Specificity of Practice Effects in the Classic Stroop Color-Word Task

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 15(0)

Authors

Clawson, Deborah M.

King, Cheri L.

Healy, Alice F.

et al.

Publication Date

1993

Peer reviewed

Specificity of Practice Effects in the Classic Stroop Color-Word Task[1]

Deborah M. Clawson, Cheri L. King, Alice F. Healy, & K. Anders Ericsson
Department of Psychology
University of Colorado
Boulder, CO 80309-0345
dclawson@clipr.colorado.edu

Abstract

Specificity effects of practice on the classic Stroop color-word task were explored using two different practice tasks, practice on the Stroop task itself and practice on simple color naming. Clear evidence for specificity effects was found, and this specificity persisted across a one-month delay. Stroop practice, but not color-naming practice, led to a pattern of improvement pointing to an advantage for practiced stimuli over unpracticed stimuli on both Stroop and color-naming tests but a disadvantage for practiced stimuli on reading and "reverse-Stroop" tests. The advantage for practiced stimuli was maintained on versions of the Stroop test that used orthographic manipulations of the stimuli. This pattern of specificity is inconsistent with practice as specific to the word forms. It is consistent with practice as specific to colors, to semantic meanings of the words, or to a combination of these two.

Introduction

One of the most studied phenomena of the last two decades is the Stroop effect (Stroop, 1935). Its popularity as a research task is due both to its implications for automatic processing (e.g., Anderson, 1992) and its use in neuropsychological testing (e.g., Connor, Franzen, & Sharp, 1988). In the Stroop color-word interference task, subjects are asked to name the color of the ink in which color words are printed. The ink color and the word do not correspond. For example, given the word purple printed in red ink, the appropriate response is "red." Despite the large and increasing body of research on the Stroop effect (see Dyer, 1973; Jensen & Rohwer, 1966; MacLeod, 1991), the present study is the first to examine the extent to which practice either on the color-word interference task or on a simple color-naming task is specific to the particular colors used during training.

Although not investigated in the classic Stroop task itself, specificity of training has been explored in nonstandard versions of the Stroop task. For example, in a digit counting task, Reisberg, Baron, and Kemler (1980) trained subjects to ignore a pair of digits (e.g., 2 and 4) and found that this training did not transfer perfectly to ignoring other digits nor did it transfer to ignoring homophonic words (e.g., to and for). Some transfer was obtained, however, to the task of ignoring the digits printed as words. Ménard-Buteau and Cavanagh (1984) used a Stroop task with incongruously colored objects. Subjects practiced naming the ink color of a word representing an incongruously colored object (e.g., the word carrot printed in green ink). They found that this training did not transfer to a version of their task with drawings of the objects rather than words.

Our study involved the classic Stroop task, with training in two different situations. The color-patch training condition involved practice in simply naming color patches. The Stroop training condition involved practice in naming the colors of incongruent color words. All subjects were tested in a pretest prior to

[1] This research was supported in part by Army Research Institute Contract MDA903-90-K-0066.

training as well as in a posttest after training and in a retention test after a month-long delay. Each test session included a set of tests related to Stroop interference: one test each on reading and on simple color-patch naming plus a Stroop test and a test with Stroop stimuli but requiring reading responses ("reverse Stroop"). Additional orthographic manipulation versions of the Stroop test, one in which the letters of the color words were bracketed by asterisks and one in which the letters were in uppercase rather than lowercase letters, provided an indication of specificity to the word form. Another measure of specificity was provided by the use of two different color sets. Although the experimental subjects trained on only one color set (with the set counterbalanced across subjects), all subjects were tested on both color sets.

It was anticipated that specificity of training effects would lead to: (a) less improvement on the orthographic manipulations than on the normal Stroop test and (b) less improvement on the untrained color set than on the trained set. On the basis of our previous studies showing extremely good retention of procedural skills (Healy et al., 1992), we also predicted relatively little evidence of forgetting across the one-month delay interval. Of greatest interest is whether any specificity effects persist across this long retention interval.

Method

Subjects

Six female students from the University of Colorado participated for payment at the rate of $5.00 per hour. Each was assigned to a training condition and a trained color set on the basis of her time of arrival for testing according to a fixed rotation. All subjects demonstrated normal color vision.

Design

During the pretest, posttest, and retention sessions, 15 tests were administered, 12 of which were the focus of this investigation. The 12 focal tests were divided into two categories which partly overlapped: (a) 8 color interference measures, and (b) 6 orthographic color interference measures. A mixed-factorial design was employed for the 8 color interference measures, with one between-subjects factor, training condition, and three within-subjects factors, test time (pretest, posttest, retention test), color set (trained, untrained), and test type. The four test types were: (a) reading color words displayed in white on a black background, (b) color-patch naming, (c) naming colors of incongruous words (Stroop), and (d) reading incongruous color words (reverse Stroop). A second mixed-factorial design was employed for the 6 orthographic color interference measures, with one between-subjects factor, training condition, and three within-subjects factors, test time, color set, and orthographic test type (standard, asterisks, uppercase).

