

Toward a translational model of Stroop interference

ROBERT A. VIRZI and HOWARD E. EGETH

Johns Hopkins University, Baltimore, Maryland

Models of the Stroop effect that postulate single, centralized processing stations are fundamentally incorrect. A translational model is proposed in which parallel processing systems are hypothesized that are linked by a translation mechanism. Each system processes information using its own internal code specifically linked to the type of information the system processes. Information is passed between these systems via the translation mechanism. Evidence supporting this model comes from both the literature and the experiments reported in this article. Of primary importance is the demonstration that the pattern of interference that emerges is dependent upon the cognitive system used to process the information.

The Stroop paradigm was employed in the four experiments reported. In this paradigm, first reported by Stroop in 1935,¹ bidimensional stimuli are created such that either dimension of a stimulus may be responded to with the same overt response. Reaction time is measured of the response to one of the stimulus dimensions, while the subject attempts to ignore the other dimension, which will conflict on some trials with the correct response. For example, a subject who is shown the word RED printed in blue ink could respond either to word meaning by saying "red" or to ink color by saying "blue."²

The interesting finding is that, when subjects must read the word out loud, there is no effect of a conflicting ink color. Reaction time in the conflicting condition is as fast as in the condition in which the color word names the ink color for the trial. However, when the subject must name the ink color, reaction time in the conflicting condition is greatly increased, compared to the time required for the condition in which the color word and the ink color have the same value (e.g., RED in red).

Similar findings have been reported using different stimulus sets. For example, when word meaning and spatial position are combined (e.g., the words RIGHT or LEFT appear to the right or left of fixation), reading the word out loud is not affected by a conflicting spatial position, but naming the position the word occupies is slowed if a conflicting word appears in that position (Palef & Olson, 1975; Seymour, 1974).

The models proposed to account for the Stroop effect have differed in detail but have two underlying assumptions in common. The first is that, at some point in the

system, all the information passes through a single, centralized decision and/or response stage. This is true regardless of the particular task employed. To account for the asymmetry in the interference effects, the second assumption is made: The two dimensions are hypothesized to arrive at the decision stage at different times.

In Figure 1, a schematic of this type of model is shown for the ink color/color word form of the Stroop task. Two stimulus-analyzing stages are hypothesized: one to abstract word meaning and another to abstract ink color information (see Treisman, 1969). These stimulus analyzers work in parallel, and output from them races toward a central decision stage. Information about word meaning is assumed to arrive prior to information about ink color. As a result, when the subject must respond to word meaning, there is no interference from a conflicting ink color, because the word code is the only code active at the decision stage. When the subject must name the ink color of the stimulus, interference arises due to the prior arrival of the word code at the decision stage. Two codes are presumed to be active, and additional time is required to select the proper code for entry into the response buffer.

The account just provided of speed-of-processing models is necessarily devoid of details. It is meant to be an abstraction of the features common to models provided by Glaser and Dolt (1977), Klein (1964), Morton (1969), and Palef and Olson (1975), among others. The critical characteristics of these models are that all are based on the relative speeds of processing of the two stimulus dimensions and that a single, centralized response channel is hypothesized. Furthermore, speed-of-processing models are formulated so that the faster dimension can interfere with the slower dimension but not vice versa. For example, one would not expect a conflicting ink color to interfere with a response based on word meaning. The ink information is simply hypothesized to be processed too slowly.³

There are indications in the literature that speed-of-processing models are, at best, incomplete. Before dis-

This work was supported by NSF Grant BNS81-00842. We would like to thank Alfonso Caramazza and Michael McCloskey for helpful comments on earlier versions of the manuscript and Jordan Weinstein for his invaluable assistance in carrying out the mechanics of the studies reported in this article. R. Virzi is now a member of technical staff with Bell Communications Research, 445 South Street, Morristown, NJ 07960.

SPEED-OF-PROCESSING MODEL OF STROOP INTERFERENCE

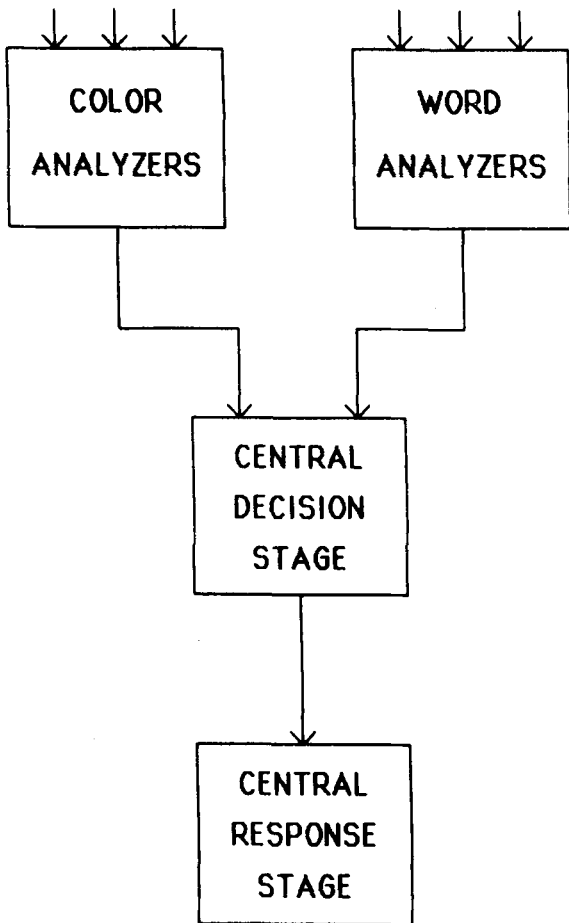


Figure 1. A generic representation of speed-of-processing models of Stroop interference is shown. Input from both stimulus analyzing stages is passed to a single, centralized decision stage. Output of the decision stage passes to a central response stage from which the overt response is initiated.

cussing the pertinent data, an outline of the model we are proposing will be provided.

AN OVERVIEW OF THE TRANSLATIONAL MODEL

The basic premise of the translational model is that there are several processing systems, each of which codes and processes information in a way that is specific to that system (see Palef, 1978). Information from one system may be translated into the code of another system (hence the name we have given this model). This translation process takes time. Furthermore, some of these systems are replete with an output or response stage suited to the type of information the system processes, so input to one system must undergo translation if it is to be processed within another system. In Figure 2, a generic representation of the translational model is shown.

Assume that two independent systems exist: a spatial system that processes the spatial relations between objects and a linguistic system that processes information in the form of linguistic codes. If the location of an object is to be vocally reported, then the information would have to be translated from the spatial system that performed the encoding into the linguistic system for output via a vocal response. If the location were to be indicated by moving a lever in a compatible direction, then no translation would need to occur because the spatial system would be capable of encoding the stimulus and planning and executing the response. (Note that to say a translation is not required is not the same as saying that no processing is required. We are not arguing that there is some sort of direct connection between eye and hand that mediates lever movement.)

The time-of-arrival hypothesis has been cogently criticized by Dunbar and MacLeod (1984). In this section we review a set of studies that speak to the hypothesized existence of a single, central decision and/or response stage underlying the Stroop effect.

