The color-word interference test and its relation to performance impairment under auditory distraction

RICHARD I. THACKRAY, KAREN N. JONES
and ROBERT M. TOUCHSTONE
Civil Aeromedical Institute, FAA, Oklahoma City, Okla. 73125

The present study examined the relationship between performance on the Stroop color-word interference test (a suggested measure of distraction susceptibility) and impairment under auditory distraction on a random generation task. Although there was a significant decrease in "randomness" as a result of auditory distraction, the correlation between change in randomness and amount of color-word interference was nonsignificant. This suggests that the Stroop test may measure a rather restricted type of perceptual interference essentially unrelated to a possibly more general ability to maintain concentration in the presence of competing (distracting) stimuli.

A previous study by Thackray & Jones (1971) described the development of a laboratory version of the Stroop test for use in distraction research and examined the influence of simultaneously presented relevant (conflicting color names) and irrelevant (random numbers) auditory distraction on the color-word interference effect. Although the method developed for presenting the visual color-word stimuli was successful in eliciting the characteristic Stroop effect (Stroop, 1935; Jensen & Rohwer, 1966), there were no significant performance or physiological differences between the group that received the standard Stroop conditions and the groups which received the standard Stroop conditions plus relevant or irrelevant auditory distraction. This indicated that the addition of auditory "distraction" did not augment or modify the basic effect in any way.

If the color-word interference test measures a form of *general* ability to sustain attention in the presence of interfering stimuli, one might have expected the addition of the competing auditory stimuli to have resulted in at least some increase in response times to the visual stimuli. The lack of effect of the auditory stimuli suggests that the Stroop test may reflect susceptibility to a rather specific kind of perceptual interference. If this is the case, the test may possibly have limited usefulness as a measure of general distractibility.

The present study was conducted to evaluate this possibility. Susceptibility to interference, as measured by the Stroop test, was compared with extent of performance impairment on a task known to be adversely affected by distraction. The task chosen was random generation of letters of the alphabet. This task has been used in several recent studies concerned with deployment of attention and has been shown to be quite sensitive to the effects of distracting auditory stimuli (Schimek & Wachtel, 1969; Wolitzky & Spence, 1968). The usual procedure consists of instructing Ss to try to generate letters of the alphabet (or numbers) in "random" order at some E-determined rate. Amount of change in randomness under auditory distraction reflects the degree of success with which Ss are able to sustain attention to the primary task and exclude or ignore the irrelevant auditory stimuli.

METHOD

Fifty paid male university students served as Ss. All were right-handed and had no reported color-vision or hearing deficiencies.

Apparatus for the random generation (RG) task consisted of a pair of headphones for presenting the task instructions and the distraction stimuli, a microphone, and a pair of small "stimulus" lights for pacing the S's responses. The lights were located directly in front of the S and flashed momentarily every 2 sec. Leads from the microphone were connected to an amplifier and a second set of headphones to enable monitoring and recording of the S's verbal responses.

The apparatus and procedure for administering the Stroop stimuli have been discribed previously (Thackray & Jones, 1971). In essence, an automatic slide projector was employed for presenting the conventional Stroop word (W), color (C), and color-word (CW) stimuli. The S responded to each stimulus by pressing one of four buttons that had the stimulus words printed above them. Reaction time and response (correct or incorrect) to each slide were recorded on paper tape by means of a Welford Mark V SETAR.

All Ss received the RG task prior to being administered the Stroop test. This was felt desirable in order to eliminate the possible influence of the particular quasirandom order employed with the Stroop stimuli on the S's conception of randomness. In the instructions for the RG task, the S was told that his task would be to generate a series of random letters, using all 26 letters of the alphabet. The S was given Baddeley's standard instructions (Baddeley, 1966), in which he was asked to imagine that on each trial he was drawing a letter from a hat, saying the letter out loud, and returning the letter to the hat so that on each trial every letter would be present and have an equal chance of being chosen. He was also asked to keep in mind that such a series of letters would be completely random and would not be likely to consist of words, alphabetic sequences, etc. The S was informed that the whole task would take about 25 min and would be divided into three parts, with a short rest period between parts. (Each part contained 150 trials and lasted approximately 5 min with 2-min rest periods between parts.) Following a practice series of 20 trials, the S performed the first part in silence. At the beginning of the second part, he was informed that he would hear random letters through his headphones, but that he was to try to ignore them. A continuous 5-min tape recording consisting of the letters B, D, F, G, I, K, M, N, Q, R, T, V, and Y, arranged in a random order, was presented to the S during this part. Intervals between letters varied randomly from approximately 0.5 sec to 1.0 sec. The third part was identical to the first. At the end of the RG task, the instructions for the Stroop test were presented and the test administered.