Apparatus

A DTK Data-1000 personal computer with a Zenith Data Systems color monitor was employed for training and tests. A MEL (Schneider, 1988) Version 5.0 voice key-button box and an Electret microphone were used for measuring the subjects' verbal response latencies and for recording the experimenter's indications of response accuracy.

Materials

Test materials were identical for pretest, posttest, and retention test, and consisted of 12 relevant tests, each consisting of 24 trials. The tasks were presented in a fixed order: color-patch naming, Stroop, asterisk Stroop, uppercase Stroop, reading, and reverse Stroop, with each task being first tested on the trained color set, then on the untrained color set. The color stimuli from Set 1 were: pink, blue, and orange. The colors from Set 2 were: purple, green, and red. Stroop tests and training used all six possible incongruous stimuli from a set. One of the two subjects in each condition was trained on Set 1, and the other was trained on Set 2.

There were two conditions of training: color-patch and Stroop. These conditions of training correspond to the tests on color-patch naming and on Stroop, respectively. Each training session consisted of 240 trials. Four versions of the training stimulus sequence were constructed, which differed only in the pseudorandom order of the stimuli. The four versions were shown in a fixed rotation, with every version shown once in each successive set of four sessions.

Procedure

The pretest was administered in the first experimental session. Training took place during the second through thirteenth sessions, on successive weekdays. During the fourteenth session, subjects were given the posttest, and they returned one month later for the retention test. All instructions and trials were self paced. Throughout all test and training trials, subjects were given a warning message whenever their vocal response was not registered by the voice key. Any response not registered by the voice key was discarded from the data analyses.

The instructions for each test were shown on the computer screen. For each test the instructions described the stimuli to be presented and the expected response. For the tests involving incongruous color words, the instructions included a single example from the appropriate color set. All instructions directed subjects to respond as accurately and as quickly as possible. The procedure on a given test trial was as follows: First, the subject pressed the space bar, after which the screen became blank for 300 ms, then the stimulus appeared on the screen. The subject responded, then the experimenter, who sat behind the subject, indicated the accuracy of the subject's vocal response by pressing either a "correct" or an "incorrect" key on the button box. The training procedure was the same as the procedure used for the tests except that feedback was provided, both on accuracy and reaction time, and verbal protocols were collected after the first and last 12 trials of each session.

Results

Performance on the color interference tests, averaged across the three training conditions, is displayed in Figure 1 in terms of log correct reaction times as a function of test time (pretest, posttest, retention test) and test type (reading, reverse Stroop, color-patch naming, Stroop). As in previous research, reaction times were faster for the test types involving reading than those involving color naming and were slower for the test types involving incongruous stimuli than for those that did not ($\underline{F}(3, 9) = 27.56$, $\underline{MS_e} = .0041$, $\underline{p} < .001$, on an analysis of the pretest and posttest; $\underline{F}(3, 9) = 21.84$, $\underline{MS_e} = .0036$, $\underline{p} < .001$, on an overlapping analysis of the posttest and retention test). Specifically, the test types in order of fastest to slowest were: reading, reverse Stroop, color-patch naming, and Stroop. Overall, subjects improved with practice; they were faster on the posttest than on the pretest ($\underline{F}(1, 3) = 16.37$, $\underline{MS_e} = .0047$, $\underline{p} < .05$), and there was no significant forgetting evident from the posttest to the retention test.

The specificity of training is reflected by three related observations. First, reaction

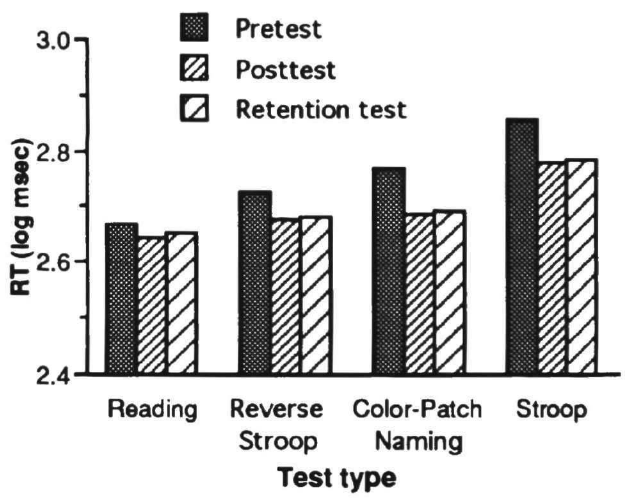

Figure 1. Correct reaction times, across training conditions, on the color interference tests as a function of test time.

