Acta Psychologica 58 (1985) 225-235
North-Holland

THE EFFECTS OF EFFORT ON STROOP INTERFERENCE *

David P. MacKINNON, R. Edward GEISELMAN
and J. Arthur WOODWARD
University of California, Los Angeles, USA

Accepted August 1984

Stroop interference was defined as the difference in time needed to name the ink colors of printed color and color-related words versus control plus signs. The effect of effort on Stroop interference was studied using an inter-subject competition procedure designed to manipulate effort. In experiment 1, subjects in the competition group were successful at inhibiting Stroop interference when compared to the performance of subjects in the no-competition group. This result is consistent with theories that postulate attentional effects on Stroop interference. In experiment 2, the significant decrease in Stroop interference was accompanied by a significant reduction in recognition memory for Stroop list items. Therefore, Stroop interference was reduced at a stage during the processing of word meaning. This result is consistent with theories that locate Stroop interference before response output.

The purpose of this research is twofold: first, to investigate the effect of effort on Stroop interference; and second, to study the locus of the mechanism by which effort may influence Stroop interference. Stroop (1935) demonstrated that the time for naming the color of ink with which color words are incongruently written is greater than the time for naming the colors of squares or other color-irrelevant symbols. The difference between the performance times with these two kinds of stimuli is called Stroop interference, or simply the Stroop effect. In the 49 years since the discovery of this phenomenon, the Stroop effect has received considerable attention in cognitive psychology; first of all, in an attempt to understand the effect itself (Seymour 1977; Stirling 1979) and alternatively, in the study of other constructs such as automatic

* This research was supported in part by Grant BNS 7910053, funded jointly by the National Science Foundation and Office of Naval Research.
Mailing address: R.E. Geiselman, Dept. of Psychology, University of California, Los Angeles, CA 90024, USA.

processes (Posner and Snyder 1975), hemispheric dominance (Perret 1974), and cognitive styles (Broverman 1960). Jenson and Rohwer (1966) and Dyer (1973) provide extensive reviews of the Stroop effect and its applications.

In this research, we are asking two related questions about the nature of the effect of effort on Stroop interference. First, what happens when subjects expend considerable effort to perform as rapidly as possible on all materials (i.e., on color word items as well as on control plus signs)? Is the performance difference between the types of items (the interference) equal to, greater than, or less than that observed for subjects performing the task under a lower level of task-specific effort? Although Stroop performance is known to be malleable to some extent, such as with practice (Stroop 1935), it is not clear from previous research what effects effort would have on Stroop interference. Subjects might be able to reduce Stroop interference by narrowing attention to the relevant color cues and inhibiting semantic processing. The second question concerns the mechanism by which effort may affect Stroop interference. If an effect of effort on Stroop interference is found, then it is important to learn about the locus of the effect; specifically, to learn if Stroop interference occurs at response output or at a prior stage of processing. To address this latter question, subjects at different levels of effort were measured for their memory of the color-word stimuli, as well as on Stroop interference in experiment 2.

Effort and Stroop interference

The effect of effort on Stroop interference has been studied directly, and indirectly, in a variety of experimental paradigms. At present, the results of these studies are equivocal, with some studies reporting a positive effort–interference relationship, and others suggesting a negative effect of effort on Stroop interference. An added degree of ambiguity in the literature stems from the fact that in some studies the connection with effort is indirect. Thus, one purpose of the present experiment was to devise an innovative method for manipulating task-specific effort in the context of the Stroop task. The method devised is based upon an inter-subject competition procedure.

Hartley and Adams (1974) reported increased Stroop interference with greater arousal, where high arousal was produced with 30 minutes of 100 db noise. Pollack et al. (1975) also found increased Stroop interference with greater arousal produced by the threat of impending electric shock. Other researchers have found a decrease in Stroop interference with effort. Logan (1980), for example, used a priming paradigm to illustrate that the interference can be brought under some attentional control. Agnew and Agnew (1963) increased drive with threat of electric shock contingent upon performance and also found significant reductions in ink-color naming times.

Pilot work

Prior to the experiments reported here, two pilot studies were necessary to develop the experimental procedure for manipulating effort. Attempts to manipulate effort through payment of a one or five dollar reward were unsuccessful as was a negative-reinforcement procedure in the second pilot study. In these two procedures, the low and high effort groups did not differ significantly on the self-report measure of effort administered following the experimental procedure. The current experimental manipulation of effort, which is detailed in the next section, evolved out of this pilot work. One final suggestion of the pilot work was that previously cited gender differences (McKeachie 1961) might be relevant here. Thus, in the two experiments reported, gender served as a blocking factor in order to identify effects that might otherwise be mediated by or masked by gender.

