Journal of Reading Behavior
1979, Vol. XI, No. 3

THE DEVELOPMENT OF AUTOMATIC WORD RECOGNITION SKILLS[1]

Richard F. West
James Madison University

Keith E. Stanovich
Oakland University

Abstract. Kindergarteners, first graders, and third graders performed a discrete-trial Stroop task in which they named the colors of stimuli that either matched or did not match items that were being concurrently held in memory. Letters, high-frequency words, and low-frequency words were used as stimuli. There was a developmental trend toward the color being named faster when the stimulus matched the item held in memory. This finding was unexpected. While the color-naming times of the first and third graders did not depend on stimulus type, the kindergarteners named colors slower when the stimuli were letters and showed a tendency to respond slower to high-frequency words than to low-frequency words. Apparently, the kindergarteners had fully automated the recognition of only the letters and were beginning to automate the recognition of high-frequency words. In contrast, the older children had automated the recognition of letters, high-frequency words, and low-frequency words to an equal extent.

Numerous factors are undoubtedly involved in the acquisition of proficiency in reading. Reading researchers have often argued that one of the most important of these factors is the ability to use prior contextual information to facilitate word recognition (e.g., Goodman, 1967; Smith, 1971). However, several recent studies have suggested that such factors may play a less important role than single-word processing skills in accounting for differences in reading ability (Biemiller, 1977-1978; Perfetti, Finger, & Hogaboam, 1978; Perfetti & Hogaboam, 1975; Shankweiler & Liberman, 1972; West & Stanovich, 1978). For example, contextual factors have been found to have more of an influence on the word reading latencies of less skilled than more skilled readers (Perfetti, Goldman, & Hogaboam, 1979; Roth, Perfetti, & Lesgold, Note 1; Schvaneveldt, Ackerman, & Semlear, 1977; West & Stanovich, 1978).

While word reading skills are frequently assessed in terms of accuracy criteria, consideration of the speed and automaticity of these skills may prove informative. A process is considered to be automatic when it can take place without attention being directed to it (LaBerge & Samuels, 1974). The Stroop effect (Stroop, 1935) has recently been viewed as an index of automatic word reading and may provide a means for studying automatic processes (Posner & Snyder, 1975). In a typical Stroop task subjects are asked to name the color of the ink in which a string of letters is printed. When the letter string forms a non-word, color-naming latencies are relatively short. However, when the string of letters spells the name of a conflicting word (e.g., the word "red" written in blue ink), the printed word interferes with, or slows down, the color-naming response. Color-word interference is normally explained in terms of the competition between vocal responses to the printed word and the ink color. Since the subjects engaged in a Stroop task are attempting to attend only to the color of the ink, color-word interference is presumably the result of the word having been read automatically (Posner & Snyder, 1975).

If skilled readers process words more automatically than less skilled readers, a larger Stroop effect might be expected for the more skilled readers. Evidence somewhat supportive of this prediction for adults has been reported by Martin (1978). However, studies using Stroop-type paradigms with children have generally failed to find the expected relationship between reading skill and the Stroop effect. For example, Schiller (1966) found that color-word interference was maximal in the second grade and decreased with further development. Using a picture-word interference task that is a variant of the Stroop paradigm, Golinkoff and Rosinski (1976) found that the picture naming times of unskilled third- and fifth-grade readers were delayed by the presence of incongruent words just as long as the times of skilled readers of the same age. In another study, Rosinski, Golinkoff, and Kukish (1975) found no increase in the interference effect from second grade through adulthood.

There are probably three reasons why the developmental research has failed to find the expected relationship between reading ability and automatic processing. The first two reasons are interrelated, and concern the reading level of the subjects used in the experiments and the type of words that have been employed. The words used in the Golinkoff and Rosinski (1976) and Rosinski, et al. (1975) studies were simple and highly familiar. The use of highly familiar words, coupled with the fact that even the second graders and unskilled third-grade subjects in these studies were well beyond the initial stages of reading acquisition, may have resulted in a sort of "ceiling effect" whereby the levels of the words and subjects were not in a range where the reading ability-automaticity relationship could be uncovered. This conjecture is supported by research suggesting that the expected relation begins to become apparent when either more difficult words are employed (Pace & Golinkoff, 1976) or poorer readers are tested (Ehri, 1976).

