Journal of Experimental Psychology:
Human Perception and Performance
1998, Vol. 24, No. 1, 3–19

0096-1523/98/$3.00

The Effects of Stimulus–Response Mapping and Irrelevant Stimulus–Response and Stimulus–Stimulus Overlap in Four-Choice Stroop Tasks With Single-Carrier Stimuli

Huazhong Zhang
Indiana University Kokomo

Sylvan Kornblum
University of Michigan

The purpose of this study was to investigate whether and how stimulus–stimulus (SS) and stimulus–response (SR) consistency and SR congruence effects combine to produce the Stroop effect. Two experiments were conducted with 4-choice tasks in which SS and SR consistency and SR congruence effects were examined in isolation as well as in the Stroop task. The experiments were so designed as to remove the confound between SS and SR consistency that is ordinarily found in standard Stroop tasks and to pit SS consistency against the logical recoding hypothesis (A. Hedge & N. W. A. Marsh, 1975). The results indicate that SS and SR consistency both contribute to the Stroop effect and that they interact. This finding supports models such as the dimensional overlap model (e.g., S. Kornblum & J. W. Lee, 1995) that distinguish between SS and SR overlap. Simulation results from an interactive activation network, modeled after the dimensional overlap model, provide reasonable fits to the experimental data.

The Stroop task (Stroop, 1935/1992), together with its many variants (see MacLeod, 1991, for a review), is one of the most widely used experimental paradigms in cognitive psychology. In today's standard version of the Stroop task, subjects are shown a series of color words, written in various color inks, and are instructed to name the color of the ink while ignoring the word itself. The word and the ink color may correspond (e.g., the word "RED" written in red) or not (e.g., the word "RED" written in green). When they correspond, the reaction time (RT) for naming the color is faster than when they do not correspond. This RT difference is usually called the Stroop effect (Dyer, 1973; MacLeod, 1991; Stroop, 1935/1992)[1] and ranges between 100 and 130 ms (Dyer, 1971, 1974; Glaser & Glaser, 1982; Hintzman et al., 1972).

Many theoretical accounts of the Stroop effect fit into one of two general classes of models: those attributing the effect to stimulus conflicts—so-called early-selection models; and those attributing it to response conflicts—late-selection models.[2] An example of an early-selection account was given by Hock and Egeth (1970), who argued that color perception is slowed when the ink color is inconsistent with the color word. Similarly, Seymour (1977) and Simon and Berbaum (1990) maintained that stimulus conflicts may occur during memory retrieval and comparison of the relevant and irrelevant stimuli. In contrast, according to one version of the late-selection account—the relative speed-of-processing view (Dyer, 1973)—the ink color and the color word produce two potential responses that then race against each other. The winner of this race is the response eventually made. The argument is that because word reading is the faster of the two processes, it is more likely to interfere with the slower process (i.e., color naming) than vice versa. From this viewpoint it follows that if relative timing were the crucial factor, then if color naming were given a head start, naming the ink color would interfere with word reading. The data, however, do not support this prediction (e.g., Glaser & Glaser, 1982; MacLeod & Dunbar, 1988). A difference in processing speeds is, therefore, unlikely to be the principal determinant of the Stroop effect. Another late-selection account—the automaticity view (Logan, 1978; Posner &

This article is based on Huazhong Zhang's dissertation submitted to the Department of Psychology at the University of Michigan. This research was funded by National Institute of Mental Health Grant R01 MH43287, by Air Force Office of Scientific Research Grant F49620-94-1-0020, and by a predoctoral fellowship and a Horace H. Rackham Dissertation Fellowship from the University of Michigan. The preparation of this article was assisted by a Summer Research Fellowship from Indiana University Kokomo. We thank the members of the doctoral committee, Keki Irani, David Meyer, and Jun Zhang, for their comments and Anthony Whipple for his help in computer programming. We are grateful to Carol A. Fowler, Colin MacLeod, and Robert Proctor for their suggestions on an earlier draft of this article.

Correspondence concerning this article should be addressed to Huazhong Zhang, Department of Social and Behavioral Sciences, Indiana University Kokomo, 2300 South Washington Street, Kokomo, Indiana 46904-9003 or to Sylvan Kornblum, Mental Health Research Institute, 205 Zina Pitcher Place, University of Michigan, Ann Arbor, Michigan 48109-0720. Electronic mail may be sent via Internet to hzhang@iukfs1.iuk.indiana.edu. or to Kornblum@umich.edu.

[1] The condition in which the color and the word correspond was not included in Stroop (1935/1992). Thus, technically the Stroop effect referred to an interference, the RT difference between a noncorresponding condition and a control condition.

[2] In this article, we divide the information processing sequence into two global stages—stimulus processing and response production. The stimulus processing stage may be composed of perceptual encoding, memory retrieval, conceptual encoding, and stimulus comparison, and the response production stage may be composed of response selection, motor programming, and execution.

Snyder, 1975; Shiffrin & Schneider, 1977; Washburn, 1994)—contends that because reading is automatic and color naming is not, reading may interfere with color naming, but not vice versa. Because the automatic process does not require attention, according to this strong automaticity view, attentional allocation should not influence the Stroop effect, but it does (Kahneman & Chajczyk, 1983). A weaker version that assumes a gradient of automaticity seems more promising (Kahneman & Chajczyk, 1983; MacLeod & Dunbar, 1988).

The Dimensional Overlap Model

These two proposals, early and late selection, focus on one particular aspect of the Stroop task to the exclusion of the other. The early-selection account focuses on the similarity between the relevant stimulus and the irrelevant stimulus, whereas the late-selection account focuses on the similarity between the irrelevant stimulus and the response. Both similarity relationships are, of course, present in the Stroop task—in fact, they constitute a confounding that makes distinguishing empirically between the two accounts difficult. Recently, these and other similarity relationships between relevant and irrelevant stimulus sets and response sets have been analyzed in detail and now form the basis of the dimensional overlap model (Kornblum, 1992; Kornblum, Hasbroucq, & Osman, 1990; Kornblum & Lee, 1995) that attempts to encompass stimulus–stimulus (SS) and stimulus–response (SR) compatibility tasks ranging from the simplest ones, studied by Fitts and his colleagues (e.g., Fitts & Deininger, 1954), to the more complex ones represented by the Stroop and Stroop-like tasks.

Dimensional overlap (DO) is defined as the occurrence of perceptual, conceptual, or structural similarity between stimulus sets, stimulus and response sets, or both. Dimensional overlap may occur between the relevant stimulus and response dimensions (called *relevant SR overlap*), between the irrelevant stimulus and response dimensions (called *irrelevant SR overlap*), or between the relevant and irrelevant stimulus dimensions (called *SS overlap*). Clearly, given any SR ensemble, DO may occur between none, one, two, or all three of these dimensions, thus giving rise to eight different classes of potential SR ensembles. These eight classes make up the taxonomy that we have been using recently to distinguish between various compatibility tasks (see Kornblum & Lee, 1995, for the most recent version of the taxonomy together with illustrative examples).

According to this taxonomy, an SR ensemble in which the relevant and irrelevant stimulus dimensions do not overlap either with each other or with the response dimension is a Type 1 ensemble. When only the relevant stimulus dimension overlaps with the response, and there are no other overlapping dimensions, the SR ensemble is a Type 2 ensemble. When only the irrelevant stimulus dimension overlaps with the response, and there are no other overlapping dimensions, it is a Type 3 ensemble. When only the relevant and irrelevant stimulus dimensions overlap with each other, and there are no other overlapping dimensions, it is a Type 4 ensemble. When the same relevant and irrelevant stimulus dimensions overlap with each other as well as with the response, it is a Type 8 ensemble. (For present purposes, we ignore Types 5, 6, and 7.)

Illustrative examples of these various ensemble types are easily constructed. Imagine an SR ensemble in which the relevant stimulus dimension is colors, the irrelevant stimulus dimension is color words, and the response dimension is color names. This description obviously refers to the Stroop task and is clearly a Type 8 ensemble in the taxonomy. Imagine now an SR ensemble with the same stimulus dimensions, color, and color words, but one in which the response dimension has been changed from color names to digit names. Although the stimuli would still be called "Stroop-like," clearly the task has been altered. According to our taxonomy, this is a Type 4 ensemble. Next, imagine taking this Type 4 ensemble and changing the irrelevant stimulus dimension from color words to digits. The relevant stimulus dimension is still colors, and the response dimension is still digit names, but the irrelevant stimulus dimension is now digits. The irrelevant stimulus dimension no longer overlaps with the relevant stimulus dimension (colors), but it overlaps with the response dimension (digit names). This is a Type 3 ensemble, of which the spatial version, and probably the best known example, is the so-called "Simon task" (Simon, 1990). To construct a Type 2 ensemble, a researcher should reverse the relevant and irrelevant stimulus dimensions of the Type 3 ensemble. That is, the stimuli still consist of colored digits, and the responses still consist of digit names, but instead of colors being relevant and digits being irrelevant, digits are now relevant and colors are now irrelevant. This is a Type 2 ensemble and represents the standard SR compatibility tasks. Finally, an SR ensemble in which colors are the relevant stimulus, shapes (of the color patch) are the irrelevant stimulus, and digit names are the response is a Type 1 ensemble. When properly designed (see Kornblum & Lee, 1995), Type 1 ensembles may serve as neutral baselines that researchers can use to evaluate the effects of dimensional overlap in other ensembles.

Whenever two dimensions in an SR ensemble overlap, the particular instances of these dimensions, as they occur in individual trials, either match or mismatch. Matches and mismatches between the relevant stimulus and the response are called congruent and incongruent, respectively, and matches and mismatches between the irrelevant stimulus and the relevant stimulus or response are called consistent and inconsistent, respectively. Thus, given the appropriate DO conditions, some trials may be SS consistent (SS^+), SS inconsistent (SS^-), SR consistent (SR^+), SR inconsistent (SR^-), or any pairwise combination of SS and SR consistency.

Empirically, the RT for congruent SR mappings is faster than for incongruent mappings (in Type 2). The DO model attributes this difference to two processes in the response production stage: (a) the process of automatic activation of the congruent response by the overlapping stimulus, and (b) the process of identification of the correct response, given the stimulus and the SR mapping. According to the model, the congruent response is automatically activated irrespective of the SR mapping. Response identification is faster for congruent mapping than for incongruent mapping. Furthermore, if the mapping is incongruent, the automatically activated, erroneous, congruent response needs to be aborted before the correct response can be programmed and executed. Thus, compared with the congruent mapping condition, incongruent responses are subject to two sources of delay (for more detail, see Kornblum et al., 1990; Kornblum & Lee, 1995).