Response times to the 72 stimuli in each part of the Stroop test were obtained for each S, and means were computed. Randomness over the 150 trials in each part of the RG task was measured by the entropy formula $H = \log_2 N - (1/N) \Sigma n_i \log_2 n_i$, where N is the number of trials and n_i is the frequency of usage of each letter of the alphabet. The higher the value of H, the more random the series (Attneave, 1959).

RESULTS

Mean H values for the three parts of the RG task are shown in Table 1. As

Table 1
Mean H Values for the Predistraction, Distraction, and Postdistraction Parts

Part	H Values	SD
Predistraction	4.4052	0.1496
Distraction	4.3558	0.1282
Postdistraction	4.4034	0.1443

expected, the effect of the auditory distraction was to reduce randomness.

A repeated-measures analysis of variance revealed this reduction to be significant ($F = 5.12$, $df = 2,98$, $p < .01$). Although the magnitude of the effect appears small, the H values obtained for the nondistraction and distraction parts are virtually identical to those obtained by Schimek & Wachtel (1969) under comparable conditions.

A baseline measure of randomness was obtained for each S by combining his pre- and postdistraction H values. This was deemed appropriate, since there was a reasonable degree of response consistency between these two parts ($r = .58$, $p < .01$), and Tukey's HSD test (Kirk, 1968) revealed the differences to be nonsignificant ($p > .05$). Change in randomness was determined by subtracting H values for the second part from mean values for the pre- and postdistraction parts. Although reliability of this derived score was not determined, Schimek & Wachtel (1969), using conditions comparable to those of the present study, found individual differences in change in H values to be consistent (product-moment correlations of 0.51 to 0.74) across a variety of distraction conditions.[1]

For the color-word interference test, mean response times were 851 and 1,015 msec for the C and CW parts, respectively. These values closely approximate those obtained for the comparable stimulus conditions in the previous study by Thackray & Jones (1971).

Although a variety of scores have been suggested as measures of the color-word interference effect, a factor analysis of these measures by Jensen (1965) has demonstrated a simple difference score (CW-C) to be the most effective measure of the interference effect. Consequently, the product-moment correlation between this measure of color-word interference and the difference scores on the RG task was computed.[2] Although the correlation was positive, it was quite low and nonsignificant ($r = .12$, $p > .05$). No improvement was obtained when the same scores for both tests were expressed in terms of percent change.

DISCUSSION

The results of the present study confirm earlier findings (Schimek & Wachtel, 1969; Wolitzky & Spence, 1968) that the ability to generate random letters or digits is significantly impaired when Ss are required to perform this task in the presence of auditory distraction. Individual differences in the extent of this impairment, however, were found to be completely unrelated to differences in the magnitude of color-word interference on the Stroop test. This lack of relationship supports the implications of the results obtained in the previous study by Thackray & Jones (1971), that the Stroop test reflects susceptibility to a limited kind of perceptual interference that may be essentially unrelated to what is commonly thought of as distractibility.

In a factor analytic study designed to investigate possible correlates of field dependence-independence, Karp (1963) identified two clusters of factors that were associated with two rather different types of visual distraction situations. One cluster of factors was represented in general by tests in which the critical stimulus is presented in the presence of irrelevant stimuli that compete with, but do not distort or modify, the basic properties of the central stimulus. An example of such tests would be the digit symbol subtest of the Wechsler Adult Intelligence Scale. Presumably, tests loading on this cluster reflect an ability typically implied by the common conception of concentration, i.e., the ability to sustain attention in the presence of potentially interfering ("distracting") stimuli.

The second cluster of factors was represented by tests in which the figural properties of the central stimulus are actually changed by the irrelevant stimuli and new, competing Gestalts are formed. An example would be the embedded-figures test. Although some degree of correlation exists between these two clusters of factors, Karp apparently feels that the ability to overcome the effects of embedding contexts represents an ability that is factorily different from the ability to sustain concentration in the presence of "distracting" stimuli.