In an auditory Stroop task, McClain (1983) has shown a clear reversal of the pattern of interference that can readily be interpreted in terms of the direction in which a translation process is assumed to run. She presented the words

GENERAL TRANSLATIONAL MODEL

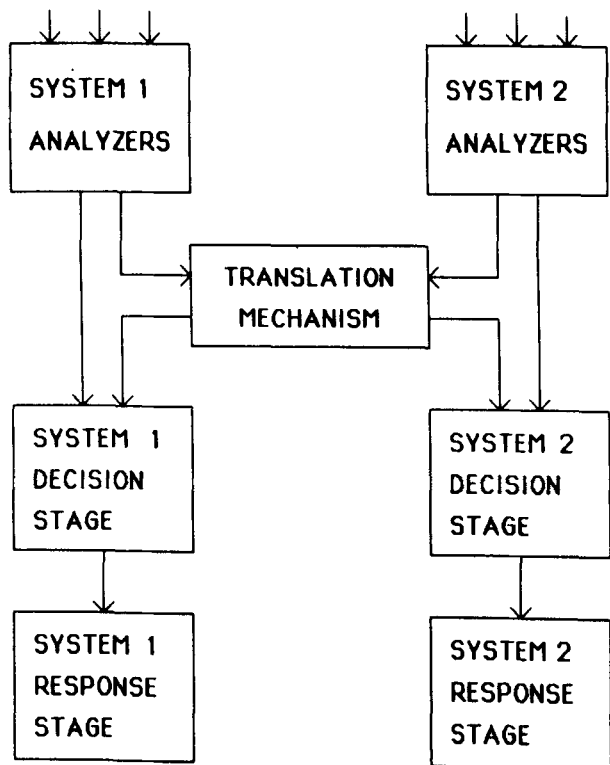


Figure 2. The framework for the translational model is presented. Generic systems are indicated that will be elaborated for the particular tasks as necessary.

HIGH and LOW in either a high or low pitch to her subjects. Four conditions were created by having subjects respond to either word meaning or pitch by saying "high" or "low" or by humming in a high or low pitch. In the four resulting conditions, the irrelevant dimension could be either compatible with or in conflict with the relevant dimension.

We can postulate two systems: a linguistic system that processes words and is linked to a vocal response stage and a tonal system for the processing of pitch that is linked to a response system for the output of pitch, in this case humming. Based on these systems, we can generate several predictions. When word meaning is to be output by the humming of a correspondingly high or low pitch, the word code must undergo translation into the form of representation used by the tonal system. Once in the tonal system, the code is subject to interference from conflicting information already in that system due to the irrelevant tone in which the word was presented. Interference will also arise when the subject must name the pitch of the stimulus. For this condition, the tonal information is translated into the linguistic system where it may be interfered with by a competing linguistic code.

When a subject responds to a word by saying it, or to pitch by humming, processing is contained within the linguistic and tonal systems, respectively. Interference from a conflicting irrelevant cue has no chance to occur because the irrelevant cue is processed by a different system and does not compete with the representation of the word at any point.

The results substantiate the predictions derived from the model. When a verbal response was required to the word or a humming response was required to the pitch, the value on the irrelevant dimension had no effect on reaction time. Conflicting trials were responded to as rapidly as compatible trials. When subjects were required to make a vocal response to the pitch of the stimulus or a humming response to the word, interference from the irrelevant dimension obtained. Conflicting trials in these two conditions took significantly longer than compatible trials. Thus, the relationship between the relevant dimension and the response mode and, by extension, the existence or absence of a translation, determined the conditions in which interference obtained. (Note that McClain interpreted her results in terms of a speed-of-processing model.)

Similar results were obtained by Flowers, Warner, and Polansky (1979) for the dimensions of numeric value (the digits 1, 2, and 3 or the words ONE, TWO, and THREE) and numerosity. They found a markedly different pattern of results that depended upon the mode of responding subjects were asked to use. In vocal-naming procedures and card-sorting tasks, a conflicting numeric value interfered with responses based on the numerosity value, but conflicting numerosity values did not interfere with sorting based on the numeric value.

When Flowers et al. changed the response mode to manual tapping or "vocal tapping" (e.g., saying duh-duh to indicate two "taps"), numeric value no longer interfered with numerosity-based judgments. However, a con-

flicting value on the numerosity dimension interfered with judgments of numeric value in this condition. This experiment illustrates the point that a change in response mode can drastically alter the pattern of results. The tapping response appears to be tied to numerosity-based judgments in much the same way that the vocal naming response is related to judgments of numeric value. They each have their own form of representation, and, when translation between codes must occur for output, interference from the irrelevant dimension can occur.

Additional evidence that interference effects are strongly dependent on translation processes comes from studies by Greenwald (1972), Pritchatt (1968), and Treisman and Fearnley (1969).

In summary, speed-of-processing models are insufficient, as currently outlined. Evidence suggesting the existence of several separate cognitive systems has been provided that is fatal to these simpler models predicated on a single, central decision and/or response stage. To account for these data, the translational model, incorporating the principle of independent cognitive systems, was postulated.

In Experiments 1 and 2 of this article, we set out to directly test predictions derived from the model using first the standard color word/ink color form of the task and then the locative/spatial position form of the task. Changes in the mode of responding are expected to reverse the direction of the interference effects. In the final two experiments, the translation mechanism was probed in an attempt to make as explicit as possible the conditions under which it operates.

EXPERIMENT 1

The translational model adequately accounts for results from conditions in which the response is initiated via the linguistic system, but so do simpler models based on the relative speeds of processing of word and color information (see Glaser & Dolt, 1977; Morton, 1969; Palef & Olson, 1975, for examples and Figure 1 for a schematic of this type of model). The problem was to devise a task that could be used to control the flow of processing so that the criterial dimension would arrive at the linguistic system's decision stage or the alternate system's decision stage, depending on the experimental condition. We could then examine whether or not this processing leads to the standard and/or reversed patterns of interference.

A card-sorting task was used to control the type of processing needed to complete the task. To mimic the standard Stroop situation, the bins into which the cards were placed were labeled with color words printed in black ink. We hypothesized that these labels would require processing to occur in the linguistic system's decision stage because the criterial dimension on the card would be matched to a linguistic cue. If this were the case, then the results should be the same as when subjects were asked to respond vocally: Responses to the meaning of the word should not be affected by a conflicting ink color,

but responses based on the color of the stimulus should suffer in the presence of a conflicting word cue.

To create a set of conditions in which processing would occur within the color system's decision stage, the bins into which the cards were to be sorted were labeled with patches of color. Rather than matching the criterial dimension on the card with a verbal label, the subject must match the criterial dimension on the card with a color label. We hypothesized that this manipulation would require processing to be carried out within the color system because the flow of information would be changed so that all codes must pass through the color system's decision stage. At the very least, linguistic processing would not be necessary at the response stage because the criterial dimension was not matched to a linguistic cue. This manipulation should result in a reversal in the patterns of interference: Responses to word meaning should suffer in the presence of a conflicting color cue, but responses to ink color should not be affected by a conflicting word cue.