As is true of the effects of DO when the relevant stimulus dimension overlaps with the response (e.g., Type 2), when the irrelevant stimulus dimension overlaps with either the response (e.g., Type 3) or the relevant stimulus dimension (e.g., Type 4), RT is generally faster for consistent conditions (SR^+ or SS^+) than for inconsistent conditions (SR^- or SS^-). In the case of irrelevant SR overlap (e.g., Type 3), the DO model attributes this effect to the same automatic response activation process that it postulated for the relevant SR overlap condition. In the case of SS overlap (e.g., Type 4), the DO model attributes this effect mainly to the stimulus processing stage (see Kornblum, 1992, 1994), which is globally construed to encompass encoding, retrieval, and comparison of the relevant and irrelevant stimulus features. Even though some have argued that SS and SR overlap effects are attributable to the same processing stage (e.g., Lu & Proctor, 1995), Kornblum (1994) and Stoffels and van der Molen (1988) have found that SS and SR consistency effects are additive, which suggests that these effects arise from two separate processing stages. Some researchers have reported psychophysiological data (e.g., lateralized readiness potential) that appear to indicate that SS consistency affects the response production stage (Coles, Gratton, Bashore, Eriksen, & Donchin, 1985; Gratton, Coles, Sirevaag, Eriksen, & Donchin, 1988; Lu & Proctor, 1995). Therefore, either the SS consistency effect is not a pure measure of the stimulus processing stage, the psychophysiological measure (e.g., lateralized readiness potential) is not a pure measure of the response production stage, or both. In either case, however, we know from the findings showing additivity of SS and SR effects that any response stage at which SS consistency effects arise is nonetheless distinct from the response stage at which SR consistency effects arise.

Parsing the Stroop Task Into Its Constituent Components

Because the Stroop task is a Type 8 ensemble, the relevant and irrelevant stimulus dimensions overlap not only with each other but also with the response. Furthermore, the overlapping dimension is the same for all three relations, which in the standard Stroop task is color. Using the DO taxonomy as a framework, it is now possible to parse the Stroop task into its theoretical constituent components; these components turn out to be Types 2, 3, and 4. That is, the unique overlap occurring between the relevant and the irrelevant stimulus dimensions (colors and color words) is the identifying property of Type 4; the unique overlap occurring between the irrelevant stimulus dimension and the response (color words and color names) is the identifying property of Type 3; and the unique overlap occurring between the relevant stimulus dimension and the response (colors and color names) is the identifying property of Type 2.

Unconfounding Confounded Effects in the Stroop Task

The standard Stroop task is normally run with congruent SR mapping instructions. That is, given the Type 2 properties of the task, subjects are usually instructed to respond to the color with its color name. This process necessarily locks the irrelevant SR (Type 3 constituent) and SS (Type 4 constituent) to the same consistency value: Either they are both consistent (SR^+/SS^+)—the color and color word correspond—or they are both inconsistent (SR^-/SS^-)—the color and the color word do not correspond (see Table 1). This confounding needs to be addressed directly if one is to determine whether the "Stroop effect" is due to the SS overlap, the irrelevant SR overlap, or possibly, both.

This confounding can be eliminated with incongruent SR mapping instructions where the relevant stimulus and the response never correspond. Therefore, if the irrelevant SS relation is consistent (SS^+), the irrelevant SR relation is necessarily inconsistent (SR^-), as in row C of Table 1, or if

Table 1
Stroop Conditions and Illustrative Examples

Condition	Illustrative stimuli and responses: Relevant stimulus (color)	Response (color name)	Irrelevant stimulus (color word)	SR mapping	Consistency values of the two irrelevant stimulus relations (SR/SS)[a]
A	Blue	Blue	Blue	Congruent (+)	SR^+/SS^+
B	Blue	Blue	Green	Congruent (+)	SR^-/SS^-
C	Blue	Green	Blue	Incongruent (−)	SR^-/SS^+
D	Blue	Green	Green	Incongruent (−)	SR^+/SS^-
E	Blue	Green	Red	Incongruent (−)	SR^-/SS^-

Note. SR = stimulus–response; SS = stimulus–stimulus.
[a]Superscript plus sign indicates consistent, and superscript minus sign indicates inconsistent.

SS relation is inconsistent (SS^-), SR relation is consistent (SR^+), as in row D of Table 1. Simon and Sudalaimuthu (1979) were the first to use incongruent SR mapping instructions with a two-choice Stroop task. They found that irrespective of the SR mapping, RT was faster when the relevant and irrelevant stimuli matched (SS^+; rows A and C in Table 1) than when they did not match (SS^-; rows B and D in Table 1). Green and Barber (1981) and Kornblum (1992) obtained similar results with other two-choice, Stroop-like tasks that used noncolor stimuli and responses. Because the result with incongruent mapping conditions appears to be a reversal of the SR consistency effect, this finding with incongruent mapping conditions would seem to support an SS over an SR account of the Stroop effect.

However, there is another possible explanation that needs to be considered. Note that coinciding with each instance of the SS^+ conditions that produced the faster RTs (rows A and C in Table 1), the values of SR congruence and SR consistency are identical in each of the two rows: congruent/consistent in row A and incongruent/inconsistent in row C. Thus, if subjects had used a rule (identity or reversal) to arrive at the correct response from the relevant stimulus, they might have used the same rule to deal with the irrelevant stimulus and the response and possibly have been faster in rows A and C than in rows B and D on that account. This theory, of course, is an application of Hedge and Marsh's (1975) "logical recoding hypothesis" to the Stroop task. These authors, who also found a "reverse SR consistency effect" with incongruent mapping, albeit with a different task type (Type 5), argued the following:

> For a given logical recoding (identity or reversal) of the relevant attribute . . . , responding was faster for trials in which the recoding of the irrelevant attribute . . . was of the same logical type as that of the relevant attribute, than for trials in which the logical recoding of the irrelevant attribute was opposite in type. (p. 435)

Thus, even though it may have seemed that, in principle, the confounding that we pointed to at the beginning of this article could be resolved if we used incongruent SR mapping instructions, a new confound has emerged, and there is no way to disentangle the SS account from the logical recoding hypothesis on the basis of two-choice tasks, even though we are using congruent and incongruent SR mapping instructions.

This issue is resolved by going to three or more choices. The additional condition furnished by increasing the choices beyond two is illustrated in row E (SR^-/SS^-) of Table 1. If logical recoding accounts for the results, then the RT in rows C (SR^-/SS^+) and E (SR^-/SS^-) should not differ significantly from each other because the same rule—reversal—is applicable in both; if SS consistency accounts for the results, then there should be a significant difference in RT between rows C (SR^-/SS^+) and E (SR^-/SS^-), with the RT in rows E (SR^-/SS^-) and D (SR^+/SS^-) not being significantly different. Because one of the principal objectives of the present study was to address the confounds present in the Stroop task, all the experiments to be reported used four-choice Stroop tasks with both congruent and incongruent mappings.

Are Types 2, 3, and 4 Additive Constituent Components of the Stroop Task?

A second major objective of this study was to examine the extent to which performance on the Stroop task could be accounted for by the effects of its constituent components: SS and SR consistency. This issue can be addressed in at least two different ways: One can either examine the effects of SS and SR consistency in isolation—in Types 3 and 4—or one can calculate these effects from the various conditions in the Stroop task itself (see Table 1). We did both.

We used the same stimulus and response dimensions to construct Types 2, 3, and 4 tasks as we used in the Type 8 Stroop task. To avoid contamination of the data by extraneous factors such as, for example, differences between the speeds of reading and naming, we used single-carrier stimuli (Glaser & Glaser, 1982, 1989) in which both the relevant and the irrelevant aspects of the stimulus were of the same type; for example, they were both words, digits, or color patches. We showed in an earlier study that when Types 3 and 4 are combined into a single task (Type 7), the SS and SR consistency effects in this new task are additive (Kornblum, 1994) and not significantly different from what they are in isolation (see also Simon & Berbaum, 1990; Stoffels & van der Molen, 1988). This additivity is in agreement with the DO model, which postulates that SS and SR consistency effects are generated in different stages. In Type 8, the DO model postulates that both the relevant and the irrelevant SR overlap are processed by the same stage, the response production stage. Therefore, we expected these to interact.

Experiment 1

The stimuli in this experiment consisted of three words presented one above the other. The middle word was the relevant stimulus, to which the subject was instructed to attend; the top and bottom words were always the same and were the irrelevant stimuli that the subjects were instructed to ignore. The responses were vocal. Four ensembles were constructed: Types 2, 3, 4, and 8. Types 2, 3, and 4 were used to examine the main effects of SR mapping, SR consistency, and SS consistency in isolation, respectively. We then used the results from these ensembles to try to calculate performance in Type 8.

Method

Subjects. Eight male University of Michigan undergraduate students were tested individually. They were right-handed, had normal vision and hearing, and had perfect color vision as tested by the concise edition of *Ishihara's Tests for Colour-Blindness* (Ishihara, 1991). They were all native English speakers with normal reading skills as tested by the WRAT R^2 reading test. Each of them participated in six 90-min sessions and was paid \$60 for the completion of the study. In addition, they each received a bonus of approximately \$1 per session, which was contingent on their performance.

Apparatus. An IBM 386 PC controlled the presentation of stimuli and the measurement of responses. Subjects sat at a table in a dimly lit room and viewed a computer screen from a distance of approximately 75 cm. During the experiment, subjects wore headphones. Vocal responses were registered by a software voice key and were recorded by the computer. The experimenter, who was in another room and was also wearing headphones, listened to the subject's responses and keyed them into the computer. The computer monitored the subject's performance for accuracy.

Stimuli and responses. The stimuli were constructed of three English words presented in the center of a computer screen in white, against a black background, one above the other, inside a rectangle of 4.5 cm × 3 cm (subtending a visual angle of 3.43° × 2.29°). The centers of the three words were aligned vertically, and the center of the middle word coincided with the center of the rectangle. The middle word was the relevant stimulus. The top and bottom words were identical and constituted the irrelevant stimulus. The words were prepared with the Aldus PhotoStyler program and were written in bold typeface with Helvetica 32 point font. Each word measured 0.8 cm vertically (0.61°) and, depending on the length of the word, 1.6–4.0 cm horizontally (1.22°–3.05°). Words were separated vertically by an intercontour distance of 0.15 cm (0.11°).

There were four color words (i.e., RED, GREEN, BLUE, and YELLOW) and four digit names (i.e., TWO, FOUR, SIX, and EIGHT) used as relevant and irrelevant stimuli. As shown in Table 2, depending on the ensemble type, the relevant and irrelevant stimuli were either both color words or both digit words (Types 4 and 8), or the relevant stimuli were one type and the irrelevant stimuli were another type (Types 2 and 3). The relevant stimulus as well as the irrelevant stimulus could be any one of those eight words, which generated 64 (8 × 8) unique stimulus triplets.