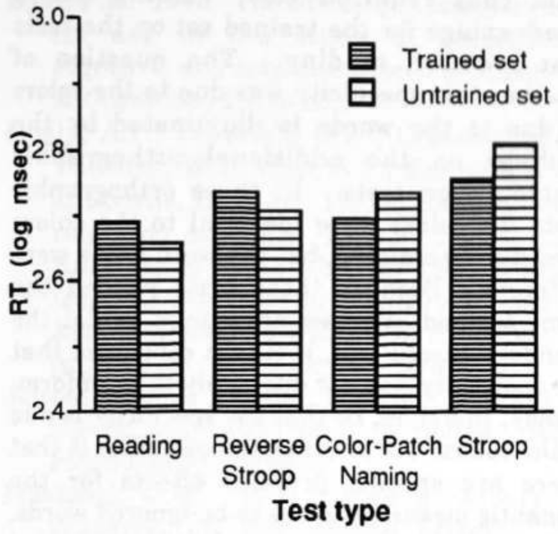

Figure 2. The Stroop-trained group's correct reaction times, after training, on the trained and untrained color sets of the color interference tests.

times decreased from pretest to posttest more for the color set on which subjects had trained than for their untrained color set, $F(1, 3) = 38.51$, $MS_e = .0001$, $p < .05$ (trained set: $M = 2.765$ pretest, $M = 2.698$ posttest; untrained set: $M = 2.744$ pretest, $M = 2.698$ posttest). Second, as illustrated by Figure 2, which presents data only from the Stroop-trained condition after training (i.e., on the posttest and retention test), subjects were faster on the trained set than on the untrained set when naming colors, that is in the Stroop test and in the color-patch naming test, but were not faster in either the reading test or the reverse Stroop test ($F(3, 9) = 6.91$, $MS_e = .0002$, $p < .05$, for this two-way interaction of color set and test type). In these last two tests, reaction times were actually faster for the untrained set than for the trained set. Third, this advantage for the trained set on color-naming responses and the advantage for the untrained set on reading responses after training were only found for subjects in the Stroop training condition, not for subjects in either the color-patch training or control conditions ($F(6, 9) = 6.95$, $MS_e = .0002$, $p < .01$, for the three-way interaction).

The results of the orthographic manipulation tests also were analyzed in terms of log correct reaction times. Importantly, there was no effect of orthographic test type ($M = 2.807$ for standard Stroop, $M = 2.810$ for asterisks, $M = 2.808$ for uppercase) and that factor did not enter into any significant interactions, suggesting that

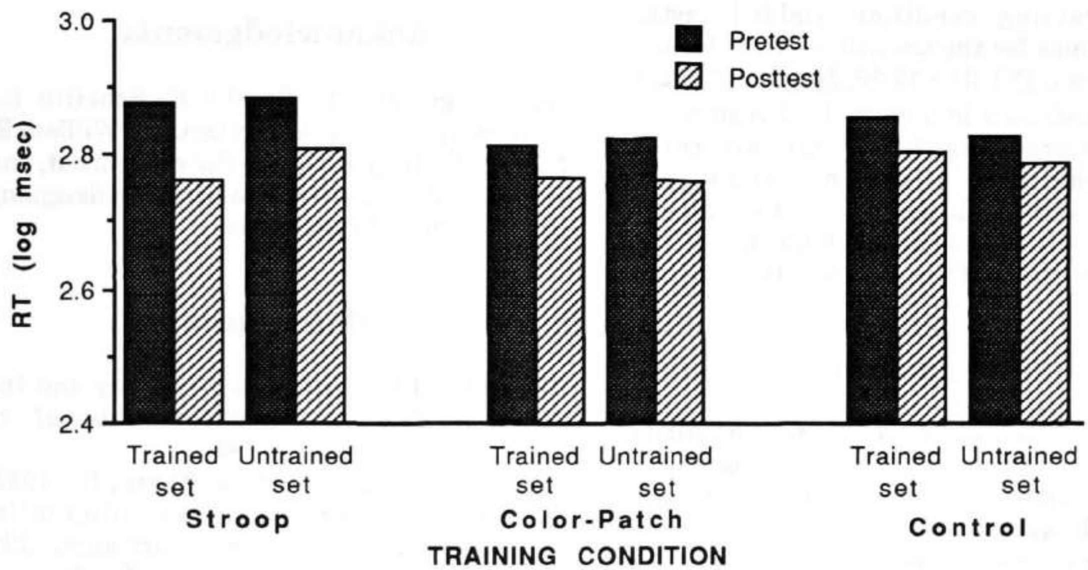

Figure 3. Correct reaction times, across the orthographic Stroop tests (standard, asterisks, uppercase), as a function of training condition, color set, and test time.