This predicted reversal in the direction the interference takes is called the reversed Stroop effect. The translational model predicts this change in the pattern of interference when the way a subject responds is altered. Earlier models (e.g., Palef & Olson, 1975) do not call for the ink color to interfere with responses to the meaning of the word. Nor would they predict that responses to the ink color of a stimulus would be buffered from interference from a conflicting word, because information about word meaning is expected to reach the central decision stage prior to information about color. Thus the direction of interference is immutably fixed by assumptions about the speeds at which the two types of information are processed. The consequence of this assumption is that these earlier models predict no reversal in the pattern of interference.

Two groups of subjects participated in Experiment 1. For half the subjects the bins were labeled with color words printed in black ink, and for the other half the bins were labeled with rectangular patches of colored inks. The values of the color word labels and the ink color labels were the same and thus provided two alternative modes of responding. (For half the subjects, the labels above the bins were the words RED, BLUE, GREEN, and ORANGE printed in black ink; for the remaining subjects the labels were changed to patches of red, blue, green, and orange ink). Both groups of subjects performed five card-sorting tasks. The tasks differed with respect to the information on the cards or to the aspect of the stimulus to which the subject was asked to respond. We could then examine the speed with which the various sorts were performed using the two modes of responding and determine whether the translational model's predictions (i.e., a reversal in the pattern of interference) match these obtained reaction times.

Method

Subjects. Thirty subjects were recruited from the Johns Hopkins University subject pool and received either class participation credit or a nominal fee.

Stimuli. Four decks of 48 cards each were prepared in the following way using Cosco rubber stamp kit #R07 and Excelsior Dri-kwik stamp pads. Each card measured 7.5×12 cm. Centered in the middle of each card was a single stimulus, which varied in length from 2.5-5.0 cm and was approximately 0.7 cm in height.

The first deck was a control deck that consisted of the color words used in the experiment printed in black ink.⁴ These words were RED, BLUE, GREEN, and ORANGE. Each of the four words appeared equally often, one word per card. The second deck was made by placing a row of colored Xs on each card. The ink colors used for the Xs were named by the color words from the experiment. Twelve cards were prepared using each ink color with the following constraint: The number of Xs on each card varied from three to six to match the number of letters in the color words. As an example, consider the 12 cards that had rows of green Xs on them. Three cards each were prepared with three, four, five, and six Xs on them to match the length of the words RED, BLUE, GREEN, and ORANGE, respectively. Similar procedures were followed for the other three colors. The third deck was made up of conflicting color/word pairs (e.g., the word ORANGE appeared in green ink). Each of the four color words appeared four times in each of the three ink colors incompatible with the word. The final deck consisted of compatible color/word stimuli (e.g., the word RED would always appear in red ink). Twelve cards were prepared for each color word in the ink color named by the word.

Procedure. Subjects were seated at a table in front of four bins. The top of each bin was open and measured approximately 10×15 cm. The bins were deep (16 cm) so that, once a card was placed into the bin, the card dropped completely out of sight. For half the subjects, the bins were labeled with color words printed in black ink. For the remaining subjects, the labels were rectangular patches of the colored inks used in the experiment and were approximately 5 cm wide \times 12 cm long. The purpose of this manipulation was to permit the match between the criterial dimension and the bin label to be carried out by the linguistic and color systems, respectively. In either case, the labels were affixed directly over the bins so that they were clearly visible to the subject. (Note that each subject saw only labels that were either all color words or all color patches.)

The same five sorts were carried out by each subject. For the subjects who had the bins labeled with color words, the criterial dimension from the cards was to be matched with the word above the bin. For the subjects who had the bins labeled with color patches, the criterial dimension from the cards was to be matched to the color of the patch above the bins.

For the deck in which color words were printed in black ink, the criterial dimension was word meaning. The cards were matched to the words above the bins or to the corresponding color patches, depending on how the bins were labeled for the subject. For the rows of colored Xs, the criterial dimension was the ink color of the stimulus. Again, half the subjects matched the color on the card to word labels, and half matched the color on the card to color-patch labels.

Two types of sorts were performed by each subject for the conflicting color/word deck. Depending on which sort the subject was asked to perform, the deck was sorted with either ink color or word meaning as the criterial dimension. In either case, the irrelevant dimension provided a conflicting cue. The purpose of this manipulation was to examine the subject's ability to respond to the criterial dimension when the bins' labels were changed, and by extension when the system used to process the information must be changed.

The last deck of cards contained compatible color/word pairs. The subject was free to base the response on either the meaning of the word from the card or the ink color on the card. No conflicting cue was present, as in the decks having words in black ink and colored Xs. Unlike those decks, the irrelevant dimension was compatible with the relevant dimension. The question was whether subjects use this redundancy to facilitate response times.

Three orthogonal Latin squares were defined so that each row represented a unique ordering of the five types of sorts each subject performed. The sort order represented by each row was per-

formed by one subject when the bins were labeled with color words and by another subject when the bins were labeled with color patches. Any given type of sort was performed three times in succession, and the mean of these three replications was entered into the analyses. If the sort order for a given subject was A, B, C, D, E; the subject actually performed 15 sorts in the order A, A, A, B, B, B,

After each sort, the labels above the bins were randomly reordered to insure that subjects used the labels in performing the task. Also, to avoid biasing the subjects toward linguistic processing, a minimum of verbal instruction was given at the beginning of the experiment. After that, all instructions were given by manually demonstrating each sort, without verbal instruction. When the subject understood the task, he or she indicated this by nodding.

A digital stopwatch recorded reaction times to the nearest .1 sec. Timing began with the experimenter tapping the table at which the subject was seated, and timing terminated when the subject tapped the table after placing the last card into its bin.

Accuracy was emphasized in the instructions, and subjects made few uncorrected errors. These will not be discussed further.

Results

A one-way analysis of variance (ANOVA) with repeated measures was performed on the five sorts in which the bins were labeled with color words. A separate ANOVA was performed on the five sorts in which the bins were labeled with color patches.

Bins labeled with color words. The mean times to sort the decks of 48 cards are shown in Table 1. The ANOVA revealed a highly significant difference among the sorting times for the five conditions [$F(4,56) = 14.87, p < .001$]. A Tukey honestly significant difference (HSD) test was carried out to determine the causes of the differences between the means. The analysis confirms that sorting the conflicting color/word pairs based on ink color took significantly longer than all other sorts (HSD = 5.66, $p < .01$). The remaining sorts did not differ significantly from one another.

Bins labeled with color patches. The results for the sorts in which the bins were labeled with color patches also are shown in Table 1. The ANOVA showed a highly significant effect of the type of sort [$F(4,56) = 21.63, p < .001$]. A Tukey HSD test showed that sorting the conflicting color/word pairs on the basis of word meaning took significantly longer than the other four sorts

(HSD = 3.65, $p < .01$), which did not differ significantly from one another.