Upon seeing a stimulus, subjects made a response by saying either a color name (e.g., "RED," "GREEN," "BLUE," or "YELLOW") or a digit name (e.g., "TWO," "FOUR," "SIX," or "EIGHT"). The two response sets, color names and digit names, were paired with different combinations of the stimulus sets, color words and digit words, as relevant or irrelevant stimuli, to form two alternative SR ensembles for each ensemble type (see Table 2). These alternative ensembles will henceforth be designated by their type and response set (e.g., Type 4—digit names). Within each ensemble type, color words and digit words appeared once as the relevant (in one of the alternative ensembles) stimulus sets and once as the irrelevant (in the other alternative ensemble) stimulus sets, and color names and digit names appeared once as the response set. Thus, any differences in the stimuli and responses across ensemble types were eliminated when the results for the two alternative ensembles within a type were combined in a grand average for that type.

Table 2
Stimulus and Response Sets Used in Experiment 1

Ensemble type	Relevant stimulus set	Response set	Irrelevant stimulus set
2	Color words	Color names	Digit words
	Digit words	Digit names	Color words
3	Digit words	Color names	Color words
	Color words	Digit names	Digit words
4	Digit words	Color names	Digit words
	Color words	Digit names	Color words
8	Color words	Color names	Color words
	Digit words	Digit names	Digit words

Note. Color words (names) were RED, GREEN, YELLOW, and BLUE, and digit words (names) were TWO, FOUR, SIX, and EIGHT.

SR mapping. For Types 2 and 8, relevant stimuli were mapped onto individual responses either congruently or incongruently. With congruent mapping, subjects were instructed to say the relevant stimulus word. With incongruent mapping, subjects were instructed to say a word that was different from the relevant stimulus but was in the same category. Each subject performed with one congruent mapping and one incongruent mapping for Types 2 and 8. For the incongruent mapping, given a relevant stimulus (e.g., TWO), subjects could give three possible responses (i.e., "FOUR," "SIX," or "EIGHT"). Because one of these incongruent mappings would have constituted the reversal rule (i.e., TWO-"EIGHT," FOUR-"SIX," SIX-"FOUR," and EIGHT-"TWO"), it was not used in this experiment.

For both Types 3 and 4, the relevant stimuli could be mapped onto the responses in four different ways. If the relevant stimulus was a color word, then the response could be any one of the four digit names; if the relevant stimulus was a digit word, then the response could be any one of the four color names. All four possible mappings were used, but each subject was run with only one.

Experimental conditions. In Type 2, there were two SR mapping conditions: congruent and incongruent. Given a response set (e.g., color names), the irrelevant stimulus words (e.g., digit word) were neutral with respect to both the relevant stimuli (e.g., color words) and the responses (e.g., color names). The only difference between the mappings was the SR congruence relation (congruent vs. incongruent).

In Type 3, there were two consistency conditions: SR^+ and SR^-. In both conditions, given a response set (e.g., color names), the relevant stimulus (e.g., digit word) was neutral with respect to both the irrelevant stimulus words (e.g., color words) and the response (e.g., color name). The only difference between them was the SR consistency relation (SR^+ vs. SR^-).

In Type 4, there were also two consistency conditions: SS^+ and SS^-. In both conditions, the response (e.g., digit name) was neutral with respect to both the relevant stimuli (e.g., color word) and the irrelevant stimuli (e.g., color words). The only difference between them was the SS relationship (SS^+ vs. SS^-).

In Type 8, there were two SR mappings: congruent and incongruent. In the congruent mapping, subjects said the relevant stimulus word; in the incongruent mapping, subjects said a word that was different from the relevant stimulus. Within the congruent mapping, there were two conditions: A and B (see Table 1). In Condition A, the relevant and irrelevant stimuli were consistent with each other (SS^+) as well as congruent or consistent with the response. In Condition B, the relevant stimulus was congruent with the response, but the irrelevant stimulus was inconsistent with both the response and the relevant stimulus (SR^-/SS^-). Within the incongruent mapping, there were three conditions: C, D, and E (see Table 1). In Condition C, the irrelevant stimulus was inconsistent with the response but was consistent with the relevant stimulus (SR^-/SS^+); in Condition D, the irrelevant stimulus was consistent with the response but inconsistent with the relevant stimulus (SR^+/SS^-); and in Condition E, the irrelevant stimulus was inconsistent with both the response and the relevant stimulus (SR^-/SS^-).

Design. We used a within-subjects design that included six sessions per subject. All four ensemble types were run within a session. The SR ensembles in Table 2 were subdivided into two groups. Group 1 included the following: Type 2—color names, Type 3—color names, Type 4—digit names, Type 8—color names; Group 2 included the rest. Half the subjects were run with Group 1 on the first three sessions (sessions 1–3) and with Group 2 on the last three sessions (sessions 4–6). For the other half of the subjects, this order was reversed. There was a break, ranging from one to several days, between the third and fourth sessions.

Each session consisted of 12 blocks. The following conditions were each run for two blocks: Type 2—congruent, Type 2—incongruent, Type 8—congruent, Type 8—incongruent, Type 3, and Type 4. This particular order was used for all subjects during practice (Sessions 1 and 4). Data from the practice sessions were not analyzed. For Session 2, four ensemble types and 8 subjects were counterbalanced with two balanced Latin squares, and for Session 3, the reverse Latin squares were used. For Session 5, four ensembles and 8 subjects were counterbalanced with two different balanced Latin squares, and for Session 6, the reverse Latin squares were used.

Each block contained 48 trials: three repetitions of 16 unique stimuli. The order of presentation was randomized within each block. In Types 2 and 8, all the trials within a block were either mapped congruently or incongruently. In Type 3, the irrelevant stimulus and response were consistent (SR^+) on one quarter of the trials and inconsistent (SR^-) on the remaining three quarters. In Type 4, the relevant and irrelevant stimuli were consistent (SS^+) on one quarter of the trials and inconsistent (SS^-) on the remaining three quarters. In Type 8 with congruent mapping, SR and SS were consistent (A: SR^+/SS^+) on one quarter of the trials and inconsistent (B: SR^-/SS^-) on the remaining three quarters (see Table 1). In Type 8 with incongruent mapping, one fourth of the trials were Condition C (SR^-/SS^+), one fourth were Condition D (SR^+/SS^-), and the remaining half of the trials were Condition E (SR^-/SS^-).

Procedure. At the beginning of each trial, a 1000-Hz tone was presented over the headphones that subjects wore; at the same time, a display containing four corners that formed a rectangle appeared in the center of the computer screen. Subjects were instructed to fixate the plus sign in the center of this outlined rectangle. After 700 ms, the display was replaced by the stimulus (word triplet) inside the rectangle. The subjects' task was to utter the correct response word based on the identity of the middle word in the stimulus array. They were explicitly told to ignore the top and bottom words (irrelevant stimuli) because these words had no bearing whatsoever on the responses that subjects had to produce. Subjects were instructed to make their responses as accurately and as rapidly as possible. Once they made a response, the stimulus display disappeared. The subjects' response was then keyed into the computer, and response accuracy and RT were recorded. On the basis of the accuracy and speed of the response, a score was calculated. If the response was incorrect, a penalty of −30 points was assessed; if the response was correct, a score ranging from 0 to 100 points was calculated. A fast and correct response earned a score of 100 points; a slow but correct response earned a score of 0 points; and an intermediate and correct response earned a score between 0 and 100 points that was linearly proportional to RT. Shortly after the subject made a response, the subject received feedback regarding his or her performance on this trial. If the subject made a correct response, then he or she saw the visual message "CORRECT, EARN X POINTS" on the computer screen, where X was replaced by the actual points the subject earned on the given trial; if the subject committed an error, however, he or she heard a white noise over the headphones as a penalty and saw the visual message "INCORRECT, LOSE 30 POINTS" on the computer screen. The visual feedback remained on the computer screen for 700 ms. After a 1-second intertrial interval, the next trial began.

Error messages were presented if subjects responded before the stimulus onset ("WAIT FOR STIMULUS") or if they did not respond within 2.5 seconds after the stimulus onset ("TOO SLOW").

Results

The error rate was low, averaging 2.5% in Type 2, 2.3% in Type 3, and 2.2% in both Types 4 and 8. These error rates were not statistically different across ensemble types, $F(3, 15) = 0.23$, $p = .87$, $MSE = 0.000041$. Thus, further data analysis was based on the correct responses only. For each subject and condition, the mean and standard deviation of the correct responses were calculated. Then, responses that were 3.3 *SDs* above or below the mean were trimmed from further analysis.

There were two groups of alternative SR ensembles. In order to eliminate possible stimulus and response differences across ensemble types, we combined these two groups in the data analysis. We combined Sessions 2 and 5 (first presentation) and Sessions 3 and 6 (second presentation). Presentation (instead of session) was the independent variable in the statistical calculations.

We first analyzed the results separately for each ensemble type and then compared the results across ensembles.

Type 2. In Type 2, the relevant stimuli overlapped with, and thus were mapped either congruently or incongruently to, the responses. We analyzed RTs in terms of three dichotomous factors: SR mapping (congruent vs. incongruent), presentation (first vs. second), and response set (color names vs. digit names). The data in Table 3 show the main effects for all these factors. The mean RT for the congruent mapping was significantly faster (425 ms) than that for the incongruent mapping (672 ms), $F(1, 7) = 127.46$, $p < .0001$, $MSE = 7{,}644$. The mean RT in the second presentation was significantly faster (534 ms) than that in the first presentation (564 ms), $F(1, 7) = 12.65$, $p < .01$, $MSE = 1{,}122$. Digit responses to digit stimuli were significantly faster (519 ms) than were color responses to color stimuli (579 ms), $F(1, 7) = 22.43$, $p < .01$, $MSE = 2{,}467$. There was a significant interaction between presentation and SR mapping, $F(1, 7) = 11.38$, $p = .012$, $MSE = 213$, with presentation having a greater effect on incongruent than on congruent mapping (43 vs. 17 ms). Response set and SR mapping also interacted, $F(1, 7) = 16.34$, $p < .005$, $MSE = 1{,}080$, with the mapping effect being 281 ms for color responses and 213 ms for digit responses. There was no interaction between presentation and response set, $F(1, 7) = .12$, $p = .74$, $MSE = 821$, and no three-way interaction among presentation, response set, and SR mapping, $F(1, 7) = .08$, $p = .79$, $MSE = 763$.