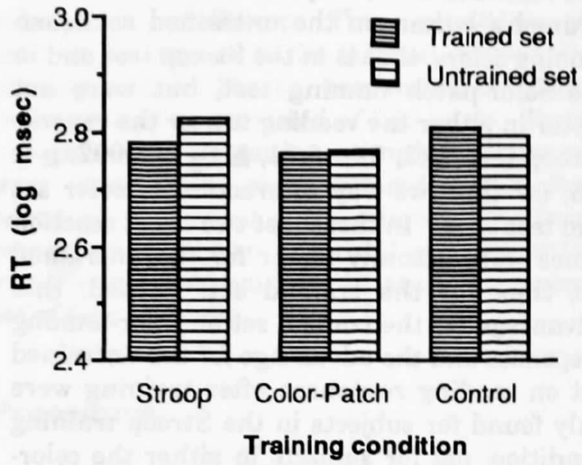

Figure 4. Correct reaction times, after training, across the orthographic Stroop tests (standard, asterisks, uppercase), as a function of training condition and color set.

the effects of training were not specific to the word form. In contrast, specificity of color set was again evident by two observations: First, as shown in Figure 3, the greatest decrease in reaction time from the pretest to the posttest occurred for the Stroop training condition with the trained color set, $F(2, 3) = 11.50$, $MS_e = .0001$, $p < .05$. Second, after training, only the Stroop training condition yielded faster reaction times for the trained set than for the untrained set, $F(2, 3) = 12.56$, $MS_e = .0005$, $p < .05$, as illustrated in Figure 4. Because the same pattern was found for all three orthographic test types, the results are consistent with the hypothesis that training is specific to the colors employed but not to the orthographic form of the color words.

Discussion

We have found clear evidence for specificity effects in the Stroop task, and this specificity persisted across a one-month delay interval. Subjects showed differential effects of color set, depending on whether or not it was the set on which they had trained. Furthermore, after training, the Stroop-trained subjects showed a pattern of advantage for the trained set on the tests that required color naming but a disadvantage for the trained set on the tests that required reading. The question of whether this specificity was due to the colors or due to the words is illuminated by the findings on the additional orthographic manipulation tests. In these orthographic tests the colors were identical to the colors used during training, but the word forms were different. Because these tests yielded the same trained color set advantage as did the standard Stroop test, it can be concluded that the specificity was not due solely to word form. It may, therefore, be that the specificity is due to the colors. An alternative possibility is that there are specific practice effects for the semantic meanings of the to-be-ignored words, as suggested by Reisberg et al. (1980). A third possibility is that practice is actually specific to the individual color-word combinations. That is, practice may improve performance at the level of individual combinations rather than at the level of words or colors. This possibility is currently being examined using verbal protocols collected during Stroop practice. Whatever the cause, the present findings have the clear implication that skill learning, even for relatively simple tasks like the Stroop, cannot be assumed to generalize over variations in all task dimensions.

Acknowledgments

We are grateful to Brooke E. Schefrin for conducting the color vision tests, to William R. Marmie for help designing the experiment, and to Robert J. Crutcher for assistance designing verbal protocol instructions.

References

Anderson, J.R. 1992. Automaticity and the ACT* theory. *American Journal of Psychology*, 105:165-180.

Connor, A., Franzen, M., & Sharp, B. 1988. Effects of practice and differential instructions on Stroop performance. *The International Journal of Clinical Neuropsychology*, 10:1-4.

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

Healy, A.F., Fendrich, D.W., Crutcher, R.J., Wittman, W.T., Gesi, A.T., Ericsson, K.A., & Bourne, L.E., Jr. 1992. The long-term retention of skills. In A.F. Healy, S.M. Kosslyn, & R.M. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 87-118). Hillsdale, NJ: Erlbaum.

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

MacLeod, C.M. 1991. Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109:163-203.

Ménard-Buteau, C., & Cavanagh, P. 1984. Localization of the form/colour interference at the perceptual level in a Stroop task with stimuli drawings. *Canadian Journal of Psychology*, 38:421-439.

Reisberg, D., Baron, J., & Kemler, D.G. 1980. Overcoming Stroop interference: The effects of practice on distractor potency. *Journal of Experimental Psychology: Human Perception and Performance*, 6:140-150.

Schneider, W. 1988. Micro Experimental Laboratory: An integrated system for IBM PC compatibles. *Behavior Research Methods, Instruments, & Computers*, 20:206-217.

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