While Karp did not employ the Stroop test in his factor analysis, other investigators have examined the relationship between this test and the embedded-figures test (Jensen, 1965). Moderate correlations, ranging from 0.36 to 0.54, have generally been reported. This would suggest that the Stroop test might well have loaded on the same factors as the embedded figures test, had it been included in Karp's study. It might also suggest that had the embedded-figures test been employed in the present study, it would have been unrelated to performance change on the random generation task under auditory distraction. This, in fact, was one of the findings of the Schimek & Wachtel (1969) study. Their results failed to support the hypothesis that field-dependent Ss (as determined by scores on the embedded-figures test) would show greater impairment on the random generation task than would field-independent Ss. No relationship whatsoever was found between any of the measures of field dependence and either baseline levels or change in randomness under distraction.

Wachtel (1967) has noted that a controversy exists as to whether such tests as the Stroop test and the embedded-figures test primarily reflect the ability to extract items from embedding contexts or whether they represent a more general capacity to direct attention selectively to relevant rather than competing irrelevant stimuli. The findings of the studies reviewed here, taken together with those of the present investigation, strongly suggest that "distractibility" as measured by the color-word interference test may be more closely related to the rather restricted ability to overcome the effects of embedding contexts than to the more general capacity to attend to a task in the presence of competing irrelevant stimuli. Additional support for this is provided by Mandell (1966), who found that performance of children on the Stroop test was unrelated to teacher ratings of distractibility.

More promising, perhaps, as a measure of distractibility is the task used in the present study as the "criterion" measure. The ability to generate random letters or digits has been clearly shown to be impaired in the presence of auditory distraction. As Schimek & Wachtel (1969) suggest, the measure of randomness appears to be a promising one for the study of individual differences in attention deployment. Further research, using change in randomness under distraction as a predictor variable, would seem to be indicated.

REFERENCES

ATTNEAVE, F. *Application of information theory to psychology.* New York: Holt, Rinehart & Winston, 1959.

BADDELEY, A. D. The capacity for generating information by randomization. Quarterly Journal of Experimental Psychology, 1966, 18, 119-129.

FRANKENHAEUSER, M., FROBERG, J., HAGDAHL, R., RISSLER, A., BJORKVALL, C., & WOLFF, B. Physiological, behavioral, and subjective indices of habituation to psychological stress. Physiology & Behavior, 1967, 2, 229-237.

JENSEN, A. R. Scoring the Stroop test. Acta Psychologica, 1965, 24, 398-408.

JENSEN, A. R., & ROHWER, W. D. The Stroop color-word test: A review. Acta Psychologica, 1966, 25, 36-93.

KARP, S. A. Field dependence and overcoming embeddedness. Journal of Consulting Psychology, 1963, 27, 294-302.

KIRK, R. E. *Experimental design: Procedures for the behavioral sciences.* Belmont, Calif: Brooks/Cole, 1968.

MANDELL, S. J. *Children's resistance to competing and distracting stimuli in the classroom.* (Doctoral dissertation, University of Southern California) Ann

Arbor, Mich: University Microfilms, 1966, No. 66-11, 578.

SCHIMEK, J. G., & WACHTEL, P. L. Exploration of effects of distraction, competing tasks and cognitive style on attention deployment. Perceptual & Motor Skills, 1969, 28, 567-574.

STROOP, J. R. Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 1935, 18, 643-662.

THACKRAY, R. I., & JONES, K. N. Level of arousal during Stroop performance: Effects of speed stress and "distraction." Psychonomic Science, 1971, 23, 133-135.

WACHTEL, P. L. Conceptions of broad and narrow attention. Psychological Bulletin, 1967, 68, 417-429.

WOLITZKY, D. L., & SPENCE, D. P. Individual consistencies in the random generation of choices. Perceptual & Motor Skills, 1968, 26, 1211-1214.

NOTES

1. It was felt that, because of the nature of the data, the use of a split-half technique to determine reliability of H-value change scores in the present study would be based upon too few responses to yield a valid estimate of response consistency.

2. Split-half reliability of the CW-C score as well as reliabilities of the C and CW scores were 0.65, 0.93, and 0.92, respectively ($p < .01$).