Because SR mapping interacted with presentation and with response set, we performed additional analyses. We found congruent SR mapping to be faster than the incongruent SR mapping when the responses were color names (438 ms in congruent vs. 719 ms in incongruent mapping), $F(1, 7) = 110.16$, $p = .0001$, $MSE = 5{,}700$, or digit names (413 ms in congruent vs. 626 ms in incongruent mapping), $F(1, 7) = 120.40$, $p = .0001$, $MSE = 3{,}035$. Furthermore, we found the congruent SR mapping to be faster than the incongruent SR mapping for both the first presentation (434 ms in congruent vs. 694 ms in incongruent mapping), $F(1, 7) = 132.77$, $p = .0001$, $MSE = 4{,}050$ and second presentation (417 ms in congruent vs. 651 ms in incongruent

Table 3
Mean Reaction Times and Standard Deviations (in Parentheses) in Milliseconds From Experiment 1

Condition and response set	Type 2		Type 3		Type 4	
	1st pres.	2nd pres.	1st pres.	2nd pres.	1st pres.	2nd pres.
Con[a]						
Color	448 (52)	428 (67)	628 (114)	603 (94)	643 (131)	607 (101)
Digit	419 (46)	406 (51)	578 (96)	551 (94)	576 (102)	553 (89)
Incon[a]						
Color	741 (170)	697 (160)	665 (124)	632 (107)	676 (127)	629 (99)
Digit	647 (113)	604 (109)	622 (127)	583 (112)	612 (119)	582 (109)

Note. pres. = presentation.
[a]Con and Incon, when applied to Type 2, mean "Congruent" and "Incongruent"; when applied to Types 3 and 4, they mean "stimulus–response consistent or inconsistent" and "stimulus–stimulus consistent or inconsistent," respectively.

mapping), $F(1, 7) = 115.34, p = .0001, MSE = 3{,}818$. Thus, we obtained a reliable, large SR mapping effect for every response set and for every presentation.

Type 3. In Type 3, the irrelevant stimuli and responses were either consistent (SR^+) or inconsistent (SR^-). Table 3 presents both mean RTs and standard deviations for different presentations and response sets. We analyzed these results in terms of SR consistency (consistent vs. inconsistent), presentation (first vs. second), and response set (color names vs. digit names), and we obtained a main effect for each of these factors. The mean RT for SR consistent conditions was significantly faster (590 ms) than that for SR inconsistent conditions (625 ms), $F(1, 7) = 27.44, p = .0012, MSE = 687$. The mean RT in the second presentation (592 ms) was significantly faster than that in the first presentation (623 ms), $F(1, 7) = 15.08, p = .006, MSE = 932$. Digit responses to color stimuli were slightly faster than were color responses to digit stimuli (584 ms vs. 632 ms), $F(1, 7) = 4.67$, $p = .068$, $MSE = 8{,}043$. None of the interactions were significant.

Type 4. In Type 4, the relevant and irrelevant stimuli were either consistent (SS^+) or inconsistent (SS^-). Table 3 presents the mean RTs and standard deviations separately for various presentations and response sets. Results in Type 4 were similar to those in Type 3. We obtained a main effect for SS consistency, for presentation, and for response set. RT was faster for the SS consistent condition (594 ms) than for the SS inconsistent condition (625 ms), $F(1, 7) = 59.65, p = .0001$, $MSE = 244.3$, indicating a main effect of 31 ms for SS consistency. The second presentation produced a significantly faster RT (593 ms) than the first presentation produced (627 ms), $F(1, 7) = 19.94, p = .003, MSE = 888$. Digit responses to color stimuli (581 ms) were significantly faster than were color responses to digit stimuli (639 ms), $F(1, 7) = 7.82, p = .027, MSE = 6{,}695$. No interaction was significant.

Type 8 (the Stroop task). In the Stroop task, the overall mean RT for the incongruent mapping (Conditions C, D, and E) was slower (690 ms) than for the congruent mapping (Conditions A and B; 442 ms), $F(1, 7) = 122.83, p < .0001$, $MSE = 7{,}789$, indicating a mapping effect of 248 ms. (One could argue that the difference between Conditions B and E, for which the SR and SS consistency values were equal, is a purer measure of the mapping effect in Type 8. On this measure, the mapping effect is 265 ms.)

Within the congruent mapping, Condition A was faster (425 ms) than Condition B (448 ms), a significant Stroop effect of 23 ms, $F(1, 7) = 100.6, p = .0001, MSE = 91$. Digit name responses to digit stimuli were significantly faster (425 ms) than were color name responses to color stimuli (448 ms), $F(1, 7) = 8.17, p = .024, MSE = 1{,}015$. The second presentation (432 ms) was faster than the first presentation (441 ms), but not significantly so, $F(1, 7) = 2.44, p = .16, MSE = 444$. No interaction was significant.

Within incongruent mapping, digit name responses to digit stimuli were significantly faster (646 ms) than color name responses to color stimuli (720 ms), $F(1, 7) = 11.48$, $p = .012, MSE = 10{,}591$; the second presentation (668 ms) was significantly faster than the first presentation (698 ms), $F(1, 7) = 5.48, p = .052, MSE = 3{,}588$. Conditions C, D, and E were also statistically different, $F(2, 14) = 11.05, p = .0013, MSE = 2{,}212$. Further analyses indicated significant differences between them all: Condition C (SR^-/SS^+) faster (656 ms) than both Condition D (SR^+/SS^-; 679 ms), $F(1, 7) = 14.71, p = .0064, MSE = 519$ and Condition E (SR^-/SS^-; 713 ms) $F(1, 7) = 15.43, p = .0057, MSE = 3{,}125$; and Condition D ($SR^+/SS^-$) faster than Condition E (SR^-/SS^-), $F(1, 7) = 5.84, p = .046, MSE = 2{,}993$. Thus, we obtained reliable effects for both SR and SS consistency. None of the interactions were significant. Table 4 presents mean RTs and standard deviations for different conditions, presentations, and response sets.

The only ensemble for which there were any significant interactions was Type 2. Here, SR mapping interacted with presentation as well as with response set. However, the mapping effect was significant for both presentations and both response sets—it just differed in degree. It, therefore, seemed reasonable to collapse RTs over all presentations and response sets. Table 5 presents the resulting grand mean RTs, averaged over all subjects, presentations, and response sets. In summary, main effects were obtained for SR mapping in Types 2 and 8 (247 ms and 265 ms, respectively), SR consistency in Type 3 (35 ms), and SS consistency in Type 4 (31 ms). In the congruent Stroop tasks, we obtained a

Table 4
Mean Reaction Times and Standard Deviations (Milliseconds) in Type 8 From Experiment 1

Condition	SR mapping	SR/SS consistency	Response Set	1st presentation *M*	1st presentation *SD*	2nd presentation *M*	2nd presentation *SD*
A	Congruent	SR^+/SS^+	Color	444	49	432	60
			Digit	414	51	410	45
B	Congruent	SR^-/SS^-	Color	465	61	452	60
			Digit	441	62	434	49
C	Incongruent	SR^-/SS^+	Color	710	179	671	146
			Digit	634	133	612	104
D	Incongruent	SR^+/SS^-	Color	734	167	701	147
			Digit	659	122	623	102
E	Incongruent	SR^-/SS^-	Color	769	216	733	172
			Digit	682	134	667	121

Note. Superscript plus sign indicates consistent and superscript minus sign indicates inconsistent. SR = stimulus–response; SS = stimulus–stimulus.

significant Stroop effect of 23 ms (RT difference between Conditions A and B). We obtained significant effects in the incongruent Stroop tasks for both SS consistency (57 ms; RT difference between Conditions C and E) and SR consistency (34 ms; RT difference between Conditions D and E).

We would like to note that the congruent Type 2 may be considered as a baseline condition for congruent Type 8 conditions (Conditions A and B, see Table 1), the RT difference between congruent Type 2 and Condition A (SR^+/SS^+) should reveal a facilitation effect, and the RT difference between congruent Type 2 and Condition B (SR^-/SS^-) should reveal an interference effect. Because RT was the same (425 ms) in congruent Type 2 and in Condition A, no facilitation occurred. Therefore, the Stroop effect seemed to be due mainly to interference.

Table 5
Grand Mean Reaction Times and Standard Deviations (Milliseconds) From Experiment 1

Ensemble type and mapping/consistency	*M*	*SD*
2		
Congruent	425	57
Incongruent	672	150
3		
SR consistent (SR^+)	590	104
SR inconsistent (SR^-)	625	121
4		
SS consistent (SS^+)	594	112
SS inconsistent (SS^-)	625	119
8		
Condition A: Congruent (SR^+/SS^+)	425	53
Condition B: Congruent (SR^-/SS^-)	448	59
Condition C: Incongruent (SR^-/SS^+)	656	147
Condition D: Incongruent (SR^+/SS^-)	679	143
Condition E: Incongruent (SR^-/SS^-)	713	170

Note. Superscript plus sign indicates consistent and superscript minus sign indicates inconsistent. SR = stimulus–response; SS = stimulus–stimulus.

Discussion

In this experiment we produced several important findings. First, we obtained significant main effects for SR mapping, SR consistency, and SS consistency in Types 2, 3, and 4, respectively. Second, both SR and SS consistency were significant constituents of the Stroop effect. Third, the Stroop effect appears to be due mainly to interference.

The size of the Stroop effect reported in the literature is roughly 100 ms or more (Dyer, 1971, 1974; Hintzman et al., 1972). The effect (23 ms) obtained in this experiment is, therefore, relatively small. Two factors may have contributed to this. First, the physical arrangement of our stimulus display may have caused the Stroop effect to be smaller than in a standard Stroop task. We used a vertical version of the flanker paradigm (Eriksen & Eriksen, 1974; Shaffer & LaBerge, 1979) in which the relevant and irrelevant stimuli were spatially separated. In a typical Stroop task (Stroop, 1935/1992) the relevant and irrelevant stimuli are spatially integrated into a single perceptual object. Even though the spatial separation between the relevant and irrelevant stimuli in our experiment was small, subjects may have been able to ignore the irrelevant information in our study more easily than they would have if the two had been spatially integral (Eriksen & Eriksen, 1974, 1979; Miller, 1991), as they are in standard Stroop tasks. Second, and most important, both the relevant and irrelevant stimuli were words, which, according to both the speed-of-processing and the automaticity views, produces two equal-strength response tendencies. In the typical Stroop task, in contrast, the ink color and color word each generates a response tendency, but these response tendencies differ in strength. Stronger processes are expected to interfere more with weaker ones (as in the typical Stroop task) than vice versa. Equal strength processes (as in Experiment 1) are expected to have small effects on each other. Despite its small size, however, the Stroop effect in our experiment is reliable.