One final reason for the failure to find a relationship between reading ability and automaticity in the studies reviewed above may involve the use of the continuous-list procedure, whereby the subject names a series of items and his score is the total time to name the entire list. The continuous-list procedure involves complex articulatory and sequential-response processes (Proctor, 1978), in addition to making possible nonperceptual strategies for overcoming interference that might change with age (Posnansky & Rayner, 1977). The discrete-trial procedure, where a reaction time to a single stimulus is recorded, is a much cleaner measure and is therefore preferred over the continuous-list procedure.

The study reported here attempts to remedy some of the problems discussed above in the following ways. An age range is investigated (kindergarten, first grade, and third grade) in which large changes in automatic processing of written material would be expected. A color-naming task is used in which interfering stimuli are letters, high-frequency words, and low-frequency words (i.e., stimuli that differ in frequency of exposure, and presumably automatization). Thus, changes in the automatic processing of these different types of written items can be observed across the age range under consideration. A ceiling effect would not be expected with these stimuli, since the low-frequency words should be difficult for the first-grade and kindergarten subjects. Finally, a discrete-trial presentation procedure is used that is more precise and is less amenable to response strategies than is the continuous-list procedure.

One additional manipulation of interest was included in the study to be reported. Warren (1972, 1974) found that interference increased in a discrete-trial Stroop task when the stimulus word matched a word that was concurrently being held in memory for later recall. This effect presumably occurs because the information from both the auditorily and visually presented words is combined at a memory location (see Morton, 1969) and strengthens the word reading response that competes with the color to be named. In the present study, the hypothesis that word recognition becomes increasingly automatic with skill development is tested by the use of a procedure adapted from Warren (1972, 1974). Subjects from three grade levels were asked to name the colors of verbal items that either matched or did not match an item that was being concurrently held in memory. Stimuli of three difficulty levels were used. Since it was expected that the reading of familiar material is automated at an earlier age than the reading of unfamiliar material, it was predicted that the older readers would name the colors of the three stimulus types equally fast, while the younger readers would take longer to name the colors of letters and highly frequent words, because only these stimuli are automated at the younger age levels. Based on Warren's (1972, 1974) findings, it was expected that these results would be most apparent when the colored target and word held in memory matched.

METHOD

Subjects

The subjects were 24 kindergarteners (10 males and 14 females), 24 first graders (15 males and 9 females), and 24 third graders (9 males and 15 females). The children, who were recruited from a predominately middle-class elementary school, were tested near the end of the school year. The kindergarteners had a mean age of 5.9 years, and their mean reading ability was at the 1.1 grade level as tested by the Reading Subtest Level I of the Wide Range Achievement Test or WRAT (Jastak, Bijou, & Jastak, 1965). The first graders had a mean age of 7.2 years, and their mean reading ability was at the 2.2 grade level as tested by the WRAT. The third graders had a mean age of 9.2 years, and their mean reading ability was at the 4.8 grade level as tested by the WRAT.

Stimuli and Apparatus

The stimuli consisted of the 26 letters of the alphabet, 26 high-frequency words, and 26 low-frequency words. The selection of the words was based on an inspection of the Dolch list, the Kucera and Francis (1967) corpus, and a number of reading primers. The low-frequency words had a mean frequency of 16.0 based on the Kucera and Francis (1967) count and 24.0 based on the Carroll, Davis, and Richman (1971) count of third-grade literature. The high-frequency words had a mean frequency of 326.5 based on Kucera and Francis and 509.2 based on Carroll, Davis, and Richman (1971). The stimuli were typed in IBM Courier 72 font. Only the lowercase was used. Black and white negatives of the stimuli were mounted on slides that were colored either red, yellow, blue, or green with acetate films. The slides were projected onto a white screen by a Kodak Carousel 760H projector. Subjects sat approximately 90 cm from the screen, and the size of the projection was such that a five-letter word subtended a horizontal visual angle of approximately 3 degrees. Stimulus onset was controlled by a Vincent Associates Uniblitz Shutter that was positioned over the lens of the projector. When the experimenter pushed a control button, the shutter was electronically opened, and the projected image of the stimulus item appeared. Simultaneously, a National Electronic Systems Crystal Stopwatch was started by the same push of the control button. When the subject verbally responded, a voice activated relay stopped the stopwatch and closed the shutter. The microphone that led to the voice activated relay was held by the subject.

