}{4} \times 1\frac{1}{4}$ in. for RED. Color names on W were in 12 rows of seven each, printed in black ink and $\frac{1}{4}$ -in. Gothic capital letters. Positions were chosen randomly, with the same restrictions as for C. Color names on CW were also in 12 rows of seven each, but were printed in an incongruous color. Positions were chosen randomly, as with C and W, with the additional incongruity constraint. Each color name and each ink color appeared 21 times.

Procedure

Sixty undergraduates from the introductory psychology subject pool at Michigan State University were randomly assigned to conditions. Ss stood in front of an easel on which cards were presented, one at a time, covered by a sheet of glass. With each card, they scanned first for brown, as a warmup, and then for green and blue. Half (Group 1) scanned CW for ink colors and half (Group 2) for color names. To control for possible practice effects, half of each group scanned cards in the order C, W, CW, and half had the order W, C, CW. Approximately half scanned for green, then blue, while the others scanned for blue, then green on each card. This basic 2 by 2 by 2 factorial design (CW Group by Card Order by Response Order), with repeated measures, was analyzed using unweighted means solutions for unequal cell sizes.

Each of the three trials on each card was

timed with a stopwatch to the nearest second, and interference scores were computed for each S. For Group 1, interference = CW - C (seconds); for Group 2, interference = CW - W. This score was chosen from among many possible interference scores on the basis of analogy with Jensen's (1965) factor analysis of standard Stroop measures.

Ss were instructed to complete each card *correctly*, as quickly as possible, marking *all* appropriate stimuli. Instructions for C were: "Your task will be to scan the card and pick out all of the blocks of one particular color. You must scan each row from left to right, beginning with the first row and moving down one row at a time from the top of the card to the bottom. As you scan the card in this manner, make a checkmark below each rectangle of the specified color, and say the name of the color. Complete the card in this manner as quickly as possible." Ss checked on the glass with a black litho-crayon; the checks were wiped off with a damp cloth after each trial (brown, green, and blue). For W, "color names" was substituted for "blocks (or rectangles) of color." For CW, Group 1 was told to find, mark, and *name* all of the instances of one particular ink color on each trial, disregarding the color names printed on the card. Group 2 was told to find, mark, and *read* all of the names of one particular color on each trial, disregarding ink colors.

RESULTS

Preliminary Analyses

To check for expected C and W scanning-time differences and possible initial differences between CW groups and practice effects, a 2 by 2 by 2 by 4 ANOVA (CW Group by Card Order by Response Order by Trials) was done on scanning times for C-green, C-blue, W-green, and W-blue. There were significant effects within Ss for trials ($F = 779.2, p < .001$),¹ CW Group by Trials ($4.65, p < .01$), Response Order by Trials ($47.9, p < .001$), and a four-way interaction ($5.06, p < .01$). Analysis of simple effects yielded the following conclusions.

First, W scanning times were longer than C times (F s from 79 to 155, $ps < .001$, within each cell), confirming Lund's (1927) finding (see Fig. 1). Second, there were significant practice effects within each card (Response Order by Trials, F s from 11.0 to 17.9, $ps < .001$), shown clearly in Fig. 1, but no differences between green and blue times under similar conditions. Third, there were small practice effects (0.3-0.4 sec) between C and W, and W and C, but only under some conditions (largest $F = 3.28, p < .05$). And finally,

*The authors wish to thank Professors Paul Bakan and Lawrence Messé for their helpful suggestions.

Fig. 1. Mean scanning times on second and third trials for each card: C, W, and CW.

there were a few differences between groups (CW Group by Trials simple interactions) for a few conditions, but none were large or consistent enough across conditions to alter the analyses presented below.

Similar analyses of CW times yielded only the expected overall main effect for the CW group ($F = 213.8$, $p < .001$)² and the practice effects found above within C and W (Response Order by Trials $F = 103.0$, $p < .001$).

Interference on CW

The major hypothesis was that interference in scanning CW would be greater for Group 2 (reading words) than for Group 1 (naming colors). A 2 by 2 by 2 by 2 ANOVA (one repeated measure; interference scores for green, then blue) yielded a significant main effect for the CW group, confirming the major hypothesis ($F = 470.0$, $p < .001$); a CW Group by Card Order interaction ($F = 4.84$, $p < .05$); and a CW Group by Trials interaction within Ss ($F = 5.67$, $p < .05$). Response-order effects were completely cancelled out by the way in which interference scores were computed.

As Fig. 1 suggests, interference differences between groups were huge, in the expected direction. In fact, there was no overlap. All those in Group 1 had zero or negative interference, probably through practice. All those in Group 2 had positive interference.

The CW Group by Card Order interaction reflects practice effects between C and W. For Group 1 mean interference (CW - C) was -1.42 for Card Order C-W, and -1.00 for W-C (simple

main effect; $F = 5.04$, $p < .05$). For Group 2 mean interference (CW - W) was 1.86 for C-W and 1.68 for W-C (n.s.). Differences for both groups are in the direction practice effects would produce, given no differences on CW (see above).

The CW Group by Trials interaction reflects a simple main effect for trials in Group 2 (mean interference = 1.96 for green, 1.58 for blue; $F = 6.86$; $p < .05$). This resulted from a combination of individually nonsignificant differences in trial times on W and CW, and in no way compromises the major results.