Next, we consider an analysis of the Stroop effect according to the dimensional overlap model. In the congruent Stroop task, when a trial is consistent (in row A of Table 1), it is consistent for both SS and SR; when it is inconsistent (in row B of Table 1), it is inconsistent for both SS and SR. As we pointed out at the beginning of this article, the RT difference between these two conditions (Conditions A and B) is, therefore, attributable to either SS or SR consistency. This ambiguity is eliminated in the incongruent Stroop tasks. First, we note that the RT for Condition C (SR^-/SS^+) was faster than for Condition D (SR^+/SS^-). This difference is in agreement with the findings in the two-choice literature (Green & Barber, 1981; Kornblum, 1992; Simon & Sudalaimuthu, 1979) and is attributable either to the effect of SS consistency dominating over the effect of SR consistency or to logical recoding (Hedge & Marsh, 1975). This difference also raises the question of whether SR consistency has any effect at all—some have maintained that it does not (e.g., Stoffels, van der Molen, & Keuss, 1989). This last question is answered by comparing Condition D (SR^+/SS^-) with Condition E (SR^-/SS^-), for those conditions differed only in terms of SR consistency. The fact that Condition E (SR^-/SS^-), in which SR is inconsistent, was slower than Condition D (SR^+/SS^-), in which SR is consistent, indicates that SR consistency does have an effect in the Stroop task, contrary to what Stoffels et al. (1989) have maintained. Similarly, the fact that Condition E (SR^-/SS^-) was slower than Condition C (SR^-/SS^+) indicates that SS consistency also has an effect in the Stroop task and that logical recoding, while it may be occurring, is not the whole story. Furthermore, the effects of SS and SR consistency obtained in the incongruent Stroop task are not additive, for a simple additive model would require the RT difference between Conditions A and B to be the sum of SS and SR consistency effects, 91 ms (57 ms plus 34 ms). But the actual difference is 23 ms, $t(7) = 37.359$, $p < .001$. We, therefore, conclude that the Stroop effect is attributable to both SS and SR consistency.

Given that the effects in Ensemble 8 are attributable to SR mapping as well as to SS and SR consistency, can the effects of these various overlap relations, when obtained in isolation (in Types 2, 3, and 4), be combined to produce the Stroop effect? From the data in Table 5, the effect of these factors in isolation are as follows: From Type 2 we note an effect of 247 ms for SR mapping; from Type 3, an SR consistency effect of 35 ms; and from Type 4, an SS consistency effect of 31 ms. We note that in Type 8, Condition D (SR^+/SS^-) differed from Condition E (SR^-/SS^-) only in terms of SR consistency. Using a simple, linear additive model, we would therefore predict that these two conditions would differ by the SR consistency effect obtained in Type 3, which is 35 ms. The actual difference obtained was 34 ms (see Table 5). This particular comparison thus seems to be consistent with an additive model.

However, an examination of other aspects of the data reveals a totally different picture. First, if the SS and SR consistency effects were additive in Type 8 (as they appear to be in Type 7, see Kornblum, 1994), then the Stroop effect would simply be the sum of the SS and SR consistency effects. However, the size of the Stroop effect is 23 ms, and the sum of the SS and SR consistency effects is 66 ms, and this difference is statistically significant, $t(7) = 18.113$, $p < .01$. Second, if the effects of SS and SR consistency were additive, then the RTs for Conditions C and D should be almost identical; the consistency effect for SS is 31 ms and for SR is 35 ms, and under a linear model they would cancel each other. However, they differ by 23 ms, which is significantly different from 0, $t(7) = 4.56$, $p < .01$. An additive model thus seems to be an oversimplification.

In summary, the various effects in the Stroop task appear to be attributable to both SS and SR overlap. Furthermore, the effects for SR mapping, and SR and SS consistency, when obtained in isolation in Types 2, 3, and 4, respectively, do not add in a simple, linear manner to yield the Stroop effects; they appear to be interactive.

Experiment 2

The main purpose of Experiment 2 was to test the generality of the results obtained in Experiment 1 in nonverbal tasks. A second purpose was to include neutral baseline conditions for Ensemble Types 2, 3, and 4 so that we could calculate facilitation and interference effects for these ensembles.

Method

Subjects. Ten male University of Michigan undergraduate students were tested individually. They were selected according to the same criteria and were paid the same amount as those in Experiment 1. None of them had participated in Experiment 1.

Apparatus and procedure. The apparatus and procedure were the same as for Experiment 1.

Stimuli and responses. The stimulus display was composed of three side-by-side rectangles inside an overall rectangular frame of 2.2 cm × 1.2 cm (subtending a visual angle of 1.68° × 0.92°), which was presented in the center of a black computer screen. The middle rectangle was the relevant stimulus; the left and right flanking rectangles were the irrelevant stimuli and were identical. Each rectangle measured 0.55 cm × 0.95 cm (0.42° × 0.73°). They were separated horizontally by an intercontour distance of 0.12 cm (0.09°).

The relevant stimulus was either a digit (e.g., 2, 4, 6, or 8) or a color patch (e.g., red, green, blue, or yellow color). The irrelevant stimuli were digits (e.g., 2, 4, 6, and 8), false fonts (e.g., ■, ■, ■, and ■), color patches (e.g., red, green, blue, and yellow colors), or gray patches.

The responses were vocal and consisted of color names (e.g., "RED," "GREEN," "BLUE," and "YELLOW") and digit names (e.g., "TWO," "FOUR," "SIX," and "EIGHT"). The relevant stimulus, irrelevant stimulus, and response sets were selected so as to generate instances of Ensembles 1, 2, 3, 4, and 8. These ensembles are shown in Table 6.

Experimental conditions. With the exception of Type 1, the experimental conditions in this experiment were identical to those in Experiment 1. Because the relevant stimulus, irrelevant stimulus, and response sets in Ensemble 1 did not have dimensional overlap, all the trials in Ensemble 1 were neutral.

Design. Similar to Experiment 1, we used a within-subjects design that included six sessions per subject. All five ensemble types were run within a session. The SR ensembles in Table 6 were divided into two groups. Group 1 included the following: Type 1—color names, Type 2—color names, Type 3—color names, Type

Table 6
Stimulus and Response Sets Used in Experiment 2

Ensemble type	Relevant stimulus set	Response set	Irrelevant stimulus set
1	Digits	Color names	False fonts and gray-patches
	Color patches	Digit names	False fonts and gray-patches
2	Color patches	Color names	Digits
	Digits	Digit names	Color patches
3	Digits	Color names	Color patches
	Color patches	Digit names	Digits
4	Digits	Color names	Digits
	Color patches	Digit names	Color patches
8	Color patches	Color names	Color patches
	Digits	Digit names	Digits

Note. Digits were 2, 4, 6, and 8, and color patches were red, green, yellow, and blue.

4—digit names, Type 8—color names; Group 2 included the balance. Half the subjects were run with Group 1 on the first three sessions (sessions 1–3), and with Group 2 on the last three sessions (sessions 4–6). For the other half of the subjects, this order was reversed. There was a break, ranging from one to several days, between the third and fourth sessions.

Each session consisted of 14 blocks. The following conditions were each run for two blocks: Type 2—congruent, Type 2—incongruent, Type 8—congruent, Type 8—incongruent, Type 1, Type 3, and Type 4. This particular order was used for all subjects during practice (Sessions 1 and 4). Data from the practice sessions were not analyzed. On Session 2, five ensemble types and 10 subjects were counterbalanced with two balanced Latin squares, and on Session 3, the reverse Latin squares were used. On Session 5, five ensembles and 10 subjects were counterbalanced with two different balanced Latin squares, and on Session 6, the reverse Latin squares were used.

Each block contained 48 trials and, for Types 2, 3, 4, and 8, had the same composition as in Experiment 1. In Type 1, all trials were neutral. In one block, the relevant stimuli consisted of four digits, and the responses were four color names; in the other block, the relevant stimuli were four colors, and the responses were four digit names. In both blocks, the irrelevant stimuli were four false fonts or four gray patches. Thus, there were 32 unique stimulus triplets (4 relevant stimuli by 8 irrelevant stimuli). All stimulus triplets were shown once, and some of them were presented twice to form a block of 48 trials.

Results

The error rate was low (2.2% in Types 1 and 3, 2.1% in Type 2, 1.6% in Type 4, and 1.8% in Type 8) and identical in different ensembles, $F(4, 36) = 1.13$, $p = .36$, $MSE = 0.000061$. Thus, as in Experiment 1, we restricted further data analysis to correct responses only. In addition, presentation (instead of session) was an independent variable in the statistical calculations.

Type 1. In Type 1, RT was significantly faster with digit responses to color stimuli (534 ms) than with color responses to digit stimuli (583 ms), $F(1, 9) = 9.68$, $p = .0125$, $MSE = 1{,}091$. When colors were the relevant stimuli, there was no significant difference between gray patches and false fonts as the irrelevant stimuli (533 vs. 534 ms). On the other hand, when digits were the relevant stimuli, a longer RT (602 ms) was produced with false fonts than with gray patches (564 ms) as the irrelevant stimuli, $F(1, 9) = 18.13$, $p = .0021$, $MSE = 379$. The mean RTs are presented in Table 7.

Type 2. RTs were analyzed in terms of SR mapping (congruent vs. incongruent), presentation (first vs. second), and response set (color names vs. digit names). The mean RT for congruent mapping was significantly faster (444 ms) than for incongruent mapping (663 ms), $F(1, 9) = 245.35$, $p < .0001$, $MSE = 3{,}947$. The mean RT in the second presentation was significantly faster (536 ms) than in the first presentation (571 ms), $F(1, 9) = 30.03$, $p < .001$, $MSE = 789$. The digit responses to digit stimuli (497 ms) were significantly faster than the color responses to color stimuli (610 ms), $F(1, 9) = 102.49$, $p < .0001$, $MSE = 2{,}487$. There was a significant interaction between SR mapping and presentation, $F(1, 9) = 6.36$, $p < .033$, $MSE = 1{,}245$ and a marginally significant interaction between SR mapping and response set, $F(1, 9) = 3.99$, $p < .077$, $MSE = 1{,}226$. There was no two-way interaction between presentation and response set, $F(1, 9) = .18$, $p = .68$, $MSE = 1{,}265$, nor was there a three-way interaction between SR mapping, presentation, and response set, $F(1, 9) = .0$, $p = .974$, $MSE = 789$. Table 8 presents the mean RTs and standard deviations for Type 2. The major finding in Type 2 is a reliable, large SR mapping effect (219 ms).

Type 3. We analyzed RTs in terms of SR consistency (consistent vs. inconsistent), presentation (first vs. second), and response set (color names vs. digit names), and we obtained main effects for each of these factors. The mean RT for SR consistent condition was significantly faster (536 ms) than for SR inconsistent condition (573 ms), $F(1, 9) = 32.03$, $p = .0003$, $MSE = 824$. The mean RT in the second presentation (539 ms) was significantly faster than in the first presentation (571 ms), $F(1, 9) = 15.14$, $p = .0037$, $MSE = 1{,}381$. In contrast to the results of Experiment 1, there was no significant effect of response set, $F(1, 9) = 1.45$, $p = .26$, $MSE = 10{,}222$. However, there was a significant interaction between SR consistency and response set, $F(1, 9) = 5.60$, $p = .0421$, $MSE = 1{,}292$ and a marginally significant interaction between presentation and response set, $F(1, 9) = 4.23$, $p = .07$, $MSE = 763$. Neither the two-way interaction between presentation and SR consistency nor the three-way interaction between presentation, SR consistency, and response set was significant. Table 8

Table 7
Mean Reaction Times and Standard Deviations (Milliseconds) in Type 1 From Experiment 2

	Irrelevant stimulus set			
	Gray patches		False fonts	
Relevant stimulus set	*M*	*SD*	*M*	*SD*
Color patches	533	113	534	123
Digits	564	120	602	117

Table 8
Mean Reaction Times and Standard Deviations (in Parentheses) in Milliseconds From Experiment 2

Condition and Response set	Type 2 1st pres.	Type 2 2nd pres.	Type 3 1st pres.	Type 3 2nd pres.	Type 4 1st pres.	Type 4 2nd pres.
Con[a]						
Color	501 (73)	483 (82)	580 (102)	542 (110)	590 (154)	549 (127)
Digit	401 (47)	389 (47)	526 (110)	498 (104)	603 (140)	541 (116)
Incon[a]						
Color	757 (214)	698 (188)	601 (124)	553 (113)	629 (156)	589 (133)
Digit	624 (167)	574 (119)	576 (153)	563 (155)	595 (129)	575 (126)

Note. pres = presentation.
[a]Con and Incon, when applied to Type 2 mean "congruent" and "incongruent"; when applied to Types 3 and 4, they mean "stimulus–response consistent or inconsistent" and "stimulus–stimulus consistent or inconsistent," respectively.

presents mean RTs and standard deviations in Type 3. The main results were similar to those in Experiment 1—a reliable main effect for SR consistency.