Procedure

Subjects were individually tested in a session that lasted approximately 30 minutes. Preceding each visual presentation of a stimulus the experimenter orally presented the subject with a word or a letter. While they were instructed to try to remember all of the orally presented items, no recall was actually required of the subjects. Approximately .5 seconds after the experimenter presented the auditory item, the stimulus was exposed, and the subject named the color of the stimulus as rapidly as possible.

After six practice trials the subjects completed two blocks of 72 experimental trials. Each block contained a random ordering of 24 trials of each of the three stimulus types (letters, high-frequency words, low-frequency words). Each subject saw the same stimulus once in each of the blocks. On half of the trials within a block the visual stimulus matched the auditory item that had preceded it at the beginning of the trial. When a stimulus was matched in one block it was not matched by the preceding auditory item in the second block, and vice versa. The order in which the blocks were performed was counterbalanced across subjects.

RESULTS

Approximately 5% of the trials were dropped from the data analysis due to some type of experimental malfunction. Trials on which the subject articulated the wrong word or had a response time longer than 2000 msec were scored as subject errors and dropped from the analysis. Errors were made on less than 4% of the trials. There were no sex differences on any of the performance measures, and the data from males and females were pooled in the analysis that follows. Grade level was a between subject factor. Stimulus type and match condition were within subject factors. The ordering of blocks was included as a factor in all of the analyses so as to decrease the error term. Since neither the main effect of order nor any of the interactions involving this factor were significant, it will not be considered further.

The mean color-naming times for each of the grade levels, stimulus type, and match conditions are presented in Table 1. A four-way analysis of variance (grade × stimulus type × match × order) on the color-naming times indicated a significant effect of grade level ($F(2,60) = 13.53$, $p < .001$), stimulus type ($F(2,120) = 4.28$, $p < .025$), and match condition ($F(1,60) = 6.35$, $p < .025$). The grade level by stimulus type interaction ($F(4,120) = 9.73$) was significant at the .001 level. There was a tendency for matches to become faster than nonmatches as the children got older. However, the grade level by match condition interaction did not quite reach accepted levels of statistical significance ($F(2,60) = 2.45$, $p < .10$). None of the remaining interactions was significant. Virtually identical results were obtained when an analysis of covariance was carried out on the reaction times with the number of errors as a covariate. A planned orthogonal contrast indicated that the color-naming latencies of the first graders were significantly slower than those of the kindergarteners and the third graders ($p < .05$). The latencies for the kindergarteners and the third graders did not differ significantly.

TABLE 1

Mean Color-Naming Times as a Function of Grade Level, Stimulus Type, and Match Condition

Grade level	Stimulus type	Match condition	
		Match	Nonmatch
	Letters	1213	1158
Kindergarten	High-frequency words	1002	1031
	Low-frequency words	955	956
	Letters	1190	1306
First grade	High-frequency words	1232	1264
	Low-frequency words	1308	1353
	Letters	961	1070
Third grade	High-frequency words	934	1063
	Low-frequency words	989	1076

In order to further explore the grade level by stimulus type interaction, separate analyses were run on the naming times of each of the age groups of children. An analysis of variance on the color-naming times of the kindergarten children indicated a significant effect of stimulus type ($F(2,40) = 41.55$, $p < .001$). A planned orthogonal contrast indicated that colors were named more slowly for the letter stimuli than for the two types of word stimuli ($p < .001$). Mean reaction times were longer for the high-frequency words than for the low-frequency words, but not significantly so ($p < .15$). Neither the main effect of match condition ($F < 1$) nor the stimulus type by match condition interaction ($F(2,40) = 1.22$) was significant.

An analysis of variance on the color-naming times of the first graders indicated no significant effect of stimulus type ($F(2,40) = 1.38$), match condition ($F(1,20) = 1.20$), nor stimulus type by match condition interaction ($F < 1$). Turning to the color-naming times of the third graders, an analysis of variance indicated no significant effect of stimulus type ($F(2,40) = 2.27$). Matches, collapsed across the stimulus type factor, were a significant 108 msec faster than the nonmatches ($F(1,20) = 31.98$, $p < .001$). The stimulus type by match condition interaction was not significant ($F(2,40) = 1.17$).

