Perception & Psychophysics
1983, 33 (3), 266-270

Stimulus-response compatibility affects auditory Stroop interference

LUCINDA McCLAIN
Marquette University, Milwaukee, Wisconsin

The contribution of stimulus-response compatibility to Stroop interference was tested in an auditory version of the Stroop test. The words "high" and "low" were presented in high and low pitches with either the pitch or the word designated as the relevant dimension. College students categorized the relevant stimulus dimension with a verbal response, a buttonpress, or a pitched hum. Significant interference occurred in the incompatible conditions (pitch-verbal, word-hum, pitch-button), but not in the compatible conditions (pitch-hum, word-verbal, word-button). The results indicated that stimulus-response compatibility was an important determinant of observed interference. It was suggested that stimulus-response incompatibility contributed to the response competition that presumably occurs in Stroop tasks.

In the Stroop (1935) task, subjects are asked to name as rapidly as possible the color in which various stimuli are printed. The color of congruent color words (e.g., "red" printed in red ink) and neutral words (e.g., "dog" printed in red ink) can be named quickly, but naming time is dramatically increased when incongruent color words (e.g., "blue" printed in red ink) are used. Increased response times to incongruent stimuli are produced by some sort of "interference," but in the 50 years since this phenomenon was first reported, a completely adequate theoretical explanation of it has not been formulated.

The favored explanation attributes interference to response competition (e.g., Dyer, 1973a; Keele, 1972; Morton, 1969). Stroop stimuli have two dimensions: color and word. The color is relevant to the task at hand; the word is irrelevant. Although the two dimensions can be processed in parallel, Morton's competition model (Morton, 1969; Morton & Chambers, 1973) maintains that the word tends to be more quickly analyzed and, consequently, most often enters the single response channel before the color does; the two responses are then assumed to compete for execution. It has been demonstrated that subjects can read words faster than they can name colors (Beller, 1975; Fraisse, 1969; Murray, Mastronardi, & Duncan, 1972; Stroop, 1935). Therefore, the Stroop task requires subjects to actively inhibit the first available response (word name) and to respond with the color name. The competition model predicts that incongruent stimuli will produce interference whenever the response to the task-irrelevant stimulus dimension becomes available more rapidly than the response to the task-relevant dimension. This model also predicts that interference should be minimized when the response to the task-relevant dimension becomes available first. It has been shown (Stroop, 1935) that incongruent stimuli produce strong interference when color is the relevant dimension, but minimal interference when the word is the relevant dimension.

Stimulus-response incompatibility could contribute to response competition. Stimulus-response incompatibility occurs in the Stroop task because one type of stimulus (color) must be responded to with a different type of response (word). According to Treisman and Fearnley (1969), Stroop interference results from the lack of identity (incompatibility) between stimulus type and response type. Their incompatibility model predicts that interference will be greatest when subjects must respond to the task-relevant dimension with a response that not only is incompatible with that dimension, but also is perfectly compatible with the task-irrelevant dimension. Although Treisman and Fearnley did not describe the mechanism by which incompatibility produces interference, stimulus-response incompatibility could exert its influence by retarding response availability; that is, incompatible responses would tend to enter the single response channel after compatible responses. The incompatibility variant of the competition model predicts that incongruent stimuli will produce interference when the relevant dimension requires an incompatible response, but that incongruent stimuli will not produce interference when a compatible response is required. Treisman and Fearnley speculated that Stroop interference would disappear if subjects were able to produce color re-

I thank J. Y. Shih Huang, J. Goerss, G. Wallinga, and C. Degree for assistance in data collection and the editors for thoughtful suggestions which improved the paper. Requests for reprints should be sent to Lucinda McClain, Department of Psychology, Marquette University, Milwaukee, Wisconsin 53233.

sponses to incongruent stimuli rather than the usual verbal (word) responses.