Type 4. We analyzed RTs in terms of SS consistency (consistent vs. inconsistent), presentation (first vs. second), and response set (digit names vs. color names) and main effects were obtained for each of these factors. The mean RT for the SS consistent condition was faster (571 ms) than for the SS inconsistent condition (597 ms), $F(1, 9) = 72.33, p = .0001$, $MSE = 195$. The second presentation (564 ms) produced a significantly faster RT than the first presentation (604 ms), $F(1, 9) = 20.03, p = .0015$, $MSE = 1{,}523$. Digit responses to color stimuli (578 ms) were not different from color responses to digit stimuli (589 ms), $F(1, 9) = .26, p = .624$, $MSE = 10{,}877$. There was a significant interaction between SS consistency and response set, $F(1, 9) = 8.06$, $p = .0194$, $MSE = 382$ and a marginally significant interaction between SS consistency and presentation, $F(1, 9) = 3.90, p = .08$, $MSE = 416$. Neither the two-way interaction between presentation and response set nor the three-way interaction between presentation, SS consistency, and response set was significant. Table 8 presents the mean RTs and standard deviations in Type 4. As in Experiment 1, we obtained a reliable main effect for SS consistency.

Type 8 (the Stroop task). In the Stroop task, the overall mean RT for incongruent mapping (Conditions C, D, and E) was significantly slower (685 ms) than that for congruent mapping (Conditions A and B; 465 ms), $F(1, 9) = 187.46$, $p < .0001$, $MSE = 5{,}106$, indicating an SR mapping effect of 220 ms. (As in Experiment 1, one could argue that the difference between Conditions B and E, 225 ms, is a purer measure of the mapping effect in Type 8.)

Within congruent mapping, Condition A was faster (446 ms) than Condition B (471 ms), a significant Stroop effect of 25 ms, $F(1, 9) = 37.96$, $p = .0002$, $MSE = 315$. Digit responses to digit stimuli produced a significantly faster RT (421 ms) than did color responses to color stimuli (496 ms), $F(1, 9) = 54.41$, $p = .0001$, $MSE = 2{,}088$. The second presentation resulted in a significantly faster RT (451 ms) than did the first presentation (466 ms), $F(1, 9) = 14.04, p = .0046$, $MSE = 296$. There were no significant two-way interactions; the three-way interaction between presentation, condition (between Conditions A and B), and response set was marginally significant, $F(1, 9) = 4.19, p = .07$, $MSE = 39$.

Within incongruent mapping, digit responses to digit stimuli were significantly faster (619 ms) than were color responses to color stimuli (741 ms), $F(1, 9) = 56.93$, $p = .0001$, $MSE = 7{,}625$. The second presentation yielded a significantly faster RT (655 ms) than did the first presentation (706 ms), $F(1, 9) = 29.91, p = .0004$, $MSE = 2{,}400$. Conditions C, D, and E were also statistically different, $F(2, 18) = 20.64$, $p = .0001$, $MSE = 562$. Further analysis revealed a significantly faster RT for Condition C (SR^-/SS^+; 662 ms) than for Condition D (SR^+/SS^-; 684 ms), $F(1, 9) = 23.30, p = .0009$, $MSE = 373$; a faster RT for Condition C (SR^-/SS^+) than for Condition E (SR^-/SS^-; 696 ms), $F(1, 9) = 36.57, p = .0002$, $MSE = 623$; and a marginally significantly faster RT for Condition D than for Condition E, $F(1, 9) = 4.82$, $p = .056$, $MSE = 690$. However, no interaction was significant. Table 9 presents mean RTs and standard deviations in Type 8.

As we did for the data in Experiment 1, we averaged RTs over all subjects, all presentations, and all response sets to generate grand mean RTs for various conditions. Table 10 shows the results. In summary, main effects were obtained for SR mapping in Types 2 and 8 (219 ms and 225 ms, respectively), SR consistency in Type 3 (37 ms), SS consistency in Type 4 (26 ms), and a Stroop effect of 25 ms for the congruent mapping condition in Type 8.

As in Experiment 1, we used congruent Ensemble 2 as a baseline condition for congruent Stroop tasks and found that the Stroop effect (25 ms) was mainly an interference (RT difference between congruent Type 2 and Condition B), not a facilitation (no difference between congruent Type 2 and Condition A). In Type 8—incongruent mapping, we found a significant effect of SS consistency (34 ms, RT difference between Conditions C and E) and SR consistency (12 ms, RT difference between Conditions D and E).

Using Type 1 as a baseline for Type 2, 3, and 4, we found both facilitation (119 ms) and interference (100 ms) of SR mapping in Type 2, a large facilitation (27 ms) and a small interference (10 ms) effect of SR consistency in Type 3, and no facilitation with an interference effect (24 ms) of SS

Table 9
Mean Reaction Times and Standard Deviations (Milliseconds) in Type 8 From Experiment 2

				1st Presentation		2nd Presentation	
Condition	SR mapping	SR/SS consistency	Response set	*M*	*SD*	*M*	*SD*
A	Congruent	SR^+/SS^+	Color	486	61	478	73
			Digit	416	48	404	47
B	Congruent	SR^-/SS^-	Color	522	80	499	81
			Digit	439	56	423	53
C	Incongruent	SR^-/SS^+	Color	756	215	688	183
			Digit	627	126	578	136
D	Incongruent	SR^+/SS^-	Color	767	204	709	196
			Digit	646	121	611	125
E	Incongruent	SR^-/SS^-	Color	792	227	736	215
			Digit	645	122	608	135

Note. Superscript plus sign indicates consistent and superscript minus sign indicates inconsistent. SR = stimulus–response; SS = stimulus–stimulus.

consistency in Type 4. Note that the SR consistency effect was principally facilitatory, whereas the SS consistency effect was primarily interfering. Unlike the results in Experiment 1, the SS and SR consistency effects in Experiment 2 differed.

Discussion

The results from this experiment are similar, in many ways, to those of Experiment 1. The Stroop effect is still small (25 ms) but is roughly the same size (23 ms) as we found in Experiment 1. The small size of this effect is likely attributable to the same factors: first, the physical separation between the relevant and the irrelevant stimuli, and second, the fact that the relevant and irrelevant stimuli used the same carriers.

As for the analysis of the Stroop effect in terms of SS and SR consistency, the results in Experiment 2 are almost identical to those of Experiment 1, which generalizes our results to non-word tasks. The RT for Condition C (SR^-/SS^+) was faster than for Condition D (SR^+/SS^-), which supports both the notion that the SS consistency effect dominates over the SR consistency effect and the logical recoding hypothesis (Hedge & Marsh, 1975). Condition D (SR^+/SS^-) was faster than Condition E (SR^-/SS^-), which, because these two conditions differ in the sign of SR consistency only, indicates that SR consistency also has an effect (12 ms). Finally, Condition C (SR^-/SS^+) was faster than Condition E (SR^-/SS^-), which supports the hypothesis that SS consistency has an effect (34 ms). Furthermore, the effects of SS and SR consistency obtained in the incongruent Stroop task are not additive, for a simple additive model would require the RT difference between Conditions A and B to be the sum of SS and SR consistency effects obtained in isolation, 46 ms (34 ms plus 12 ms). However, the actual difference is 25 ms, $t(9) = 5.44$, $p < .01$. Therefore, our conclusions from the results of Experiment 1 that the Stroop effect is attributable to both SS and SR consistency and that these consistency effects are not linearly additive are fully supported by the results of Experiment 2.

As in Experiment 1, we compared performance in the Stroop task to that in Types 2, 3, and 4. From the data in Table 10, the effects in isolation are from Type 2, an SR mapping effect of 219 ms; from Type 3, an SR consistency effect of 37 ms; and from Type 4, an SS consistency effect of 26 ms. We note from Table 1 that in Type 8, Conditions D (SR^+/SS^-) and E (SR^-/SS^-) differed in terms of SR consistency only. A simple, linear additive model would, therefore, predict that these two conditions would differ by the SR consistency effect obtained in Type 3, which is 37 ms. The actual difference is 14 ms (see Table 10), $t(9) = 3.50$, $p < .01$. (In Experiment 1 the difference between Conditions D and E was in fact equal to the isolated SR consistency

Table 10
Grand Mean Reaction Times and Standard Deviations (Milliseconds) From Experiment 2

Ensemble type and mapping/consistency	*M*	*SD*
1		
Neutral	563	128
2		
Congruent	444	81
Incongruent	663	188
3		
SR consistent (SR^+)	536	111
SR inconsistent (SR^-)	573	138
4		
SS consistent (SS^+)	571	138
SS inconsistent (SR^-)	597	138
8		
Condition A: Congruent (SR^+/SS^+)	446	69
Condition B: Congruent (SR^-/SS^-)	471	80
Condition C: Incongruent (SR^-/SS^+)	662	181
Condition D: Incongruent (SR^+/SS^-)	684	177
Condition E: Incongruent (SR^-/SS^-)	696	195

Note. Superscript plus sign indicates consistent and superscript minus sign indicates inconsistent. SR = stimulus–response; SS = stimulus–stimulus.

effect as measured in Type 3.) If the Stroop effect with congruent mapping were simply the sum of the SS and SR consistency effects, then its size should have been the sum of these two effects in isolation, or 63 ms (26 ms plus 37 ms); however, the Stroop effect is only 25 ms, which is significantly smaller than the sum, $t(9) = 9.846, p < .01$. A second case in which one could test the additivity of the SS and SR consistency effects is the difference between Conditions C (SR^-/SS^+) and D (SR^+/SS^-): If the effects of SS and SR consistency were linearly additive, Condition D would be faster than Condition C by 11 ms. However, the opposite is the case: Condition C is faster than Condition D by 22 ms, which is significantly different from the prediction from a linearly additive model, $t(9) = 6.324, p < .01$.

