EFFECT OF CONFLICTING CUES ON INFORMATION PROCESSING: THE 'STROOP EFFECT' VS. THE 'SIMON EFFECT' *

Richard SIMON and Kevin BERBAUM
The University of Iowa, USA

Accepted February 1989

This study examined the relationship between two sources of interference in human information processing: the Stroop effect and the Simon effect. Forty subjects pressed a left- or right-hand key in response to a Stroop color word located on the left or right side of a screen. For one group, ink color was the relevant cue and, for another group, word meaning was the relevant cue. Independent variables were congruence, i.e., agreement or lack thereof between the ink color and meaning of the Stroop word; spatial correspondence, i.e., agreement or lack thereof between the location of the Stroop word and the location of the key used to make the response; and stimulus duration, i.e., 400 or 100 ms. Each of these variables had a significant effect on RT, and there were no significant interactions. According to Sternberg's additive-factor logic, these findings suggest that the Stroop effect (congruence) and the Simon effect (spatial correspondence) involve separate stages of processing. If one assumes that manipulation of stimulus duration affects the encoding stage, then results also suggest that neither the Stroop effect nor the Simon effect involves the stimulus encoding stage.

This research employed the Sternberg additive-factor method to determine the relationship between two potent sources of interference in human information processing, the Stroop effect and the Simon effect.

There is an extensive literature concerned with information processing in situations where the stimulus provides relevant as well as irrelevant cues. The most widely used stimulus of this sort is the Stroop color word (Dyer 1973; Jensen and Rohwer 1966). In a typical Stroop

This research was supported by funds from the Graduate College of the University of Iowa. The authors gratefully acknowledge the assistance of Todd Bever and Renee Smith.

Requests for reprints should be sent to J.R. Simon, Dept. of Psychology, University of Iowa, Iowa City, IA 52242, USA.

task (Stroop 1935), the stimuli consist of names of colors printed in other colors. Subjects experience interference when they attempt to name the ink colors and ignore the words. Despite literally hundreds of studies, there is still a controversy regarding the source of the Stroop interference effect (Dyer 1973; Glaser and Glaser 1982). The traditional explanation is in terms of response competition (e.g., Dalrymple-Alford and Azkoul 1972; Keele 1972; Klein 1964; Morton 1969; Warren 1972). It asserts that the Stroop effect is due to the lack of correspondence between the irrelevant attribute of the stimulus and the response. An alternative explanation is in terms of encoding. It argues that the Stroop effect is due to the lack of congruence between the relevant and irrelevant cues in the stimulus. Hock and Egeth (1970) propose that the interference occurs during perceptual encoding; i.e., the color word distracts from the identification of the ink color by diverting attention from it. Seymour (1977) argues that the interference occurs during conceptual encoding which is a stage between perceptual encoding and response activation when the color information comes in contact with semantic memory.

Simon and his associates (e.g., Craft and Simon 1970; Simon 1969; Simon 1970; Simon and Rudell 1967) have used a paradigm that is similar to the Stroop paradigm to investigate the effect of conflicting cues on information processing. In a typical task, subjects might press a left- or right-hand key, depending on the color of a stimulus light which appears on the right or left side of a display panel. Results indicate that the location of the light provides an irrelevant directional cue that interferes with processing the relevant symbolic cue which is the color of the light. In other words, reactions are faster on trials in which the location of the stimulus and response correspond than on trials in which they do not correspond. Hedge and Marsh (1975) have used the term 'Simon effect' to refer to this typical finding, and we use that label here as a convenient short-hand designation for the phenomenon. Simon et al. have suggested that their effect reflects a stereotypic tendency to respond toward the source of stimulation and, in a series of studies they have provided evidence that the interference is located in the response selection stage (Acosta and Simon 1976; Mewaldt et al. 1980; Simon 1982; Simon et al. 1975; 1976).