Two factors should theoretically affect the response competition that is presumed to occur in the Stroop task: (1) the average speed with which responses to the relevant and irrelevant dimensions become available, and (2) the stimulus-response compatibility of the relevant dimension and required response. Although both relative identification speed (Dyer, 1971; Gumenik & Glass, 1970; Harrison & Boese, 1976; Hintzman, Carre, Eskridge, Owens, Shaff, & Sparks, 1972; Keele, 1972; Morton & Chambers, 1973; Nealis, 1973; Palef & Olson, 1975) and stimulus-response compatibility (Beller, 1975; Dyer, 1973b; Egeth, Blecker, & Kamlet, 1969; Pritchatt, 1968; Stirling, 1979; Treisman & Fearnley, 1969) have each received some independent experimental support, furthering our understanding of Stroop interference ultimately requires their direct comparison. Assessing the contribution of stimulus-response compatibility to response competition demanded a task that allowed subjects to respond by *producing* either dimension of a stimulus. An auditory analog of the Stroop test similar to those previously reported (Cohen & Martin, 1975; Hamers & Lambert, 1972) met this requirement with the words "high" and "low" sung in high and low pitches. Each stimulus had two dimensions (word and *pitch*), and these dimensions were congruent (high*hi* and low*lo*) or incongruent (high*lo* and low*hi*). The task-relevant dimension was either the word or the pitch, and the response to it was verbal, a buttonpress, or a pitched hum resulting in six stimulus-response conditions (word-verbal, word-button, word-hum, pitch-verbal, pitch-button, pitch-hum). Button responses were included primarily because this response modality had been used in previously reported auditory Stroop studies (Cohen & Martin, 1975; Hamers & Lambert, 1972).

Operationally, interference is present when response times to incongruent or congruent stimuli are longer than response times to control stimuli. If relative identification speed is the primary determinant of response competition, interference should be greater in the pitch-relevant conditions (pitch-verbal, pitch-button, pitch-hum) than in the word-relevant conditions (word-verbal, word-button, word-hum), and the amount of interference should be minimally influenced by response type. These predictions were based on the assumption that, on the average, words would be identified faster than pitches. If interference is affected by stimulus-response compatibility, greater interference should occur in the incompatible conditions (word-hum and pitch-verbal) than in the compatible conditions (word-verbal and pitch-hum). Stimulus-response compatibility generated no direct prediction about performance in the button conditions, because button responses were neither maximally compatible nor incompatible with the word or pitch dimension of the stimulus.

METHOD

Materials and Apparatus

Congruent and incongruent stimuli sung by a male speaker were spectrographically analyzed. The first formant frequencies in the steady-state portion of the syllable were 285 Hz for high*hi*, 287 Hz for low*hi*, 143 Hz for high*lo*, and 148 Hz for low*lo*. Unidimensional control stimuli for the pitch-relevant conditions were the syllable "ah" sung in a high (286-Hz) and low (150-Hz) pitch. Control stimuli for the word-relevant conditions were the words "high" (203 Hz) and "low" (205 Hz) pronounced in a normal speaking voice. Stimuli were equated for length (400 msec), and the six stimuli (two congruent, two incongruent, and two control) for both the word-relevant and pitch-relevant conditions were rerecorded in each of two random orders to yield stimulus tapes containing 12 presentations of each stimulus at regular 8-sec intervals. In each block of 24 trials, each stimulus was presented four times. Electronic signals coincident with the stimuli activated a clockcounter; a voice-activated relay or buttonpress completed timing of the response.

Subjects and Procedure

Young ($M = 20.40$ years, $SD = 1.27$) undergraduate volunteers were assigned randomly to one of the six stimulus-response conditions. After receiving instructions stressing both speed and accuracy followed by 18 practice trials, the 72 stimuli were presented over earphones. Twelve subjects made more than three errors in a given block of 24 trials, and these subjects were replaced. The final group contained 60 volunteers, 10 in each condition. The number of males and females in each condition was approximately equal. Throughout the button conditions, the subjects rested their index fingers on two buttons labeled "high" and "low." Half the subjects pressed the "high" button with the dominant hand; the remainder pressed the "low" button with the dominant hand. Response times to the nearest millisecond were manually recorded. The experimenter also monitored errors, using a sheet on which correct responses were listed without notation of the stimulus type.