We now turn to another observation. In this experiment, the overall RTs in Type 4 were about 30 ms longer than in Type 3 even though the relevant stimuli, irrelevant stimuli, and responses were balanced across these two ensembles. This RT difference is partly due to a similarity differential in the stimuli between Types 3 and 4. In Type 3, when the relevant stimulus was a color patch, the irrelevant stimulus consisted of digits, and if the relevant stimulus was a digit, then the irrelevant stimulus consisted of color patch flankers. Here the relevant and irrelevant stimuli were perceptually distinct. In contrast, in Type 4, because there was dimensional overlap between the relevant and the irrelevant stimulus, if the relevant stimulus was a color patch, then the irrelevant stimulus consisted of color patches as well, and if the relevant stimulus was a digit, then the irrelevant stimuli were also digits. That is, the relevant and irrelevant stimuli belonged to the same perceptual category and were perceptually more similar than in Type 3. (Note that in Experiment 1, both the relevant and the irrelevant stimuli were words; thus, even though they differed conceptually, i.e., in meaning, they were uniformly similar at the perceptual level for both Types 3 and 4.)

This effect of stimulus similarity was also noted in Type 1, where there was no dimensional overlap between the relevant stimulus, the irrelevant stimulus, and the response sets. However, when the relevant stimulus was a digit, RT was slower if the irrelevant stimulus was a false font (which is visually similar to the digit stimulus) than if it was patches of gray.

General Discussion

Summary of Results

By using single-carrier stimuli and four-choice tasks with congruent and incongruent SR mapping instructions, we were able to address a number of important issues in the study of the Stroop effect. First, we eliminated the confounding between stimulus and response effects that is inherent in the standard, congruent Stroop task. Second, we separated SS effects from the logical recoding hypothesis, which is not possible in two-choice, incongruent Stroop tasks. Third, by using single-carrier stimuli we eliminated any effects attributable to differences in the basic processing speed of two different processing functions (e.g., reading and color naming). The results from both experiments converge on the same conclusions. The Stroop effect is the result of a combination of SS and SR consistency effects. That is, neither the processes associated with SS overlap nor those associated with irrelevant SR overlap by themselves can adequately explain the pattern of performance in Stroop tasks. These findings also fail to support the logical recoding hypothesis as either a necessary or sufficient account of the Stroop effect. We also ascertained that the effects of SR mapping, and SR and SS consistency, which we obtained in isolation in Types 2, 3, and 4, respectively, do not combine in a linear fashion to produce any of the effects in the Stroop task. We have incorporated these concerns and considerations in a new interactive network model that we summarize next.

The Interactive Activation Model

Cohen, Dunbar, and McClelland (1990) have proposed an interactive network for the Stroop task that consists of two pathways: one for processing ink colors, the other for processing color words. The ink-color pathway was made the weaker of the two. Built into the model is the feature that the stronger pathway interferes with the weaker one, but not the other way around. The results of the model are consistent with the experimental results obtained with congruent mapping for the Stroop task. However, the Cohen et al. (1990) model cannot deal with incongruent Stroop tasks, nor does it address the relation between performance on the Stroop task and performance on Types 2, 3, or 4.

Given that the dimensional overlap model (Kornblum, 1992; Kornblum et al., 1990; Kornblum & Lee, 1995) provides a qualitative account of the results for both congruent and incongruent Stroop tasks, we implemented the major assumptions of the model in an interactive activation network (Zhang, Kornblum, & Zhang, 1995; Zhang, Zhang, & Kornblum, 1997). In the next few paragraphs, we summarize and demonstrate the utility of this model in explaining the Stroop results.

Our network is similar in its basic architecture to other interactive activation models in the literature (e.g., Cohen, Servan-Schreiber, & McClelland, 1992; McClelland & Rumelhart, 1981; Phaf, van der Heijden, & Hudson, 1990). At the heart of the network is the notion of a module that represents stimulus or response dimensions (e.g., colors). Each module consists of several nodes, each of which, in turn, represents a stimulus or response feature (e.g., a red color). Associated with each node in the network are input and activation values. Nodes within the same module are negatively connected to reflect the fact that an object cannot be both red and green at the same time. The strength of the negative connection within a module was set at -0.025.

These modules and nodes are organized into a feed-forward, three-layered (input-intermediate-output) network. Modules and nodes at the input layer stand for stimuli that are carrier-specific and thus are tied to the perceptual encoding of the stimuli; those at the intermediate layer stand for carrier-free concepts and thus are tied to the semantic meaning of the stimuli; and those at the output layer are for responses. For example, ink colors and color words are different carriers and are therefore represented by different modules at the input layer. However, both ink colors and color words are related to the concept of color and therefore are represented in a common module at the intermediate layer. Different modules in the same layer represent different stimulus or response categories, hence they are not connected. When two stimulus dimensions overlap (SS overlap), the input modules that represent the stimuli converge onto a common module at the intermediate layer. This convergence assumption implements the notion of dimensional overlap between two stimulus dimensions and appears to be consistent with recent physiological evidence (Damasio & Damasio, 1992). According to the dimensional overlap model (Kornblum, 1992; Kornblum et al., 1990), when the relevant stimulus and the response overlap (SR overlap), a rule-governed, controlled process and an automatic process are activated and together determine the correct response. In the interactive activation network, the control process is implemented by control lines and the automatic process by automatic lines, which join nodes at the intermediate and output layers.

With these representational assumptions, we constructed an interactive activation network for the Stroop task, which is illustrated in Figure 1. The Stroop stimuli are two-dimensional—one dimension relevant and the other irrelevant—and each dimension consists of four individual stimulus features (e.g., red, green, blue, and yellow). The relevant and irrelevant stimulus dimensions are represented by two separate modules at the input layer. Each module is composed of four nodes, representing the four stimulus features used in the present study. These nodes receive external stimulation through task lines, the strength of which was set at 0.03 for the relevant pathway and 0.011 (in Experiment 1) or 0.006 (in Experiment 2) for the irrelevant pathway. We made the strength larger for the relevant than for the irrelevant pathway in order to simulate task instructions which require subjects to pay attention to the relevant and ignore the irrelevant stimulus. Because the relevant and irrelevant stimuli are conceptually similar (SS overlap), the two modules at the input layer converge onto the common module at the intermediate layer, and their corresponding nodes are joined through carrier lines. The strength of the carrier lines was set at 0.02 for ink colors and 0.04 for other stimuli (e.g., words and digits).

Nodes at the intermediate layer are connected to those at the output layer through the control lines: In the congruent mapping, the control lines link the nodes that represent similar concepts (e.g., both representing blue), but in the incongruent mapping, the control lines link the nodes that represent different stimuli and responses (e.g., the stimulus was blue and the response was "Green"). This is the only difference between congruent and incongruent mapping, as shown in Figure 1. The strength of the control lines was set at 0.03 for incongruent SR mapping and 0.2 (in Experiment 1) or 0.04 (in Experiment 2) for congruent mapping. The control lines were stronger in the congruent mapping than in the incongruent mapping because, according to the dimensional overlap model, the fast identity rule is used to arrive at the correct response in the congruent mapping, and a time-consuming memory search or other rule is used in the incongruent mapping. Furthermore, because the stimuli overlap with the responses (SR overlap), corresponding nodes at the intermediate and output layers were linked through the automatic lines. The strength of the automatic lines was set at 0.02 (in Experiment 1) or 0.024 (in Experiment 2). In the congruent mapping, the control and automatic lines went to the same nodes at the output layer, but in the incongruent mapping, they went to different nodes.

This network (Figure 1) was then used to simulate the results in the Stroop tasks. We assumed that at $t = 0$ all the nodes in the network had reached the quiescent state and that the activation of each node was 0. At $t = 1$, external stimulation was presented and all stimulus features were identified immediately. If a feature was present in the stimulus display, then external stimulation sent a value of 1 to the corresponding nodes at the input layer; otherwise, a value of 0 was sent. The input for node i was weighted with the input function,

$$\text{Input}_i(t) = (1 - \text{Decay}) \times \text{Activation}_i(t-1) + \sum_j W_{ji} \times \text{Activation}_j(t-1).$$

Decay was set at 0.01, and W_{ji} referred to the strength of the connection from node j to node i, for example, the strength of within-module negative connections, task lines, carrier lines, control lines, and automatic lines.

Next the input was transformed to an activation value with the activation function,

$$\text{Activation}_i(t) = \begin{cases} 1 & \text{if} \quad \text{Input}_i(t) \geq 1 \\ \text{Input}_i(t) & \text{if} \quad 0 < \text{Input}_i(t) < 1, \\ 0 & \text{if} \quad \text{Input}_i(t) \leq 0 \end{cases}$$

and then the activation value was sent to other connected nodes at the input and intermediate layers. Similarly, nodes at the intermediate layer integrated input, calculated and sent activation to other connected nodes at the intermediate and output layers. Next, nodes at the output layer integrated input and calculated their activation. Thus, the input and activation values for all nodes in the network were updated once, with the input and activation functions. Then, t was incremented by 1, and input and activation values were updated again for all the nodes in the network. This iterative process continued until a response was made. External stimulation was assumed to be clamped to the network (in Experiments 1 and 2, the stimulus was continuously displayed till the onset of a response) so that it continuously fed a value of either 1 or 0 to corresponding nodes at the input layer. As a result, the activation of some nodes increased while that in others decreased. Eventually, one node at the output layer was activated to the point where it crossed the response threshold, which was set at 0.95. At that point, we