Since the Stroop effect and the Simon effect are so similar, superficially at least, and since there is some evidence that both Stroop and Simon effects might involve the same processing stage, it is of interest to determine the relationship between these two potent interference effects. Sternberg’s additive-factor-method (1969) provides a powerful tool for studying the stages of information processing and defining the processing which is accomplished by certain stages. The method assumes that the interval between the presentation of a stimulus and the subject’s response is filled with a sequence of independent stages or cognitive processes. Each stage receives an input from a previous stage, performs a transformation on that input, and passes it along to the next stage. Total reaction time (RT) is simply the sum of the stage durations. When an experimental manipulation affects RT for a task, it does so by influencing the durations of one or more of the stages. Thus, if two experimental variables affect the same processing stage, both variables will be contributors to the duration of that stage, and their expected effect on mean RT is an interaction. On the other hand, if two experimental variables affect two different stages, they will produce independent (additive) effects on total RT. In other words, the experimental variables should not interact in a statistical sense. The additive-factor method, then, involves conducting factorial experiments to determine which variables produce significant interactions and which do not. The variables which interact are assumed to affect a stage in common, whereas the variables which do not interact are assumed to affect different stages. Specifically, in the present study, a significant interaction between the Stroop effect and the Simon effect would be interpreted as evidence that they affect a common processing stage whereas no interaction would suggest that these two phenomena involve separate stages. Stimulus duration was also manipulated in the present study in an attempt to locate the source of the Stroop and/or Simon effect.

Method

Subjects

The subjects were 40 undergraduates, 20 males and 20 females, enrolled in an introductory psychology course at the University of Iowa. All subjects were native English speakers and reported normal color vision and normal or corrected-to-normal visual acuity. Half of the males and half of the females were assigned at random to an ink-relevant condition and the other half to a word-relevant condition.
The apparatus measured keypress RT to the onset of a visual stimulus which appeared on a 60 cm square glare-resistant glass screen located 82 cm from the subject's eyes and perpendicular to the line of sight. Nine Kodak Ektographic projectors with high-speed shutters attached (Uniblitz Model 225 LOAX5) were located behind the screen and used to present the stimulus materials. Each projector contained a single stimulus slide. Opening the shutter in front of a projector resulted in the stimulus appearing on the screen. The opening and closing times for the shutters were 3 and 4 ms, respectively.

A Digital Equipment Corporation (DEC) PDP 11/34 laboratory computer provided overall control of stimulus presentation and data acquisition. The computer was interfaced with a specially-built shutter controller and was programmed to operate the shutters, thus determining the sequence of stimulus presentation, duration of exposure, and intervals between stimuli. The apparatus, then, formed a precise, highly flexible, multichannel tachistoscope. The procedure eliminated stimulus rise time as a complicating factor since projector bulbs remained on throughout the session. Also, since the stimulus slides did not move during the experiment, perfect stimulus alignment was maintained.

Two response keys, one on the right and one on the left, rested on a table in front of the seated subject. Above each key was a white cardboard label with either the word Red printed in red ink, or the word Green printed in green ink. Subjects operated the keys with their right and left index fingers. The computer recorded RT in ms and also recorded which key had been pressed. Since this information was also printed immediately on a DEC Decwriter III, the experimenter could monitor the subject's performance during the session.

Each trial began with the presentation of a small white fixation point with a luminance of 4.5 cd/m² in the center of the screen at approximately eye level. The stimulus, i.e., the word Red or Green printed in red or green ink was then presented on the right or left side of the screen at a distance of 8.9 cm from the fixation point or an arc of 6.2° at the eye. Henceforth, the subscripts R and G will be used to specify ink color; i.e., Red denotes the word Red in red ink, Redo denotes the word Red in green ink, and so on. The first letter of each word was 4.3 cm high and the remaining letters were 2.85 cm high. Stroke width for all letters was 0.79 cm. The word Red measured 9.5 cm horizontally on the screen, and the word Green measured 15.25 cm. The luminance of the individual stimulus words (10.5 and 8.1 cd/m²) for red and green ink, respectively) was adjusted by means of neutral density filters and apertures in front of the projectors so that the ink colors appeared to be equally intense and salient. The viewing screen was steadily illuminated at 3.6 cd/m² (incandescent light) in order to suppress colored after images.

Procedure and experiment design

Taped instructions informed subjects that the experiment would consist of a number of trials, each trial beginning with a small white dot (ready signal) appearing in the center of the screen. After a short interval the word Red or Green printed in red or green ink would appear to either the right or left of the dot. On some trials the word would appear very briefly and on other trials it would appear for a longer duration. Half of the subjects (ink-relevant condition) were told that their task was to respond to the color of the ink; i.e., if the ink color is red, press the key labeled “Red” as quickly as possible, and if the ink color is green, press the key labeled “Green” as quickly as possible. The other half of the subjects (word-relevant condition) responded to the identical sequence of trials but were told that their task was to respond to the meaning of the words; i.e., if the word is Red, press the key labeled “Red” as quickly as possible and if the word is Green, press the key labeled “Green” as quickly as possible. For half of the males and half of the females in each condition, the right key was labeled ‘Red’ and the left key was labeled ‘Green’. For the other half of each sex group, the key labels were reversed. Subjects placed their right index finger on the right key and their left index finger on the left key and were told to keep their fingers resting lightly on the two keys throughout the experiment.