RESULTS AND DISCUSSION

The error rate in all conditions was less than 1%. The humming responses of subjects in the word-hum and pitch-hum conditions were reliably differentiated as being high-pitched or low-pitched by three blind raters. All three raters agreed on 1,523 of the 1,540 responses (72 responses by each of 20 subjects), and the response unanimously agreed on was the correct response for the condition. In the remaining cases, the response agreed on by two raters was the correct response for the condition.

Interference occurs when response times to congruent or incongruent stimuli are longer than response times to control stimuli, and facilitation occurs when response times to these stimuli are shorter than response times to control stimuli. To determine the presence of interference and facilitation, each subject's mean correct response time to control stimuli was subtracted from that subject's mean correct response times to congruent and incongruent stimuli. The use of "corrected" response times has been advocated for the traditional Stroop task (Jensen & Rohwer, 1966; Klein, 1964), but corrected times have not previously been used in auditory analogs.

The corrected response times were analyzed with a relevant dimension × stimulus type × response type × block analysis of variance. A priori F tests were used to make planned comparisons between means, and the .05 significance level was adopted for all tests. The main effect of stimulus type was significant [$F(1,54) = 22.73$, MSe = 39,360 msec, with incongruent stimuli (140 msec) producing a greater increment in response time (interference) than congruent stimuli (40 msec). The main effect of trial block was significant [$F(2,108) = 9.68$, MSe = 17,526]. Interference in the first block (134 msec) was significantly greater than interference in the second (72 msec) and third (65 msec) blocks, which did not differ. Block did not interact with any other variable, and the effect of trial block will not be discussed further.

The main effect of response type [$F(2,54) = 3.45$, MSe = 82,514] and the significant relevant dimension × stimulus type interaction [$F(1,54) = 3.99$, MSe = 39,360] can best be understood by reference to the significant relevant dimension × stimulus type × response type interaction [$F(2,54) = 5.66$, MSe = 39,360]. The means from this interaction are shown in Table 1. One comparison of interest was the relative amount of interference produced by incongruent and congruent stimuli in each condition. Incongruent and congruent stimuli each contained two dimensions (word and pitch) and differed only in the manner in which these dimensions were combined. Comparing response times to incongruent and congruent stimuli after each has been corrected for response time to the appropriate unidimensional control stimulus revealed whether incongruent stimuli produced additional interference attributable to the presence of two conflicting stimulus dimensions. Simply comparing incongruent and control stimuli would confound the number of dimensions present and their conflicting nature. The difference in the amounts of interference produced by incongruent and congruent stimuli was called the congruency effect.

Table 1
Increments (in Milliseconds) in Response Time Produced by Congruent and Incongruent Stimuli Compared With Control Stimuli

Stimulus Type	Response Type: Verbal	Hum	Button
	Relevant Dimension: Word		
Incongruent	139	123	66
Congruent	141	−14	27
	Relevant Dimension: Pitch		
Incongruent	226	192	105
Congruent	−17	153	−49

Note–A negative increment indicates facilitation.

It was also of interest to determine whether response times for incongruent and congruent stimuli were significantly longer (interference) or shorter (facilitation) than times for control stimuli. Response times for control stimuli were 423, 1,042, 300, 759, 795, and 530 msec in the word-verbal, word-hum, word-button, pitch-verbal, pitch-hum, and pitch-button conditions, respectively. Based on the significant relevant dimension × stimulus type × response type interaction [$F(4,108) = 4.19$, MSe = 35,260] obtained in an analysis of variance of uncorrected response times to all three stimulus types (incongruent, congruent, and control), a critical difference of 95 msec was required for statistically significant interference or facilitation.