Discussion

Using the traditional Stroop color/word stimuli, Experiment 1 demonstrated that changes in the response mode can profoundly affect the direction of the interference effect. If processing of the criterial dimension is carried entirely by a single system, there is no effect of a conflicting cue encoded by the other system. It also was shown that interference effects can arise when the criterial dimension undergoes a translation. These results are consonant with the predictions of the translational model. Most notable are the results for those conditions in which the bins were labeled with patches of color. When word meaning was the criterial dimension, responses were impaired by a conflicting ink color. To complete the pattern needed for a reversed Stroop effect, the responses to the ink colors were buffered from interfering effects arising from a conflicting word.

In the analog of the standard Stroop conditions (viz., when the bin labels were color words), the usual pattern of interference obtained. Sorts based on the meaning of the word were unaffected by conflicting color cues, and sorts based on the ink color of the stimulus suffered in the presence of a conflicting word cue.

Even if one agrees that the present results are consistent with the predictions of the translational model, it is still pertinent to ask whether the results are consistent with single-channel speed-of-processing models. One way such models might be able to account for data demonstrating "translational" effects is by appeal to the concept of stimulus-response compatibility (McClain, 1983; see also Treisman & Fearnley, 1969). The argument is that stimulus-response incompatibility contributes to response competition. For example, in the ordinary type of Stroop task, one type of stimulus (color) must be responded to with an incompatible type of response (word). Stimulus-response incompatibility could exert its influence by retarding response availability. This influence would cause incompatible responses to tend to enter the single response channel after compatible responses.

It is important to note that in the present experiment there was clear evidence of Stroop interference in the absence of any hint of a compatibility effect. The lack of a compatibility effect is indicated by the fact that for both bins labeled with words and bins labeled with color patches, there was no difference between the two control decks. We conclude that stimulus-response incompatibility may play some role in Stroop-like tasks but that it does not provide a complete explanation of the phenomenon (see also McClain, 1983, p. 269).

Finally, the results demonstrate that compatible color/word pairs produce sorting times equivalent to those for unidimensional stimuli. For the word-labeled bins, sorting times were equivalent to those for words in black ink and, for the color-patch-labeled bins, sorting times were equal to those for rows of colored Xs. Thus the way

Table 1
Sorting Times (in Seconds) for Experiment 1

Sorting Condition	Bins Labeled With	
	Color Words	Color Patches
Color Words in Black Ink	64.1	62.6
Colored Xs	65.3	60.8
Conflicting Color/Word Pairs Sorted by Meaning	65.8	69.4
Conflicting Color/Word Pairs Sorted by Ink Color	75.1	62.6
Compatible Color/Word Pairs	64.8	61.3

is paved for using only compatible and conflicting stimulus pairings in subsequent experiments.

EXPERIMENT 2

Experiment 2 was performed with two goals in mind. One goal was to extend the model using a new set of stimuli and tasks. If the model only applied to the color/word form of the test, then its value would be greatly reduced. Another goal of the current study was to gain control over parameters of the stimuli that were not subject to experimental manipulation using a card-sorting task and the color/word stimuli.

The stimulus set was created by the factorial combination of two locatives, LEFT and RIGHT, with two spatial locations, to the left or to the right of fixation. Two of the resulting stimuli were compatible (LEFT to the left of fixation and RIGHT to the right of fixation), and two were conflicting (RIGHT to the left of fixation and LEFT to the right of fixation).

Our test of the translational model assumes the existence of two cognitive systems, one linguistic and one spatial. The stimuli are composed of a linguistic dimension (word meaning) and a spatial dimension (word location). The system used to encode word meaning has already been proposed in Experiment 1. In this study, however, the subject was allowed to make a vocal response as the output of the linguistic system. The subject either said "right" or "left" as a way of outputting the response via the linguistic system's response stage.

In this experiment, a spatial system is postulated instead of a color system. The first stage of this system abstracts information about the location in which the stimulus appeared. This information is coded in a form specifically suited to processing by the spatial system. Output of this position-analyzing stage is passed to the spatial-decision stage and, from there, to the spatial-response stage. There is evidence in the literature that the spatial system's code is particularly suited to a manual response (Greenwald, 1972). In the current study, this manual response is made by having the subject press a right- or left-hand button.

All the criteria for a test of the translational model have been met. The stimulus set is bidimensional, and an independent cognitive system has been hypothesized for the processing of each dimension. More to the point, perhaps, is that each system is replete with its own response stage particularly suited to the form of the code used by its system. The translational model predicts different outcomes depending upon which response mode the subject employs.

For half the subjects in Experiment 2, the criterial dimension was the meaning of the word. They responded either vocally or manually in different blocks of trials. This is unlike Experiment 1, in which the criterial dimension varied across conditions for each subject but the response mode was held constant. For the other half of the subjects, word position was the criterial dimension.

This group also responded both manually and vocally in blocked sets of trials. In all other respects, the logic of Experiment 2 follows that of Experiment 1. For the two conditions in which a translation is hypothesized (vocal responses to position and manual responses to word meaning), conflicting trials should take longer than compatible trials. For the two conditions in which processing of the relevant dimension is carried out from encoding to response by a single system (vocal responses to meaning and manual responses to position), conflicting trials should be as fast as compatible trials.

Method

Subjects. A total of 24 subjects were recruited as in Experiment 1.

Stimuli. Stimuli appeared on a cathode ray tube (CRT) under control of a Commodore PET computer. An opaque mask was placed in front of the CRT with a cutout that restricted viewing to a rectangle 3.6° in height and 10.7° in width at a viewing distance of 40 cm. A fixation point was always visible centered within this rectangle.

Stimuli were created by presenting the word RIGHT or LEFT either to the right or left of the fixation point. The words were aligned so that the nearest edge was 1.4° from the center of the field. Each word subtended 0.7° in height and averaged 2.3° in length.

Procedure. Twelve subjects responded to the meaning of the word, and 12 responded to the position of the word. Within each of these groups, half of the subjects responded manually for one block of 160 trials and responded vocally to another block of 160 trials. For the other half of the subjects in each group, vocal responses preceded manual responses.

A total of four stimuli were used in this study. Two represent compatible word/location pairings, and two represent conflicting word/location pairings. Equal numbers of each of the four stimuli appeared in each of the conditions. Thus, the 160 trials for each condition were composed of the word RIGHT to the right, LEFT to the left, LEFT to the right, and RIGHT to the left appearing 40 times each in a random order determined separately for each subject.

Each trial proceeded in the following manner. A warning tone sounded .5 sec before onset of the stimulus. The stimulus (a single word to the right or left of fixation) appeared, and the subject responded as quickly as possible. Reaction time in milliseconds was collected from the onset of the stimulus until the response was made. Manual responses were made by depressing a right- or left-hand button to indicate "right" or "left," respectively. Vocal responses closed a switch in a voice-activated relay. In either case, the response terminated the stimulus display and initiated feedback (the word CORRECT or INCORRECT appeared on the screen). After a 2-sec delay, the next trial began.