DISCUSSION

While the color-naming times of the kindergarteners and third graders did not differ from each other, their response times were significantly faster than those of the first graders. This developmental trend is of particular interest, because normal reaction times increase markedly over the age range studied here (Doehring, 1976; Wickens, 1974). However, this finding should not be considered anomalous since all of the trials in this study were, in a sense, Stroop-type interference trials. Some degree of color-word interference may be assumed on each trial. In this light, the inverted U-shaped response-time function reported here is consistent with results reported by Schiller (1966), who found that time to name the colors of words was the longest in the second grade. Thus, the equality of the kindergarten and third grade reaction times is probably due to the fact that the times of the third graders have been differentially lengthened due to the relatively greater interference experienced by that group.

Stimulus type was found to have a significant effect on color-naming latencies. In addition, there was a significant grade level by stimulus type interaction. The kindergarteners' color-naming latencies were significantly slower when the target was a letter rather than a word. For this group, there was also a trend towards slower response times for the high-frequency words as compared to the low-frequency words (however, this trend did not quite reach statistical significance). In contrast, stimulus type did not appear to have an influence on the color-naming latencies of the first and third graders. Thus, the prediction concerning the interaction between stimulus type and reading ability was supported. Apparently the kindergarteners had fully automated only the recognition of the letters and had just begun to automate the recognition of highly frequent words. In contrast, the older children had automated the recognition of all the stimuli, including the low-frequency words. Thus, the present study has demonstrated a developmental trend toward the automatization of more difficult verbal stimuli; however, it appears that developmental changes in this process occur very early in reading practice.

Colors were named significantly faster when the target matched the preceding auditory word. A tendency for matches to become faster than nonmatches with development was indicated by the nonsignificant match effect for the kindergarteners and first graders, but significant match effect for the third graders. The finding that matches resulted in faster color-naming times was unexpected and contrasts with Warren's (1972, 1974) finding that color-naming times were slower

when the target either matched or was associatively related to a word that was being concurrently held in memory. Perhaps the most important procedural difference between Warren's (1972, 1974) work and the present study is that, unlike the college students in Warren's studies, the young children in the present study were never asked to recall the auditory words.

A possible explanation of the match condition findings in the present study is the "getting out of the way" hypothesis (Dyer, 1971, 1973; Klein, 1964), proposed to explain the finding that subjects experienced reduced color-word interference when they read the colored word aloud prior to naming its color (Klein, 1964). Subjects were able to perform this double-response task almost as rapidly as they could perform the color-naming task alone, and they reported that the double-response task seemed to be much easier. Apparently, vocalizing a word prior to naming its color diminishes the interfering influence of the word. The competing vocal response is "gotten out of the way." Extensive research indicates that with development, children are more likely to adopt a strategy of spontaneously rehearsing words which they want to remember (Hagen & Stanovich, 1977). Perhaps the third graders in this study engaged in such rehearsal, and this rehearsal reduced interference somewhat. Since the present study contained no condition where an auditory stimulus was not presented it is impossible to tell whether the match condition merely reduced interference or eliminated it entirely. The latter outcome, however, seems highly unlikely since if it were the case, the match-condition reaction-times of the third graders would have been much faster than the reaction times of the kindergarten children.

In summary, the present study has provided some support for predictions made by the automaticity theory of reading put forth by LaBerge and Samuels (1974). It is indeed surprising, in light of the influence and popularity of this theory, that there have been so few attempts to test its developmental predictions directly. The research reported here (along with the work of Ehri, 1976; Guttentag & Haith, 1978; Pace & Golinkoff, 1976), although supportive of the automaticity theory, should be viewed as only a tentative first step toward a more comprehensive research effort aimed at testing the theory.

Finally, some limitations of the present work should be mentioned. Several of the experimental effects were statistically marginal, suggesting that the large variability in the reaction times of the younger children may have been contributing to a lack of power. In particular, the lack of a stimulus-type effect in the reaction times of the first graders may have been due to experimental insensitivity. Also, as previously mentioned, the fact that an auditory stimulus was presented on every trial rendered the anomalous results of the match variable difficult to interpret.

FOOTNOTES

[1] The authors wish to thank Becky Tift for her assistance in collecting the data reported in this study. Requests for reprints should be sent to Richard F. West, Department of Psychology, James Madison University, Harrisonburg, VA 22807.

REFERENCE NOTE

1. ROTH, S. F., PERFETTI, C. A., & LESGOLD, A. M. *Reading ability and children's word identification processes.* Paper presented at the Midwestern Psychological Association, Chicago, May 1979.

REFERENCES

BIEMILLER, A. Relationships between oral reading rates for letters, words, and simple text in the development of reading achievement. *Reading Research Quarterly*, 1977-1978, *13*, 223-253.