Experiment 1

In experiment 1, it was expected that competition and the existence of a desired reward would lead to greater S effort than a no-competition procedure. As a result of this task-specific effort manipulation, Stroop interference should be reduced if the interference can be brought under attentional control.

Method

Subjects

Sixty-four Ss from the introductory psychology course at UCLA participated in this experiment. They were randomly assigned to two groups of 32 with equal numbers of males and females in each group.

Materials and apparatus

Four lists of stimuli were used: a practice list consisting of color-neutral words, a control list containing groups of four same-colored plus-signs in the place of words (Stirling 1979) and two Stroop lists, each of which consisted of some color words (e.g., red, green, black) and some color-related words (e.g., sky, carrot). On each list, there were two columns of 20 items with the horizontally adjacent items being the same words in different colors. The two Stroop lists were equated for the number of color words, the number of color-related words, the length of the words, and the number of same-colored items such as black and night.

The Stroop lists are presented in Appendix A. The list items were written in four colors: red, green, yellow, and blue, with each color appearing ten times on each list. No color appeared twice in a row and no identifiable pattern of colors existed on any list. One-inch vinyl Helvetica colored letters were placed on a white board background.

The measurement instrument used in this experiment was an effort scale developed by Geiselman et al. (1982). The effort scale consists of four items in which the *S* estimates his or her level of effort in relation to other contexts, e.g., memorizing a short poem. This self-report measure was found to be correlated 0.9 with GSR and 0.5 with heart-rate variability, two physiological measures believed to be representative of the expenditure of cognitive effort (Geiselman et al. 1982).

Procedure

All *S*s were screened for color blindness using six plates from the Ishihara color system. When this test was completed, the Stroop-task portion of the experiment was performed. *S*s sat in a chair approximately 8 feet from the Stroop lists that were placed on a table at eye level. The *S*s were instructed to say aloud the ink colors of the words or plus signs on the lists as fast as possible without making any errors. They were informed that there would be 40 items on a list, colored red, green, blue, and yellow, and that they would be timed with a stopwatch. The order of presentation of the Stroop and control lists was counterbalanced across *S*s. After the last list was finished, the *S*s completed the effort questionnaire.

*S*s were randomly assigned to receive one of the two following incentive procedures.

Low-incentive procedure. The *S*s were tested individually in the low-incentive condition. On entering the experimental room the following instructions were read to them: "This is just a pilot or practice procedure and is not designed to measure your specific performance. I just want you to follow my instructions correctly. At a later time, an experiment will be performed using a procedure similar to this one. But for now, I just want to see how this procedure works so we can design the final experiment to be run later with other students."

The *S*s' verbal performance in this condition was recorder on a Sony tape recorder and was timed later with a stopwatch reliable to 0.01 sec. Since the recorder was turned on before the *S* entered the room and turned off after the experiment, the *S*s were unaware of its presence.

High-incentive procedure. *S*s were tested in pairs in the high incentive procedure. The following instructions were read to them:

"You two are about to participate in a competitive game. Both of you will perform a task that requires you to respond to items on lists. Success on these lists has nothing to do with intelligence or your sex or anything special about an individual. It just depends on how much you put into it.

Each of you will do this experiment separately. It is important to do these tasks correctly and as fast as possible. The winner, who is 'the fastest', will receive one extra credit if you want it. Good luck and do your best. So, both of you will do the experiment and one of you will leave with two credits."

As part of the requirements of the introductory psychology course at UCLA, each *S* must participate in 6 hours of experiments. The additional hour of credit offered to the winner thus was a highly salient reward.

A coin was flipped to determine which *S* would do the experiment first. While one *S* was tested in the Stroop procedure, the other waited outside for 20 minutes until it was his/her turn.

After the practice list, the *S*s were told: "That is the end of the practice. Your times on the next two lists will be used to determine the winner of the game. A correction to your time will be made for any errors you make. Remember your performance depends on how much you put into it. Good luck."

Results and discussion

There were four ink-color naming errors in total made by the *S*s in experiment 1, two in the high-incentive group and two in the low-incentive group. To compensate for these errors in the time variable, an additive correction to the times was made. The correction for each error was set equal to 1.5 times the total time on the list before correction divided by the number of items on the list (Stroop 1935). For each error, this increases the total time by 1.5 times the time taken for one list item. Because of the low number of errors, the analyses of the data without corrections yielded the same conclusions as those described below.