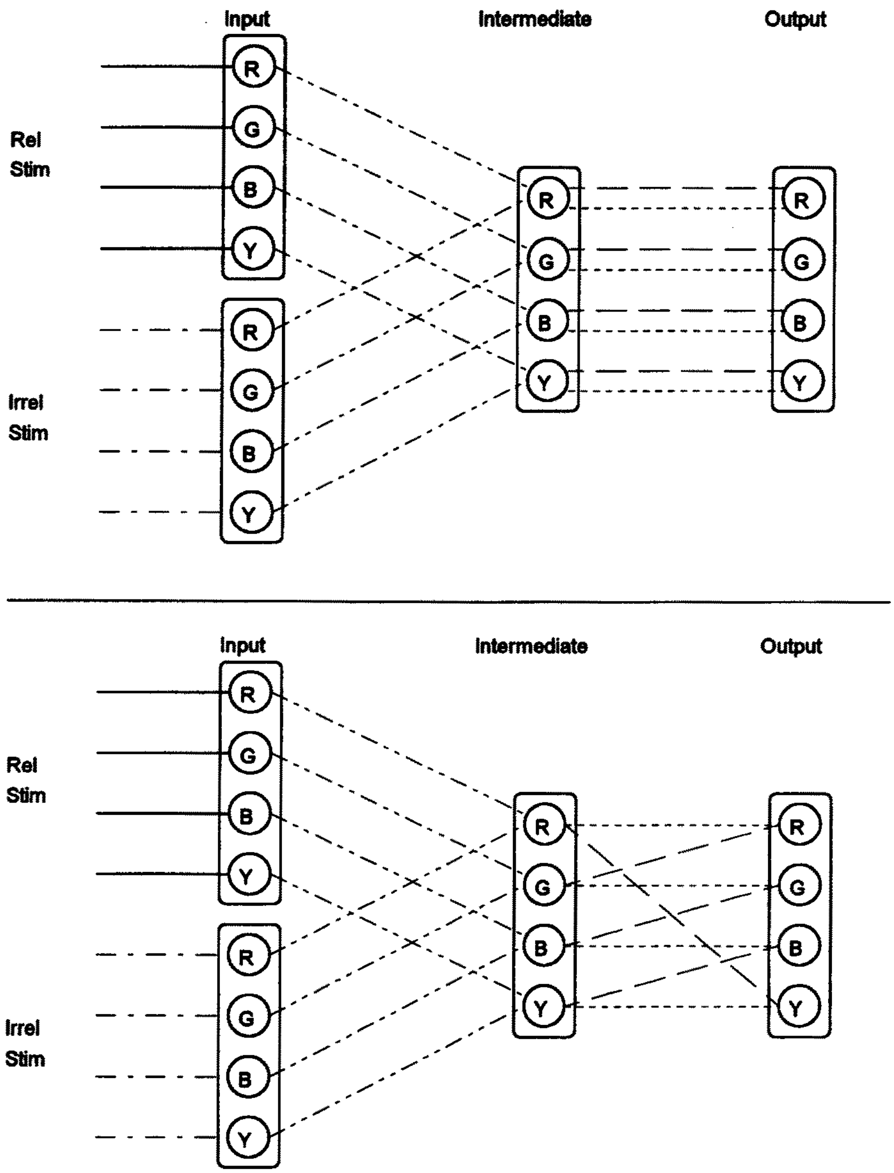

Figure 1. The interactive activation network for Stroop tasks; congruent stimulus–response (SR) mapping is on the top, and incongruent SR mapping is on the bottom. Modules are represented by large rectangles and nodes by small circles inside the rectangles. Each module was composed of four nodes (representing stimulus or response features such as red, green, blue, and yellow; R, G, B, and Y, respectively), which were negatively connected (connections omitted from this figure for ease of reading). The network was composed of two modules at the input layer and one module each at the intermediate and output layers. The modules were organized into two pathways, one for the relevant stimulus dimension (Rel Stim) and the other for the irrelevant stimulus dimension (Irrel Stim). Solid lines were the task lines for relevant stimuli, dash-dot lines were the task lines for irrelevant stimuli, dash-dot-dot lines were the carrier lines, dashed lines were the control lines, and dotted lines were the automatic lines.

considered that a response had been made and recorded the time and response. The time was converted to RT through a linear function, RT = $a + b \times t$, where t was the time taken in the simulation; the intercept, a, was the fixed amount of time that is supposed to be unrelated to the response production and decision, for example, the time taken for perception, motor programming and execution; and the slope, b, was a scale parameter. We set the intercept, a, at 270 ms (in Experiment 1) or 170 ms (in Experiment 2) and the slope, b, at 5. The network was then reset and awaited the next condition.

In the simulation the congruent mapping turned out to be

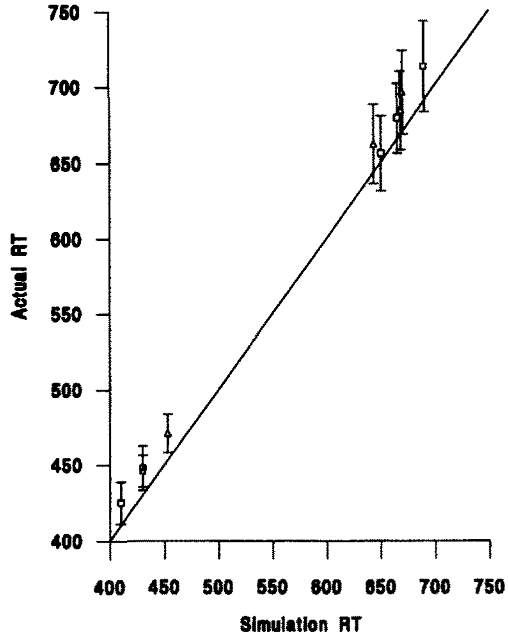

Figure 2. Experimental reaction times (RTs) from both Experiments 1 and 2 plotted against simulation RTs. The straight diagonal line indicates a perfect match between experimental and simulation RTs. Squares indicate Experiment 1 and triangles indicate Experiment 2. The error bars indicate ±1 standard errors (*N*s were 8 and 10 for Experiments 1 and 2, respectively).

much faster than the incongruent mapping. Within the congruent mapping, Condition A (SR^+/SS^+) was slightly faster than Condition B (SR^-/SS^-). Within the incongruent mapping, Condition C (SR^-/SS^+) was the fastest, Condition E (SR^-/SS^-) was the slowest, and Condition D (SR^+/SS^-) was in between. We compared the simulated RTs to the experimental results, as shown in Figure 2. If the model fit, all the points should fall in or around the diagonal line. Figure 2 clearly indicates that the interactive activation model predicted the Stroop results well. In fact, Pearson's r was .99 between the simulated and the experimental RT values. Thus, the simulation results clearly support a role for both SS and SR consistency and the importance of the interactive activation network in the Stroop tasks.

References

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* 332–361.

Cohen, J. D., Servan-Schreiber, D., & McClelland, J. L. (1992). A parallel distributed processing approach to automaticity. *American Journal of Psychology, 105,* 239–269.

Coles, M. G. H., Gratton, G., Bashore, T. R., Eriksen, C. W., & Donchin, E. (1985). A psychophysiological investigation of the continuous flow model of human information processing. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 529–553.

Damasio, A. R., & Damasio, H. (1992). Brain and language. *Scientific American, 267,* 88–95.

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

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

Dyer, F. N. (1974). Stroop interference with long preexposures of the word: Comparison of pure and mixed preexposure sequences. *Bulletin of the Psychonomic Society, 3,* 8–10.

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

Eriksen, C. W., & Eriksen, B. A. (1979). Target redundancy in visual search: Do repetitions of the target within the display impair processing? *Perception & Psychophysics, 26,* 195–205.

Fitts, P. M., & Deininger, R. L. (1954). S–R compatibility: Correspondence among paired elements within stimulus and response codes. *Journal of Experimental Psychology, 48,* 483–492.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 875–894.

Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General, 118,* 13–42.

Gratton, G., Coles, M. H., Sirevaag, E. J., Eriksen, C. W., & Donchin, E. (1988). Pre- and poststimulus activation of response channels: A psychophysiological analysis. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 331–344.

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

Hedge, A., & Marsh, N. W. A. (1975). The effects of irrelevant spatial correspondence on two-choice response time. *Acta Psychologica, 39,* 427–439.

Hintzman, D. L., Carre, F. A., Eskridge, V. L., Owens, A. M., Shaff, S. S., & Sparks, M. E. (1972). "Stroop" effect: Input or output phenomenon? *Journal of Experimental Psychology, 95,* 458–459.

Hock, H. S., & Egeth, H. (1970). Verbal interference with encoding in a perceptual classification task. *Journal of Experimental Psychology, 83,* 299–303.

Ishihara, S. (1991). *Ishihara's tests for colour-blindness* (concise ed.). Tokyo: Kanehara.

Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: Dilution of Stroop effects by color-irrelevant stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 497–509.

Kornblum, S. (1992). Dimensional overlap and dimensional relevance in stimulus–response and stimulus–stimulus compatibility. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (Vol. 2, pp. 743–777). Amsterdam: Elsevier.

Kornblum, S. (1994). The way irrelevant dimensions are processed depends on what they overlap with: The case of Stroop and Simon-like stimuli. *Psychological Research, 56,* 130–135.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus–response compatibility—A model and taxonomy. *Psychological Review, 97,* 253–270.

Kornblum, S., & Lee, J. W. (1995). Stimulus–response compatibility with relevant and irrelevant stimulus dimensions that do and do not overlap with the response. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 855–875.

Logan, G. D. (1978). Attention in character-classification tasks:

Evidence for the automaticity of component stages. *Journal of Experimental Psychology: General, 107,* 32–63.
Lu, C. H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review, 2,* 174–207.
MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.
MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 126–135.
McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review, 88,* 375–407.
Miller, J. (1991). The flanker compatibility effect as a function of visual angle, attentional focus, visual transients, and perceptual load: A search for boundary conditions. *Perception & Psychophysics, 49,* 270–288.
Phaf, R. H., van der Heijden, A. H. C., & Hudson, P. T. W. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology, 22,* 273–341.
Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.
Seymour, P. H. K. (1977). Conceptual encoding and locus of the Stroop effect. *Quarterly Journal of Experimental Psychology, 29,* 245–265.
Shaffer, W. O., & LaBerge, D. (1979). Automatic semantic processing of unattended words. *Journal of Verbal Learning and Verbal Behavior, 18,* 413–426.
Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84,* 127–190.
Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus–response compatibility: An integrated perspective* (pp. 163–180). Amsterdam: Elsevier.
Simon, J. R., & Berbaum, K. (1990). Effect of conflicting cues on information processing: The "Stroop effect" vs. the "Simon effect." *Acta Psychologica, 73,* 159–170.
Simon, J. R., & Sudalaimuthu, P. (1979). Effects of S–R mapping and response modality on performance in a Stroop task. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 176–187.
Stoffels, E. J., & van der Molen, M. W. (1988). Effects of visual and auditory noise on visual choice reaction time in a continuous-flow paradigm. *Perception & Psychophysics, 44,* 7–14.
Stoffels, E. J., van der Molen, M. W., & Keuss, P. J. (1989). An additive factors analysis of the effect(s) of location cues associated with auditory stimuli on stages of information processing. *Acta Psychologica, 70,* 161–197.
Stroop, J. R. (1992). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General, 121,* 15–23. (Original work published 1935)
Washburn, D. A. (1994). Stroop-like effects for monkeys and humans: Processing speed or strength of association? *Psychological Science, 5,* 375–379.
Zhang, H., Kornblum, S., & Zhang, J. (1995). Utilization of stimulus–response and stimulus–stimulus compatibility principles in machine designs. In K. Cox, J. Marsh, & B. Anderson (Eds.), *First international cognitive technology conference* (pp. 151–157). Hong Kong: City University of Hong Kong.
Zhang, H., Zhang, J., & Kornblum, S. (1997). *A parallel distributed processing model of stimulus–stimulus and stimulus–response compatibility.* Manuscript submitted for publication.

Received December 27, 1995
Revision received September 3, 1996
Accepted October 28, 1996 ■