CARROLL, J. B., DAVIS, P., & RICHMAN, B. *Word frequency book.* New York: American Heritage, 1971.

DOEHRING, D. G. Acquisition of rapid reading responses. *Monographs of the Society for Research in Child Development*, 1976, *41*(2, serial No. 165).

DYER, F. N. The duration of word meaning responses: Stroop interference for different preexposures of the word. *Psychonomic Science*, 1971, *25*, 229-231.

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

EHRI, L. C. Do words really interfere in naming pictures? *Child Development*, 1976, *47*, 502-505.

GOLINKOFF, R. M., & ROSINSKI, R. R. Decoding, semantic processing, and reading comprehension skill. *Child Development*, 1976, *47*, 252-258.

GOODMAN, K. S. Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, 1967, *6*, 126-135.

GUTTENTAG, R. E., & HAITH, M. M. Automatic processing as a function of age and reading ability. *Child Development*, 1978, *49*, 707-716.

HAGEN, J. W., & STANOVICH, K. E. Memory: Strategies of acquisition. In R. V. Kail and J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1977.

JASTAK, J. F., BIJOU, S. W., & JASTAK, S. R. *Wide range achievement test.* Wilmington, DE: Guidance Associates, 1965.

KLEIN, G. S. Semantic power measured through the interference of words with color-naming. *American Journal of Psychology*, 1964, *77*, 576-588.

KUCERA H., & FRANCIS, W. N. *Computer analysis of present-day American English.* Providence: Brown University Press, 1967.

LABERGE, D., & SAMUELS, S. J. Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 1974, *6*, 293-323.

MARTIN, M. Speech recoding in silent reading. *Memory & Cognition*, 1978, *6*, 108-114.

MORTON, J. Interaction of information in word recognition. *Psychological Review*, 1969, *76*, 165-178.

PACE, A. J., & GOLINKOFF, R. M. Relationship between word difficulty and access of single-word meaning by skilled and less skilled readers. *Journal of Educational Psychology*, 1976, *68*, 760-767.

PERFETTI, C. A., FINGER, E., & HOGABOAM, T. Sources of vocalization latency differences between skilled and less skilled young readers. *Journal of Educational Psychology*, 1978, *70*, 730-739.

PERFETTI, C. A., GOLDMAN, S. R., & HOGABOAM, T. W. Reading skill and the identification of words in discourse context. *Memory & Cognition*, 1979, in press.

PERFETTI, C. A., & HOGABOAM, T. Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology*, 1975, *67*, 461-469.

POSNANSKY, C. J., & RAYNER, K. Visual-feature and response components in a picture-word interference task with beginning and skilled readers. *Journal of Experimental Child Psychology*, 1977, *24*, 440-460.

POSNER, M. I., & SNYDER, C. R. R. Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1975.

PROCTOR, R. W. Sources of color-word interference in the Stroop color-naming task. *Perception & Psychophysics*, 1978, *23*, 413-419.

ROSINSKI, R. R., GOLINKOFF, R. M., & KUKISH, K. S. Automatic semantic processing in a picture-word interference task. *Child Development*, 1975, *46*, 247-253.

SCHILLER, P. H. Developmental study of color-word interference. *Journal of Experimental Psychology*, 1966, *72*, 105-108.

SCHVANEVELDT, R. W., ACKERMAN, B. P., & SEMLEAR, T. The effects of semantic context on children's word recognition. *Child Development*, 1977, *48*, 612-616.

SHANKWEILER, D., & LIBERMAN, I. Y. Misreading: A search for causes. In J. K. Kavanagh and I. G. Mattingly (Eds.), *Language by ear and eye.* Cambridge, MA: MIT Press, 1972.

SMITH, F. *Understanding reading.* New York: Holt, Rinehart & Winston, 1971.

STROOP, J. R. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 1935, *18*, 643-661.

WARREN, R. E. Stimulus encoding and memory. *Journal of Experimental Psychology*, 1972, *94*, 90-100.

WARREN, R. E. Association, directionality, and stimulus encoding. *Journal of Experimental Psychology*, 1974, *102*, 151-158.

WEST, R. F., & STANOVICH, K. E. Automatic contextual facilitation in readers of three ages. *Child Development*, 1978, *49*, 717-727.

WICKENS, C. D. Temporal limits on human information processing: A developmental study. *Psychological Bulletin*, 1974, *81*, 739-755.