The average performance values are shown in table 1. The means for the control and Stroop lists are presented for completeness. Gender (a blocking factor) and incentive group (the experimental manipulation) were the independent variables in the 2×2 analysis of variance design. The Gender by Effort Group interaction and the gender

Table 1
Performance means for experiment 1 as a function of incentive condition and gender of subject.

	Incentive					
	High			Low		
	Female	Male	Average	Female	Male	Average
Effort questionnaire (1–20 scale)	15.2	16.5	15.9	13.7	13.1	13.4
Stroop interference (in sec)	8.4	6.8	7.6	10.3	10.0	10.2
Stroop list time (in sec)	27.7	26.0	26.8	31.7	32.5	32.1
Control list time (in sec)	19.4	19.2	19.3	21.4	22.4	21.9

main effect were nonsignificant for all dependent variables. The *S*s in the high-incentive condition reported greater effort ($F(1,60) = 18.0$, $p < 0.001$) and showed lower Stroop interference ($F(1,60) = 4.51$, $p < 0.05$) than the *S*s in the low-incentive condition. Thus, the *S*s were able to inhibit Stroop interference with greater effort. This result is consistent with the view of Logan (1980) that a significant component of the Stroop task can be brought under attentional control.

Experiment 2

Explanations of Stroop interference can be divided into those that postulate interference at an early stage where degrees of semantic processing of the stimuli can be altered, and those that postulate interference at the late stage of response output. There is some evidence in the literature for each position (see Dyer 1973 for a review). Our intention is to study the locus of the effect of effort on Stroop interference. By doing so, we will provide evidence as to the locus of Stroop interference itself.

If effort affects Stroop interference at an early stage where the degree of semantic processing can be altered, then changes in Stroop interference should be accompanied by changes in long-term memory for the words in the Stroop list. If effort affects Stroop interference after meaning extraction, at the stage of response output, then changes in Stroop interference would not be accompanied by a corresponding change in memory for the list words. Memory performance was measured for each *S* in experiment 2 with a yes–no recognition test. The *S*s were not informed of this test prior to its administration at the end of the experiment.

Method

Subjects

Thirty-six *S*s from the introductory psychology course at UCLA participated in this experiment. They were randomly assigned to two groups of 18 with equal numbers of males and females in each group.

Materials and apparatus

Five lists of stimuli on a grey background were used: a practice list consisting of color-neutral words, a control list of colored plus-signs, and three Stroop lists. Each list had 12 items repeated four times for a total of 48 items. The twelve words on the Stroop lists were composed of 6 color words (e.g., red, green, black) and 6 color-related words (e.g., sky, carrot). The lists were equated on content and colors as in experiment 1.

The recognition memory test consisted of words from the three Stroop lists and words associated in color with Stroop list words. For each *S*, the recognition test consisted of 6 color words from the list, 6 color-related words from the list, 12 words associated in color with words from the list (e.g. for brown, bear and dirt) 6 non-associated, color-related words and 6 non-associated, color words. The recognition test allowed detailed analysis of recognition performance since responses could be broken down into five categories: color-word hits, color-related word hits, false alarms to list color associates, false alarms to non-associated, distractor, color words and false alarms to non-associated, distractor, color-related words.

The Stroop list words and associates are presented in Appendix B. List items were written in four colors: red, green, yellow and blue.

Procedure

The procedure for experiment 2 was identical to experiment 1 with the exception that the *S*s said the ink colors on the list twice for a total of 96 responses and then were given an unanticipated recognition memory test approximately 5 minutes after the last list was finished.

Results and discussion

There were three errors in the low incentive group and the same correction was made as in experiment 1. Again, analysis of the data without corrections yielded the same pattern of results described below.

The average performance means are displayed in table 2. Gender and incentive group were again the independent variables in a 2 × 2 design. The main effect of gender and the Gender by Incentive group interaction were not significant. *S*s in the high-incentive group reported greater effort ($F(1,32) = 10.96$; $p < 0.005$) and exhibited lower Stroop interference ($F(1,32) = 5.99$; $p < 0.05$) than *S*s in the low-incentive group. These results replicate those of experiment 1.

One possible interpretation of these results is that *S*s in the high incentive group worked faster in general than those in the low incentive group. The computation of the interference measure was used to assure that the difference would reflect Stroop interference and not the speed of naming. This measure would be problematic, however, if there was a ceiling effect on the Stroop list or a floor effect on the control

Table 2
Performance means for experiment 2 as a function of incentive condition and gender of subject.