Individual Differences

If interference is a positive linear function of the difference in habit strengths, as group results for the standard and scanning Stroop tests suggest, then the relation should also hold within groups. Two indices of habit-strength difference were computed for each S: W - C and W/C. Correlations with interference for Group 1 (where a net facilitation occurred, due to practice effects) were .124 (n.s.) and .402 ($p = .01$, one-tailed) for W - C and W/C, respectively. For Group 2, with net interference, r s were .115 (n.s.) and .331 ($p = .03$), respectively. Thus, correlations with W/C support the

differential habit-strength position within groups. It is noteworthy that several analogous correlations presented by Jensen (1965; r s between E, and F, H, and J) for the standard Stroop did not exceed .06 (n.s. for $N = 436$).

One of the appeals of the standard Stroop has been its high reliability, even though there is disagreement on the interpretation of the individual differences it measures (see Jensen & Rohwer, 1966). Jensen (1965) reported 1-day test-retest reliabilities (R_i) between .84 and .86 for C, W, and CW. An estimate of immediate test-retest reliability for the scanning Stroop was obtained by correlating the green and blue trials for each card. Results were not significantly different from Jensen's: r s = .87 for C, .90 for W (N s = 60), and .83 and .87 for CW Groups 1 and 2, respectively (N s = 30).

Correlations between cards, however, were quite different. Shown in Table 1, r s for the scanning Stroop are almost all significantly higher than corresponding r s for the standard Stroop test (Fisher r to Z , $ps < .001$, except for .84 vs .76, n.s.). For both tests the best predictor of CW time is the time on the card calling for the CW response, without the interfering stimuli (.94 > .85, one-tailed $p = .005$; .92 > .84, $p = .023$; .76 > .51, $p < .001$).

DISCUSSION

The predicted reversal of the Stroop interference effect clearly occurs when the task is scanning rather than naming each instance on the cards. Since the prediction was based on known differences in scanning times for words and colors, its confirmation seems to support the differential habit-strength theories of Stroop interference. The analyses of individual differences also offer some support for these positions; Ss with larger W/C ratios showed more interference on CW. If the interference reversal on the scanning Stroop results simply from reversed habit strengths, then individual differences in interference in Group 2 (words on CW; interference based on CW and W scores) should correlate with the same variables that standard Stroop interference (ink on CW; interference based on CW and C scores) correlates with. Thus, it should be possible to replicate Loomis

Table 1
Correlations Between Times on C, W, and CW for the Scanning Stroop and Standard Stroop Tests

	Scanning Stroop Test				Standard Stroop from Jensen, 1965 ($N = 436$)	
	Group 1 ($N = 30$)		Group 2 ($N = 30$)		C	W
	C	W	C	W		
W	.90		.92		.54	
CW	.94	.85	.84	.92	.76	.51

Note— r s for the scanning Stroop are means of the two r s for green and blue trials.

and Moskowitz's (1958) findings for flexible vs constricted control, and Broverman's (1960a, b) findings for strong vs weak automatizers, with the scanning Stroop.

However, there is reason to believe that the scanning Stroop is not just a simple reversal of the standard Stroop. With the standard Stroop, Jensen (1965) found no significant correlation between interference and W/C or W - C; and correlations between cards were significantly lower than on the scanning Stroop. Theoretically, the scanning Stroop represents a simpler task. On each card on each trial, S is provided with the correct response (e.g., "red") and must decide only whether or not to make it for each stimulus. For C he must match the stimulus with an image of the color; for W the match is with an image of the word. The difference in times on these two cards must reflect differential difficulty in the match-mismatch decision, not differences in selecting the correct verbal response. On the standard Stroop, S

must select the correct verbal response from among the three to five alternatives typically involved on C and CW. On W the correct response is, in effect, provided almost directly by the stimulus. This difference in C and W processes on the standard Stroop appears to be implicit in Broverman's (1960a, b) index of conceptual vs perceptual motor dominance, with C time representing conceptual processes and W representing simply perceptual motor. By this analysis, the scanning Stroop would not provide analogous indices.

The proposed similarity of C and CW processes on the standard Stroop is supported by the significantly higher correlation between C and CW than between W and CW (see Table 1). And the broader implication of the analysis, that the processes in the standard Stroop are more complex than in the scanning Stroop, is supported by the higher correlations among cards for the scanning Stroop, and by the fact that interference practically

doubles CW times on the standard Stroop, while having a relatively slight effect on the scanning Stroop (see Fig. 1, Group 2).

REFERENCES

- BROVERMAN, D. M. Cognitive style and intra-individual variation in abilities. *Journal of Personality*, 1960a, 28, 240-256.
 BROVERMAN, D. M. Dimensions of cognitive style. *Journal of Personality*, 1960b, 28, 167-185.
 JENSEN, A. R. Scoring the Stroop test. *Acta Psychologica*, 1965, 24, 398-408.
 JENSEN, A. R., & ROHWER, W. D., JR. The Stroop color-word test: A review. *Acta Psychologica*, 1966, 25, 36-93.
 LOOMIS, H., & MOSKOWITZ, S. Cognitive style and stimulus ambiguity. *Journal of Personality*, 1958, 26, 349-364.
 LUND, R. H. The role of practice in the speed of association. *Journal of Experimental Psychology*, 1927, 10, 424-433.

NOTES

1. All F ratio df in this section equal 3,156, until otherwise noted.
2. All F ratio df in the remainder of this paper equal 1,52.

(Accepted for publication July 6, 1970.)