Only correct reaction times were collected. Reaction times shorter than 100 msec or longer than 1,500 msec were considered to be incorrect. For the correct trials, a mean was calculated for the compatible and conflicting trials on a subject-by-subject basis. These means were then entered into the statistical analyses.

Results

A three-way ANOVA was performed on the data. Critical dimension was a between-subjects factor; response mode and trial type were within-subjects factors. Table 2 presents the results of Experiment 2. All three main effects were significant. For criterial dimension, responses to the position were faster than responses to the meaning [$F(1,22) = 4.73, p < .05$]. For response mode, manual

Table 2
Results of Experiment 2.

Vocal		Manual	
Compatible	Conflicting	Compatible	Conflicting
Response to Meaning			
463	471	445	486
Response to Position			
475	513	329	337

Note—All reaction times are shown in milliseconds (msec).

responses were faster than vocal responses [$F(1,22) = 34.93, p < .001$]. For trial type, compatible trials yielded faster reaction times than conflicting trials [$F(1,22) = 89.68, p < .001$]. Of the two-way interactions, only criterial dimension \times response mode reached significance [$F(1,22) = 33.58, p < .001$]. Both vocal and manual responses to meaning and vocal responses to position produced equivalent reaction times. Manual responses to position were much faster than the other conditions (about 150 msec faster) when collapsed over the type of trial.

The three-way interaction was highly significant [$F(1,22) = 56.96, p < .001$]. Based on data shown in Table 2, it appears that the difference between compatible and conflicting trials was greater for the two conditions in which translation was hypothesized (manual responses to word meaning and vocal responses to position) than for the two conditions in which no translation was hypothesized (vocal responses to word meaning and manual responses to position). To determine if these interference effects were significant, four *t* tests were performed. The comparisons made were between compatible and conflicting trials for each combination of response mode \times criterial dimension.

For example, the compatible and conflicting means for the 12 subjects who participated in the vocal-response-to-meaning condition were compared using a *t* test. This mean difference (conflicting trials minus compatible trials) of 8 msec was not significant [$t(11) = 1.324, p > .10$]. Thus the conflicting trials were responded to as quickly as the compatible trials. Similar results obtained for the manual-response-to-position condition [$t(11) = 0.523, p > .10$]. The 8-msec difference was not statistically reliable. For the two conditions in which a translation was hypothesized, conflicting trials were significantly slower than compatible trials. When subjects responded manually to meaning, the mean difference was 42 msec [$t(11) = 1.814, p < .05$]. Vocal responses to the position of the word yielded a 38-msec difference between conflicting and compatible trials [$t(11) = 2.305, p < .025$].

Errors were infrequent in all conditions (averaging 2%) and positively correlated with reaction times. They will not be discussed further.

Discussion

The results of Experiment 2 were consistent with the translational model. When processing of the relevant

dimension was carried out from encoding to response by a single system, as in the manual-response-to-position and vocal-response-to-word-meaning conditions, there was no evidence that a conflicting cue on the irrelevant dimension led to interference. This lack of interference is indicated by the fact that there was virtually no difference between conflicting and compatible trials, either because the irrelevant code was not translated into the system used to respond, or because it arrived after the criterial code had been passed on to the response stage.⁵ When translation of the criterial dimension was required (e.g., when the linguistic code was recoded into the spatial format for manual output, or the converse), large and reliable interference effects emerged. It is hypothesized that this interference occurs at the decision stage of the system used to respond. Two codes are in competition: the translated code to which the subject must ultimately respond and the code for the irrelevant dimension that has arrived directly from the analyzing stage of the system used to respond. When these two codes indicate the same value, no problem arises. If the codes indicate opposite values, they interact with each other to slow the decision process, thus elevating the overall reaction time.

The results of both Experiments 1 and 2 are consonant with the translational model. In accounting for the results of these studies, three cognitive systems have been identified. The properties of these systems were described a priori and conformed with experimental results. Converging evidence for at least the linguistic and spatial systems is provided by Greenwald (1972). Using a dual-response task, he showed the independence of these two cognitive systems. Derks and Calder (1969), Treisman and Fearnley (1969), and Uleman and Reeves (1971) have shown that nonlinguistic responses to the ink color of a stimulus are unaffected by a conflicting color word. These results argue for the independence of the color system and against models in which all forms of input are channeled through a single, central stage (i.e., Glaser & Dolt, 1977; Morton, 1969). Such models cannot account for the effect that changes in response mode have on obtained results.

EXPERIMENT 3

In this study, we attempted to determine the fate of the irrelevant code for those tasks of Experiment 2 in which processing of the relevant dimension was contained within a single system. What happens to the word code, for example, when the subject responds manually to position? Two possibilities seem plausible. The irrelevant code may undergo translation into the system used to respond, but it may arrive at the decision stage after the response to the relevant code has been initiated. Alternatively, the irrelevant code may not be translated into the system used for processing of the criterial dimension. Processing of the irrelevant dimension may remain within the system used to encode it; the word code may not be translated into the spatial system at all.

There is some indirect evidence that the latter alternative may be correct. Consider the condition in Experiment 2 in which subjects responded vocally to word meaning. Mean reaction time in this condition was 467 msec, yet no interference from the spatial cue obtained, even though the spatial code could have undergone a good deal of processing by this time. Indeed, the total reaction time for manual responses to the position of the word was only 333 msec, indicating that the position information had to be available sometime before then. Given that the spatial code was processed more quickly, if it were obligatorily translated into the linguistic system, it would have produced an interference effect. It did not, so it is tempting to conclude that the translation did not occur. The assumptions underlying this argument are untested, however, and it is with this in mind that Experiment 3 was performed.

In Experiment 3, the two conditions that did not show interference effects in Experiment 2 were used; subjects performed a primary task of either manually responding to the position of the stimulus or vocally responding to its meaning. For those subjects in the former condition, a linguistic cue was preexposed at fixation that perfectly predicted what the word for the trial would be. This manipulation was intended to allow processing of the irrelevant linguistic dimension to begin prior to processing of the criterial spatial dimension. If the linguistic code were always translated but simply arrived at the spatial decision stage too late to cause interference in Experiment 2, then in the current study some preexposure of the linguistic cue should yield an interference effect due to the "head start" given to it.

Similar procedures were followed for the condition requiring vocal responses to meaning, except that a spatial cue was preexposed to allow more processing time for this irrelevant cue. If preexposure leads to interference with the linguistic task, we will have evidence that employment of the translation mechanism is not under subject control; its operation would be considered obligatory. The lack of interference from the comparable condition of Experiment 2 would be attributable to late arrival of the spatial cue at the linguistic decision stage. Should no interference effects occur at any preexposure, the experiment would support the hypothesis that translation occurs only when it is necessary in order for completion of the task (e.g., when a vocal response is to be made to the location of the word). In this sense, use of the translation mechanism would be termed flexible rather than obligatory.