In the compatible conditions (word-verbal and pitch-hum), both incongruent and congruent stimuli produced significant interference, but the congruency effects in both the word-verbal (−2 msec) and pitch-hum (39 msec) conditions were not significant. A critical difference of 103 msec was required for statistical significance in the within-group comparisons of means shown in Table 1. In the incompatible conditions, a different pattern of results was found. In both the word-hum and pitch-verbal conditions, incongruent stimuli produced significant interference, while congruent stimuli produced small, nonsignificant facilitation. Correcting for facilitation, that is, assuming the interference produced by congruent stimuli was 0 msec, significant congruency effects occurred in the word-hum (123 msec) and pitch-verbal (226 msec) conditions.

In the pitch-button condition, incongruent stimuli produced significant interference and congruent stimuli produced nonsignificant facilitation. The occurrence of a significant congruency effect (105 msec) is similar to the results of previous studies (Cohen & Martin, 1975; Hamers & Lambert, 1972), which directly compared incongruent and congruent stimuli without reference to control stimuli. In the word-button condition, neither incongruent nor congruent stimuli produced significant interference and the congruency effect (39 msec) was not significant.

The present results suggest that both the relative identification speed of the two dimensions and stimulus-response compatibility affect interference, as measured by the magnitude of the congruency effect. The average magnitude of the congruency effect was greater in the pitch-relevant conditions (123 msec) than in the word-relevant conditions

(53 msec), confirming the contribution of relative identification speed. When the effects of response type are considered, the results of the word-verbal, word-hum, pitch-verbal, and pitch-hum conditions indicated that stimulus-response compatibility was an important determinant of the congruency effect. Congruency effects occurred when the relevant stimulus dimension and required response were maximally incompatible. Congruency effects were eliminated when the relevant stimulus dimension and required response were maximally compatible. A similar effect of stimulus-response compatibility in a non-Stroop, double-stimulation task was reported by Greenwald (1970) in tests of his ideomotor theory.

If we consider the nature of the button labels, a possible influence of stimulus-response compatibility also occurred in the button conditions. In the present study, the buttons were marked with *verbal* labels, the words "high" and "low." Verbal labels could be regarded as more compatible with the word-relevant stimulus dimension than with pitch-relevant dimension. If button labels were to influence stimulus-response compatibility, a greater congruency effect would be expected in the relatively incompatible pitch-button condition than in the relatively compatible word-button condition. A larger congruency effect was observed in the pitch-button condition (105 msec) than in the word-button condition (39 msec). An analogous effect of button labels was reported by Simon and Sudalaimuthu (1979). Using the traditional visual Stroop task and buttons labeled with colored lights, they found more interference in a word-button condition than in a color-button condition.

Stimulus-response compatibility definitely influenced the magnitude of the congruency effect when a verbal or hum response was required. It was argued above that the congruency effect was the appropriate measure to assess performance in the various conditions. Note that if attention were simply restricted to the interference produced by incongruent stimuli, no effect of stimulus-response compatibility would be observed, because interference in the word-verbal (139 msec), word-hum (123 msec), pitch-verbal (226 msec), and pitch-hum (192 msec) did not differ significantly. A critical difference of 132 msec was required for statistical significance in between-group comparisons of means from Table 1.

The present study has demonstrated that stimulus-response compatibility is an important determinant of congruency effects in an auditory Stroop task. Stimulus-response compatibility presumably exerted its influence by affecting the speed with which compatible and incompatible responses became available and entered the single response channel, thereby affecting response competition. The effects of response competition are not restricted to Stroop tasks (e.g., B. A. Eriksen & C. W. Eriksen, 1974; C. W. Eriksen & Schultz, 1979; O'Hara, 1980) and have occurred in experiments in which the congruent and incongruent conditions were equated for stimulus-response compatibility (C. W. Eriksen & Schultz, 1979). Although stimulus-response incompatibility is not necessary for response competition to occur, the contribution of stimulus-response compatibility to response competition is not unexpected (Greenwald, 1970; Keele, 1973, chap. 5). It has been suggested (Green & Barber, 1981) that interference operates along broadly similar lines in the auditory and visual Stroop paradigms. Response competition almost certainly affects performance in the traditional Stroop task. The present results suggest that both relative identification speed of the two stimulus dimensions and stimulus-response incompatibility may contribute to response competition in the traditional Stroop task.