Each subject performed on the same sequence of 192 test trials in which the 16 possible combinations of the four Stroop words (Red, Redo, Green, and Greeno) appeared three times in a random sequence. Prior to the test trials, there were 16 practice trials and 16 un-scored warm-up trials. Both practice and warm-up trials included one of each treatment combination.

The durations of the various events in the trial sequence were as follows: The ready signal remained on for 500 ms. It was followed by a 500 ms. blank interval after which the Stroop word appeared for either 400 or 100 ms. There was a 2-s interval between the subject’s response and the ready signal for the next trial. If the subject did not respond within 5 s, the program advanced automatically to the next trial. Subjects received no feedback regarding RT or the accuracy of their responses.

The experimental design consisted of three within-subjects factors and one between-subjects factor. The within-subjects factors were congruence (i.e., the agreement or lack thereof between the meaning and ink color of the Stroop word), spatial correspondence (i.e., the agreement or lack thereof between the location of the Stroop word and the location of the key used to make the response), and stimulus duration (i.e., 400 or 100 ms). The between-subjects factor was the relevant dimension of the Stroop word (i.e., ink color or word meaning).

Results

Table 1 shows the mean RTs for the 16 treatment conditions, i.e., congruent (Red and Green) and incongruent (Redo and Greeno) Stroop words displayed in a location (right or left) corresponding or not corresponding with the correct response key, and presented for 400 or 100 ms to the ink-relevant or the word-relevant group. Trials on which errors were made (2.6%) were excluded from the computations. The top third of the table combines the data from the group responding to ink color and the
An overall analysis of variance indicated that there was no significant difference between RT to ink color and RT to word meaning (459 vs. 435 ms), $F(1, 38) = 0.53$. The main effects of congruence, spatial correspondence, and stimulus duration were all statistically significant. Responses to incongruent stimuli were slower than to congruent stimuli (459 vs. 436 ms), $F(1, 38) = 33.01, p < 0.001$. Noncorresponding responses were slower than corresponding responses (453 vs. 441 ms), $F(1, 38) = 8.51, p < 0.01$. Responses to long-duration stimuli were slower than to short-duration stimuli (461 vs. 433 ms), $F(1, 38) = 17.13, p < 0.001$. None of the interactions between congruence, spatial correspondence, and duration was statistically significant which suggested that each factor affected a separate stage of information processing. (The $F$ values $[df 1, 38]$ for the Congruence x Correspondence, Congruence x Duration, Correspondence x Duration, and Congruence x Correspondence x Duration interactions were 0.94, 1.19, 1.07, and 3.24, respectively. Since the latter of these interactions approached significance ($p < 0.10$), separate analyses were performed for the long and short duration conditions, but these, too, failed to show a significant Congruence x Correspondence interaction.)

Separate data analyses were conducted for the group responding to ink color and the group responding to word meaning. For reactions to ink color (see middle third of table 1), responses to incongruent stimuli were slower than to congruent stimuli (474 vs. 444 ms), $F(1, 19) = 17.50, p < 0.001$, and responses to long-duration stimuli were slower than to short-duration stimuli (473 vs. 445 ms), $F(1, 19) = 9.80, p < 0.01$. The difference between noncorresponding and corresponding responses was not significant (465 vs. 453 ms), $F(1, 19) = 2.23, p > 0.15$. There were no significant interactions between any of these three independent variables. (The $F$ values $[df 1, 19]$ for the Congruence x Correspondence, Congruence x Duration, Correspondence x Duration, and Congruence x Correspondence x Duration interactions were 0.80, 0.11, 1.89, and 2.88, respectively.)

For reactions to word meaning (see lower third of table 1), responses to incongruent stimuli were slower than to congruent stimuli (442 vs. 427 ms), $F(1, 19) = 19.37, p < 0.001$; noncorresponding responses were slower than corresponding responses (442 vs. 428 ms), $F(1, 19) = 11.77, p < 0.01$, and responses to long-duration stimuli were slower than to short-duration stimuli (450 vs. 421 ms), $F(1, 19) = 7.65, p < 0.01$. None of the interactions between the three main effects was significant which again suggested that the experimental manipulations affected separate stages of information processing. (The $F$ values $[df 1, 19]$ for the Congruence x Correspondence, Congruence x Duration, Correspondence x Duration, and Congruence x Correspondence x Duration interactions were 0.17, 2.01, 0.19, and 0.56, respectively.)