There have been several studies reported in which the irrelevant dimension was previewed. For example, Dyer (1971) preexposed the word in a standard color-naming task. Interference decreased, but did not disappear, as preexposed time increased to 500 msec. Preexposure of an irrelevant dimension that produces no interference under ordinary (i.e., simultaneous onset) conditions is of greater relevance to the present experiment.

Glaser and Glaser (1982, Experiment 1) used standard color-word stimuli and preexposed color in a word-

reading task. Significant inhibition was not observed at any preview duration up to the maximum tested (400 msec), which suggests that the irrelevant code was not translated into the system used for processing the relevant dimension.

Method

Subjects. A new group of 24 subjects was recruited from the Johns Hopkins University subject pool as in earlier experiments.

Stimuli. The test stimuli used in the two primary tasks employed in this experiment were the same as those used in Experiment 2: The words RIGHT and LEFT appeared either to the right or left of fixation. Preceding exposure of the test stimuli was one of two types of cues.

When the primary task was to respond manually to the position, the preexposed stimulus was either the letter R or the letter L centered in the field of view.⁶ The preexposed letter always matched the initial letter of the test word. For example, the centered R was always followed by the word RIGHT to one side or the other side of fixation. This precue was visible for a variable length of time to allow processing of the irrelevant dimension (meaning) to begin before that of the relevant dimension (position).

For the conditions in which subjects responded vocally to the meaning of the word, a bar was preexposed to one side or the other of fixation. This precue was visible for a variable length of time, after which the test stimulus appeared. The location of the test stimulus was perfectly predicted by the bar, but each of the two words was equiprobable. Using this method, subjects could begin processing of the irrelevant spatial dimension before they could begin processing of the criterial (linguistic) dimension.

In Figure 3, the entire stimulus set for both tasks can be seen. In the top half, the stimuli for the manual response group are shown; in the bottom half, the stimuli used for the vocal response group are shown.

Experiment 3 Stimulus Set			
Manual Response to Position			
Compatible Trials		Conflicting Trials	
Pre-Cue	Test	Pre-Cue	Test
R	+ RIGHT	R	RIGHT +
L	LEFT +	L	+ LEFT
Vocal Response to Word Meaning			
Compatible Trials		Compatible Trials	
Pre-Cue	Test	Pre-Cue	Test
+ ----	+ RIGHT	+ ----	+ RIGHT
---- +	LEFT +	---- +	LEFT +

Figure 3. The complete stimulus set employed in Experiment 3 is shown. In the top half of the figure, the sequence of stimuli for the Man-P group are shown, and in the bottom half of the figure, the sequence of stimuli for the Voc-M group are shown.

Procedure. Two groups of 12 subjects each participated in this study. For one of these groups, the primary task was to respond manually to the position of the word (the Man-P group). The second group of subjects had the primary task of responding vocally to word meaning (the Voc-M group).

Man-P group. For this group of subjects, the letter R or L was preexposed at fixation. This letter remained in view for either 17, 50, 150, or 500 msec. In separate blocks of trials, the delay between onset of the letter and onset of the test stimulus (SOA) was varied. Ordering of these blocks for a given subject was determined using three orthogonal Latin squares. Reaction times were collected from onset of the test stimulus.

Voc-M group. Virtually the same procedures as used in the Man-P task were employed in the Voc-M task; however, instead of preexposing a letter at fixation, a bar marker was preexposed to either the right or left of fixation. In all other respects, the procedures followed for the Voc-M group were essentially the same as those for the Man-P group. Each subject participated in four blocks of 80 trials each, one block at each of the four preexposure durations. For further details of the experimental procedure, the reader may refer to Experiment 2.

Results

Two separate ANOVAs were performed on the data from Experiment 3, one on the Man-P group and another on the Voc-M group. Factors entering the analyses were trial type and stimulus onset asynchrony (SOA). The results from both conditions are presented in Table 3.

Man-P group. For this group of subjects, none of the effects reached significance: for SOA [$F(3,33) = 0.74$]; for trial type [$F(1,11) = 2.32$]; and for the interaction of SOA \times trial type [$F(3,33) = 0.17$, all $ps > .10$]. The lack of an interaction is of particular importance. If the linguistic cue were translated into the spatial system, we would expect to see an increasing effect of conflict as the SOA increased. This expectation is based on the assumptions that the translation times for the linguistic cue follow some distribution and that, as time passes, a greater proportion of these cues will reach the spatial decision stage in time to cause interference. (Errors were below 1% in all conditions and will not be discussed further.)

Voc-M group. The main effect of SOA and the interaction of SOA \times trial type failed to reach significance [$F(3,33) = 1.68$, $p > .10$, and $F(3,33) = 0.91$, $p > .10$, respectively]. The main effect of trial type was small but significant [$F(1,11) = 5.76$, $p < .05$]. Reaction times for compatible trials averaged 446 msec while those for conflicting trials averaged 451 msec. This result was mainly due to one aberrant subject as Table 4 shows.

Table 3
The Results for Both the Manual-P (Man-P) and
Vocal-M (Voc-M) Groups of Experiment 3

	Stimulus Onset Asynchrony (in msec)							
	17		50		150		500	
	Cp	Cf	Cp	Cf	Cp	Cf	Cp	Cf
Man-P Group	292	297	281	283	295	298	294	299
Voc-M Group	446	450	444	454	455	460	441	443

Note—Cp = compatible trials, Cf = conflicting trials.

Table 4
Compatible (Cp) and Conflicting (Cf) Trial Means are Shown
for the 12 Subjects from Experiment 3, Along with the
Difference Between the Two Trial Types

Trial Type	Subject											
	1	2	3	4	5	6	7	8	9	10	11	A
Cp	455	512	441	404	444	425	457	415	407	431	546	419
Cf	460	513	437	409	443	430	463	419	413	431	544	455
Cf-Cp	5	1	-4	5	-1	5	6	4	6	0	2	36

Note—The one aberrant subject is marked "A."

Perhaps of greater importance is the lack of an SOA \times trial type interaction. Given that the encoding and translation times for the spatial cue are distributed about a mean, we would expect an increase in the interference effect when the precue was in view for a longer period of time. This effect did not obtain. (Error rates in all conditions were below 2% and will not be discussed further.)

Discussion

The results of Experiment 3 tend to favor the hypothesis that translation of the irrelevant code into the system used for processing of the criterial dimension is not obligatory. For all the subjects in the Man-P group and for 11 out of the 12 subjects in the Voc-M group, interference effects were minimal or nonexistent, and for both groups, the interaction term was nonsignificant. This may be interpreted as indicating that the irrelevant code was not translated. It was to the subject's benefit to refrain from translating the precue because, on half the trials, it would have led to conflict and thus slowed reaction times.

It is not clear why the response pattern of one Voc-M subject was qualitatively different than the response of the other 23 subjects. We can only speculate that in the milieu of this experiment, this subject may have felt compelled to use the precue, given our efforts to make it available.