	Incentive					
	High			Low		
	Female	Male	Average	Female	Male	Average
Effort questionnaire (1–20 scale)	16.1	17.0	16.6	13.9	13.3	13.7
Stroop interference (in sec)	24.1	27.5	25.8	33.6	35.2	34.4
Stroop list time (in sec)	78.1	84.5	81.2	95.9	100.1	98.0
Control list time (in sec)	54.0	57.0	55.5	62.3	64.9	63.6

list. Evidence for such effects would be truncations of the list time distributions. Only the Stroop list time distribution for the high-incentive group in experiment 1 was skewed significantly. This distribution was skewed opposite to the prediction of the above interpretation.

The finding that increased task specific effort leads to decreased Stroop interference has important implications for studies that utilize Stroop response times to compare non-equivalent groups. For example, Wise et al. (1975) found reduced interference for college students versus elementary school students and Golden (1976) found color naming differences among normal controls, psychiatric patients and brain damaged clients. These findings could be the result of differences in effort between groups. Thus, effort is an important mediating variable to be considered in all Stroop studies.

The memory performance means are displayed in table 3. Gender and incentive group were the between factors and word type was the within factor in a $2 \times 2 \times 5$ design. The main effect of word type ($F(4,29) = 55.8$; $p < 0.0001$) and the Word type by Incentive-group interaction ($F(4,29) = 4.29$; $p < 0.05$) were significant. Both the high-incentive ($F(1,32) = 12.5$; $p < 0$ 001) and low-incentive groups ($F(1,32) = 5.13$; $p < 0.001$) had a significantly higher proportion of hits overall than false alarms. Analysis of the simple main effects revealed a significant difference between effort groups for color-related word hits ($F(1,32) = 14.3$; $p < 0.001$), but not false alarms, indicating that *S*s in the high-incentive group reduced Stroop interference by inhibiting semantic processing of the list items.

Another interpretation of this result is that the increased effort to execute a response quickly, results in less resources for memory consolidation of list items. This interpretation is inconsistent with the levels of processing framework where the depth or qualitative aspects of processing predominantly determine the level of retention, not response time, effort or other quantitative aspects (Craik and Tulving 1975). The results do suggest that with effort, *S*s can alter the depth of processing to a more shallow level which results in decreased semantic processing of Stroop list items.

*S*s did not false alarm to list color associates more than non-associated color-related words in either the high-incentive group ($F(1,32) = 0.028$; $p > 0.5$) or in the low-incen-

Table 3
Recognition memory performance (proportions).

Word type	Incentive					
	High			Low		
	Female	Male	Average	Female	Male	Average
Hits						
Color words	0.87	0.85	0.86	0.89	0.96	0.93
Color-related words	0.52	0.41	0.46	0.76	0.74	0.75
False alarms						
List color associates	0.22	0.27	0.25	0.32	0.30	0.31
Color words	0.67	0.63	0.65	0.48	0.70	0.59
Color-related words	0.28	0.22	0.25	0.20	0.26	0.23

tive group ($F(1.32) = 3.23$; $p > 0.05$). This indicates that the automatic processig of list items in the Stroop task is specific to the list items and does not generalize to words related in color. However, there was a bias to consider any color word as a hit for both the high- ($F(1,32) = 37.7$; $p < 0.0001$) and low-incentive groups ($F(1,32) = 17.47$; $p < 0.001$).

The reduction of Stroop interference at a stage prior to response output is consistent with many current findings (Williams 1977; Seymour 1977; Stirling 1979). It should be noted that our results do not identify a single locus for Stroop interference, but rather indicate that one locus of the interference is at a stage before response output.

Conclusions

Thus, the principal findings of this research are that subjects can reduce Stroop interference when their effort is directed to do so, and inhibition of Stroop interference is associated with a suppression of list word meaning. These two results taken together suggest that at least one significant component of Stroop interference can be localized at a stage of processing prior to response output.

The reduction of Stroop interference in this experiment may be the result of the task-specific nature of the effort procedure used. Previous experiments that obtained increased Stroop interference used procedures that would likely increase general arousal levels, such as noise or threat of impending shock. In contrast, the Agnew and Agnew (1963) and Logan (1980) experiments used paradigms where the incentive was specific to Stroop performance. With the present task-specific procedure, the subjects apparently were capable of restricting attention somewhat to the color cues. Semantic processing was not completely restricted as the Stroop effect remained. These results are consistent with Easterbrook's (1959) formulation where higher task-specific arousal is said to cause attention to focus.