Discussion

The major purpose of this study was to determine whether the Stroop effect and the Simon effect involved a common stage of
information processing. A second purpose was to attempt to locate the Stroop- and/or Simon-type interference in a particular processing stage. At a superficial level, the interference produced by the incongruity of ink color and word meaning in the Stroop task is similar to that observed in the Simon paradigm when an irrelevant location cue affects the time required to respond to a relevant symbolic cue (e.g., Simon 1970). Our results, however, suggest that the Stroop and Simon effects involve separate stages of processing. This conclusion is based on the absence of an interaction between congruence (Stroop effect) and spatial correspondence (Simon effect). According to Sternberg's additive-factor logic, variables that do not interact are assumed to affect different stages. However, because there are possible alternative explanations for the absence of an interaction (Pachella 1974; Sanders 1980; Taylor 1976), this conclusion must be regarded as tentative.

The stimulus duration manipulation was introduced in an attempt to specify more exactly the source of the Stroop and/or Simon interference effect. If one assumes that stimulus duration affects stimulus encoding (the stage in which a stimulus representation is formed), then an interaction of the Stroop and/or Simon effect with stimulus duration would tend to locate the effect(s) in the encoding stage. We found no interaction of either the Stroop or Simon effect with stimulus duration which suggests that neither of these interference effects involve encoding. Thus, our results agree with previous findings indicating that the Simon effect does not involve the encoding stage (Acosta and Simon 1976; Simon 1982; Simon and Pouraghhabagher 1978). Indeed, there is considerable evidence that its locus is in the response selection stage (Acosta and Simon 1976; Mewaldt et al. 1980; Simon 1982; Simon et al. 1975; 1976; Van der Molen and Keuss 1981). But what about the locus of the Stroop effect if, as the Sternberg logic suggests, it is in neither the encoding nor the response selection stage? Seymour (1977) suggests that Stroop interference occurs during conceptual encoding, a stage between perceptual encoding and response activation when color information contacts semantic memory. Simon and Berbaum (1988) conclude, based on finding a significant Stroop effect in a retrieval task, that the most likely locus of interference is the stage between perceptual encoding and response selection; i.e., the process of decoding or retrieving information from short-term memory.

In the original Stroop study (1935), subjects asked to name the ink color experienced great difficulty from incongruent word meaning, but subjects asked to read the words encountered little interference from incongruent ink colors. It is worth noting that we found not only the 'standard' Stroop effect but also the 'reverse' Stroop effect. Our results, then, support the more recent research which indicates that incongruent ink colors can also interfere with responses to word meaning (Flowers 1975; Simon and Sudalaimuthu 1979; Simon et al. 1985; Uleman and Reeves 1971).

Another interesting aspect of our results was the finding that reactions were significantly faster to the 100 ms stimulus than to the 400 ms stimulus. Welford (1961) has proposed an interesting explanation for this effect. He suggested that subjects inspect a stimulus until 'sufficient' data are gathered to make a response. Reactions are faster to a short duration stimulus because the intake of data is stopped by the cessation of the stimulus, and, therefore, subjects gain little or nothing by delaying their reaction. Studies of the effect of stimulus duration on RT have produced contradictory findings. Some report a tendency for RT to decrease with an increase in duration (e.g., Froeberg 1907). Others, like the present study, report a difference in the opposite direction (e.g., Wells 1913; Botwinick et al. 1958; Gregg and Brogden, 1950). Still other studies have found no significant effects (e.g., Deup-
ree and Simon 1963; Tolin and Simon 1968]. While the contradictory findings seem to be related, at least in part, to the specific durations studied and to the complexity of the task [Welford 1980], the conditions under which stimulus duration affects RT remain unclear.

There is a wide variety of tasks in which irrelevant cues intrude to affect information processing. Several recent studies have focused on the effect of stimulus-response compatibility in these Stroop or 'Stroop-type' tasks [Green and Barber 1983; McClain 1983; Simon and Sudalaimuthu 1979; Zakay and Glicksohn 1985]. Harvey (1984) suggests that use of the term 'Stroop-type' should be avoided since it implies that the interference may have a different etiology from that giving rise to the color-word effect. He suggests that, until contrary evidence is provided, it is more parsimonious to credit Stroop (1935) with discovering a general effect. Results of the present study strongly suggest that it would be a mistake to lump together indiscriminately all interference phenomena under the heading 'Stroop effect' since the locus of the interference; e.g., encoding or response selection may differ for different 'Stroop-type' tasks.

References


Froebre, S., 1907. The relation between the magnitude of stimulus and the time of reaction. Archives of Psychology 8, 1-38.