In general, the view that translation is not obligatory seems to be supported by the results. Even when processing of the irrelevant cue began 500 msec before processing of the criterial dimension, interference effects failed to materialize. It is unlikely that an effect failed to emerge due to an insufficient SOA between onset of the irrelevant and relevant dimensions. Had the subject been asked to translate and respond to the irrelevant dimension, he/she would have been able to do so within this 500-msec time span (see Experiment 2, manual response to meaning and vocal response to position conditions). One cannot argue, therefore, that we failed to provide ample time for the translation to occur.

One might wish to argue that the irrelevant dimension was not encoded at the stimulus-analyzing stage in this experiment or, for that matter, in the within-systems tasks of Experiment 2. We think this line of reasoning can be discounted on the grounds that it is unparsimonious. One would need to postulate a separate mechanism for encoding of the irrelevant dimension for the tasks in which the

critical dimension undergoes translation. Without such an addition, one could not account for the interference arising in these tasks.

In conclusion, the model best supported by the results is a translational model in which operation of the translation mechanism is not obligatory (i.e., when translation is not a prerequisite for completing the task, it will not occur).

EXPERIMENT 4

Experiment 3 demonstrated the nonobligatory nature of the translation mechanism. Suppose, however, that translation of an irrelevant dimension is required by the experimental task. Will it then interfere with performance in a within-system task, or will the within-system task remain insulated from interference? To answer this question, we created a situation in which the subject was able to use the precue by translating it into the system used to respond to the critical dimension. The new task required a secondary, unspeeded response to the precue within the system used to make the primary response (e.g., after a speeded vocal response to meaning, the subject was asked to name the position the word occupied). If the subject translated the precue into the system used to make the primary, speeded response, we would expect interference effects to become evident as SOA increased. The increase is expected because the extra time affords the precued code a chance to undergo translation and arrive at the decision stage used to make the primary response.

In a few previous investigations, subjects have been required or encouraged to translate an irrelevant code into a code presumably used by the system that responds to the relevant dimension. The most relevant of these studies would appear to be that by Glaser and Glaser (1982, Experiment 2). The irrelevant dimension was previewed, and SOA varied from 0-500 msec. In this experiment, they encouraged translation by making the predictive validity of the previewed dimension high (congruent stimuli occurred 80% of the time; control and incongruent stimuli each occurred 10% of the time). The comparison of the results with those of their Experiment 1, in which predictive validity was low, was quite striking. Interference effects were larger in their Experiment 2. Of particular interest is the fact that, in Experiment 2, when color was previewed, it did interfere with word reading.

Two groups of subjects participated in Experiment 4. One of these groups of subjects performed a primary task of responding manually to position. This response was speeded, and reaction time data were collected. For this group of subjects, the precue was a letter that predicted the word for the trial (see Experiment 3). After responding to the position, subjects were asked to press the button named by the word for the trial and, by extension, the precue for the trial. For example, the precue letter L might have been followed by the word LEFT to the right of fixation. The subject would press the righthand button as quickly as possible to respond to the position

and then press the lefthand button to respond to the meaning. The purpose of the secondary response was to create a situation that forced the subject to translate the irrelevant dimension into the system used to respond to the critical dimension, the spatial system in the example above. As the length of time the precue was in view increased, larger interference effects were expected, because, with the passage of time, more evidence about the word's meaning was hypothesized to be available at the spatial-decision stage. When word meaning conflicted with positional information, increased latencies should obtain.

Similar procedures were followed for the other condition used in Experiment 3, in which subjects responded vocally to meaning. For this group of subjects, however, the secondary, unspeeded response was to name the position the word occupied. The position was precued by a bar marker, as in Experiment 3. As for the previous group, the intention was to force translation of the irrelevant (spatial) information into the system used to respond to the critical dimension (the linguistic system).

Method

Subjects. Twenty-four subjects were recruited as in earlier experiments.

Stimuli. The stimuli employed in this study were the same as those used for Experiment 3. The entire stimulus set is shown in Figure 3.

Procedure. The procedures employed were the same as those for Experiment 3, with the following additions.

After the response to the relevant attribute was made, subjects were asked to indicate the value of the irrelevant dimension. (In this sense, it was not really irrelevant; however, we will refer to it as the irrelevant dimension to be consistent with Experiment 3 and because the value of the irrelevant dimension provided no information as to what the correct primary response was). The subjects were told that they did not have to make this second response quickly.

For 12 subjects, the critical dimension was the position of the word to which they responded manually (the Man-P group). This primary response was speeded. After the primary response, a secondary, unspeeded response to the word's meaning was also required. The secondary response was made manually, after the response to the position, and was perfectly correlated with the preexposed letter. For the other 12 subjects, the critical dimension was word meaning, to which they made a vocal response (the Voc-M group). A secondary response to location was also required. The secondary response was made vocally to force the subject to translate the spatial cue into the linguistic system. Recall that this response was perfectly predicted by the bar precue.

In all other respects, the procedures used were the same as those for Experiment 3. Reaction time was collected from onset of the word/location pair and terminated with the first response. Only trials in which both responses were correct were entered into the reaction-time analyses. Separate ANOVAs were performed on the error data.

Results

As in Experiment 3, separate ANOVAs were performed on the reaction times for the two groups of subjects. The error rates obtaining in this experiment were higher than in previous studies and tended to vary from condition to condition. As a result, it was deemed appropriate to also subject the error data to analysis. Factors entering all of these analyses were SOA (length of

time precue was visible: 17, 50, 150, or 500 msec) and trial type (compatible vs. conflicting).

Man-P primary task. Subjects in this group responded to the position of the word manually and then made a manual response to meaning. In Figure 4, Panel A, reaction time data for the primary response are shown. Increasing the length of time the letter precue was in view produced no change in reaction times: The main effect of SOA was insignificant [$F(3,33) = 1.88$]. Trial type was significant; compatible trials were generally responded to faster than conflicting trials [$F(1,11) = 68.80$, $p < .001$]. More important, however, is the significant interaction of SOA \times trial type [$F(3,33) = 4.12$, $p < .025$]. At longer SOAs, the difference between compatible and conflicting trials appears to be greater.

To determine exactly where the differences lay, a Tukey HSD test was performed on the SOA \times trial-type cells, resulting in eight means being subjected to all pairwise comparisons. For each of the four levels of SOA, there were two levels of trial type. In the following discussion, we will focus only on the most relevant comparisons.

There were no differences among any of the compatible trial means across SOAs. For all SOAs, the compatible trials were faster than the conflicting trials ($p < .01$). Among the conflicting trials, SOAs of 17 msec and 50 msec differed from those of 150 msec and 500 msec with $p < .05$. Differences within these two groups were not significant. In Figure 4, Panel B, error percentages for compatible and conflicting trials are shown as a function of precue duration. An ANOVA performed on these data again showed a significant interaction of SOA \times trial type [$F(3,33) = 3.50$, $p < .025$], indicating that the proportion of errors grew more rapidly for conflicting trials than compatible trials as the SOA was increased. Unlike the reaction time data, the main effect of SOA was significant [$F(3,33) = 23.27$, $p < .001$], but the effect of trial type was not [$F(1,11) = 1.91$]. These error data were subjected to a Tukey HSD test in the same manner as the reaction time data were. Of the four comparisons between compatible and conflicting trial means for a given SOA, only one reached significance. At the 500-msec SOA there were more errors on conflicting trials than on compatible trials ($p < .01$). This result is consistent with the translational model, because interference was expected to increase as the SOA increased.