A final contribution of this research concerns methods for manipulating subject intensity in experimental tasks. The present inter-subject competition procedure was the third method tried in this research. Neither monetary incentive nor negative-reinforcement procedures were found to effectively alter subject effort in the Stroop task as measured by the effort questionnaire. The procedure developed here should prove useful in other experimental contexts. These results also illustrate the importance of measuring subject intensity to ensure that the incentive procedure has, in fact, been effective.

Appendix A

Stroop list #1

Grass, White, Fire, Blood, Coal, Tree, Grey, Green, Tan, Tomato, Banana, Oil, Tar, Night, Yellow, Sky, Corn, Dirt, Plum, Purple.

Stroop list #2

Lake, Blue, Gold, Orange, Crow, Chalk, Black, Lime, Carrot, Grape, Lemon, Sun, Red, Plant, Snow, Smoke, Apple, Moss, Brown, Pea.

Appendix B

Stroop list #1

Black, Blue, Brown, Carrot, Coal, Gravy, Green, Lime, Orange, Sky, Sun, Yellow.

Stroop list #2

Blue, Brown, Cherry, Grape, Gravy, Purple, Red, Snow, Sky, Sun, White, Yellow.

Stroop list #3

Black, Carrot, Cherry, Coal, Grape, Green, Lime, Orange, Purple, Red, Snow, White.

List color associates

Lilac, Plum, Pea, Plant, Denim, Sea, Corn, Lemon, Dirt, Bear, Apricot, Pumpkin, Tar, Crow, Blood, Tomato, Milk, Salt.

References

Agnew, N. and M. Agnew, 1963. Drive level effects of tasks of narrow and broad attention. Quarterly Journal of Experimental Psychology 15, 58–62.

Broverman, D.M., 1960. Dimensions of cognitive style. Journal of Personality 28, 167–185.

Craik, F.I.M. and E. Tulving, 1975. Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General 104, 268–294.

Dyer, F.N., 1973. The Stroop phenomenon and its use in the study of perceptual, cognitive, and response processes. Memory and Cognition 1, 106–120.

Easterbrook, J.A., 1959. The effect of emotion on cue utilization and the organization of behavior. Psychological Review 66, 183–201.

Geiselman, R.E., J.A. Woodward and J. Beatty, 1982. Individual differences in verbal memory performance: a test of alternative information processing models. Journal of Experimental Psychology: General 111, 109–134.

Golden, C.J., 1976. Identification of brain disorders by the Stroop color and word test. Journal of Clinical Psychology 32, 654–658.

Hartley, E.R. and R.G. Adams, 1974. Effect of noise on the Stroop test. Journal of Experimental Psychology 102, 62–66.

Jensen, A.R. and W.D. Rohwer, Jr., 1966. The Stroop color-word test: a review. Acta Psychologica 25, 36–93.

Logan, G.D., 1980. Attention and automaticity in Stroop and priming tasks: theory and data. Cognitive Psychology 12, 523–553.

McKeachie, W.J., 1961. 'Motivation, teaching methods, and college teaching'. In: M.R. Jones (ed.), Nebraska symposium on motivation. Lincoln, NE: University of Nebraska Press. pp. 111–142.

Perret, E., 1974. The left frontal lobe of man and the suppression of habitual responses in verbal categorical behavior. Neuropsychologia 12, 323–330.

Pollack, M.S., T.S. Pittman, J.E. Haller and P. Munson, 1975. The effect of arousal on Stroop color-word task performance. Bulletin of the Psychonomic Society 6, 248–250.

Posner, M.S. and C.R. Snyder, 1975. 'Attention and cognitive control'. In: R.L. Solso (ed.), Information processing and cognition. Hillsdale, NJ: Erlbaum. pp. 55–82.

Seymour, P.H., 1977. Conceptual encoding and locus of the Stroop effect. Quarterly Journal of Experimental Psychology 29, 245–265.

Spence, J.T. and K. Spence, 1966. 'The motivational components of manifest anxiety: drive and drive stimuli'. In: D. Spielberger (ed.), Anxiety and behavior. New York: Academic Press.

Stirling, N., 1979. Stroop interference: an input output phenomenon. Quarterly Journal of Experimental Psychology 31, 121–132.

Stroop, J.R., 1935. Studies of interference in serial-verbal reaction. Journal of Experimental Psychology 18, 643–662.

Williams, E., 1977. The effect of amount of information in the Stroop color word test. Perception and Psychophysics 22, 463–470.

Wise, L.A., J.A. Sutton and P.D. Gibbons, 1975. Decrement in Stroop interference time with age. Perceptual and Motor Skills 41, 149–150.