REFERENCES

Beller, H. K. Naming, reading, and executing directions. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, **104**, 154-160.

Cohen, G., & Martin, M. Hemisphere differences in an auditory Stroop test. *Perception & Psychophysics*, 1975, **17**, 79-83.

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. (a)

Dyer, F. N. Same and different judgments for word-color pairs with "irrelevant" words or colors: Evidence for word-code comparisons. *Journal of Experimental Psychology*, 1973, **98**, 102-108. (b)

Egeth, H. E., Blecker, D. L., & Kamlet, A. S. Verbal interference in a perceptual comparison task. *Perception & Psychophysics*, 1969, **6**, 355-356.

Eriksen, B. A., & Eriksen, C. W. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 1974, **16**, 143-149.

Eriksen, C. W., & Schultz, D. W. Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics*, 1979, **25**, 249-263.

Fraisse, P. Why is naming longer than reading? *Acta Psychologica*, 1969, **30**, 96-103.

Green, E. J., & Barber, P. J. An auditory Stroop effect with judgments of speaker gender. *Perception & Psychophysics*, 1981, **30**, 459-466.

Greenwald, A. G. A double-stimulation test of ideomotor theory with implications for selective attention. *Journal of Experimental Psychology*, 1970, **84**, 392-398.

Gumenik, W. E., & Glass, R. Effects of reducing the readability of the words in the Stroop Color-Word Test. *Psychonomic Science*, 1970, **20**, 247-248.

Hamers, J. F., & Lambert, W. E. Bilingual interdependencies in auditory perception. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 303-310.

Harrison, N. S., & Boese, E. The locus of semantic interference in the "Stroop" color-naming task. *Perception & Psychophysics*, 1976, **20**, 408-412.

Hintzman, D. L., Carre, F. A., Eskridge, V. L., Owens, A. M., Shaff, S. S., & Sparks, M. E. "Stroop" effect:

Input or output phenomenon? *Journal of Experimental Psychology*, 1972, **95**, 458-459.

Jensen, A. R., & Rohwer, W. D. The Stroop Color-Word Test: A review. *Acta Psychologica*, 1966, **25**, 36-93.

Keele, S. W. Attention demands of memory retrieval. *Journal of Experimental Psychology*, 1972, **93**, 245-248.

Keele, S. W. *Attention and human performance*. Pacific Palisades, Calif: Goodyear, 1973.

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

Morton, J. Categories of interference: Verbal mediation and conflict in card sorting. *British Journal of Psychology*, 1969, **60**, 329-346.

Morton, J., & Chambers, S. M. Selective attention to words and colors. *Quarterly Journal of Experimental Psychology*, 1973, **25**, 387-397.

Murray, D. J., Mastronardi, J., & Duncan, S. Selective attention to "physical" vs "verbal" aspects of colored words. *Psychonomic Science*, 1972, **26**, 305-307.

Nealis, P. The Stroop phenomenon: Some critical tests of the response competition hypothesis. *Perceptual and Motor Skills*, 1973, **37**, 147-153.

O'Hara, W. P. Evidence in support of word unitization. *Perception & Psychophysics*, 1980, **27**, 390-402.

Palef, S. R., & Olson, D. R. Spatial and verbal rivalry in a Stroop-like task. *Canadian Journal of Psychology*, 1975, **29**, 201-209.

Pritchatt, D. An investigation into some of the underlying associative verbal processes of the Stroop colour effect. *Quarterly Journal of Experimental Psychology*, 1968, **20**, 351-359.

Simon, J. R., & Sudalaimuthu, P. Effects of S-R mapping and response modality on performance in a Stroop task. *Journal of Experimental Psychology: Human Perception and Performance*, 1979, **5**, 176-187.

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

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

Treisman, A., & Fearnley, S. The Stroop test: Selective attention to colours and words. *Nature*, 1969, **222**, 437-439.

(Manuscript received June 16, 1982;
revision accepted for publication December 6, 1982.)