Voc-M primary task. Subjects in this group responded vocally to the word's meaning and then made a vocal response to position. In Figure 4, Panel C, the reaction time data for the primary response are shown. Both main effects were significant [SOA, $F(3,33) = 6.66$, $p < .005$, and for trial type, $F(1,11) = 28.21$, $p < .001$]. The interaction of these two terms was also significant [$F(3,33) = 4.28$, $p < .01$], showing a tendency for the difference between compatible and conflicting trials to increase as the bar marker precue was in view for longer periods of time.

As in the preceding condition, a Tukey HSD test was performed on the data. At the 17-msec and 50-msec SOAs,

the difference between compatible and conflicting trials was not significant. At the 150-msec and 500-msec SOAs, the difference between compatible and conflicting trials was significant (both with $p < .01$). This difference is evidence for a buildup of interference at the longer SOAs. In addition, compatible trials at the 17-msec SOA differed from compatible trials at the 50-msec and 150-msec SOAs ($p < .01$ and $p < .05$, respectively). In Figure 4, Panel D, the percentage of errors for compatible and conflicting trials are shown as a function of SOA. Errors tended to vary with SOA [$F(3,33) = 21.06$, $p < .001$], and more errors were made on conflicting trials than on compatible trials [trial type, $F(1,11) = 8.81$, $p < .025$]. The interaction of SOA by trial type was marginal [$F(3,33) = 2.88$, with p approximately .05]. Once again, an HSD test was applied to the data. For the critical comparisons of compatible and conflicting trials within a given SOA, there were no differences at the 17-msec and 50-msec SOAs. At the 150-msec level, this difference was significant beyond the .05 level and at the 500-msec SOA, the difference between compatible and conflicting trials was significant ($p < .01$), reflecting a general trend toward greater interference at longer preexposure durations.

Discussion

The results of this experiment were generally consistent with the translational model. As the length of time the precue was in view was increased, the amount of interference generated by a conflicting irrelevant cue also increased. It appears that processing and translation of the irrelevant dimension begins with onset of the precue. Recall that the precue is a perfect predictor of what the secondary response for the trial should be. As this precue remained in view for longer periods of time, its effects on processing of the criterial dimension increased.

There are both similarities and differences between our results and those of Glaser and Glaser (1982, Experiment 2). Similarity is evident in the comparison between conditions in which subjects read words. In our case, subjects responded vocally to locatives (the Voc-M condition); position was precued. In the Glaser and Glaser study, subjects read aloud the names of color words; color was precued. In both experiments, as SOA varied from 0-500 msec, the overall Stroop effect (i.e., the difference between conflicting and compatible reaction time) increased monotonically from approximately 0 msec to over 100 msec. (In our case, the estimate of near 0 difference for 0 SOA is taken from Experiment 3, as the shortest SOA used in Experiment 4 was 17 msec.)

A clear difference between our findings and those of Glaser and Glaser (1982, Experiment 2) is evident in the comparison between our condition in which subjects responded manually to position with the word precued (the Man-P condition), and their condition in which subjects named the colors of color words with the word precued. We found that as the duration of SOA increased, so did the difference between conflicting and compatible reaction times, just as was the case in the Voc-M condition.

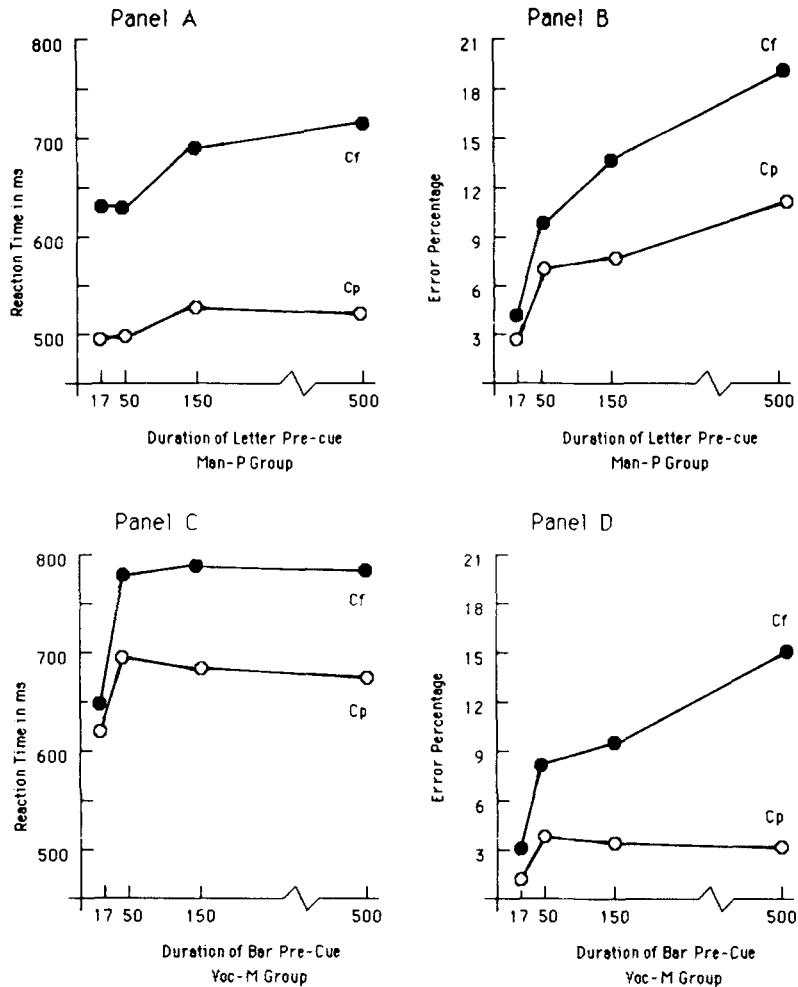


Figure 4. Results for the primary task of both conditions employed in Experiment 4 are shown. Panels A and B show reaction time and error data for the Man-P group, respectively, and Panels C and D show reaction time and error data for the Voc-M group, respectively.

However, Glaser and Glaser's data show that the overall difference between conflicting and compatible reaction times actually decreased as SOA increased in their color-naming task. There are, of course, many differences between color naming and manual response to position; it is not yet clear why manipulating SOA has such different effects on these two tasks. It seems clear that this will be worth further investigation.

The increase in the Stroop effect with increasing SOA found in Experiment 4 could result in either of two ways, depending on whether one posits a discrete stage or continuous flow type of model.⁷ If one assumes a discrete stage model, then increasing the SOA would have the effect of increasing the proportion of trials in which the irrelevant dimension reached the decision stage prior to the critical code. This argument is based on some mean time for encoding and translation of the irrelevant cue with variance about that mean. On any given trial, the irrelevant code either will or will not be available at the decision

stage used to process the critical dimension. The probability of the irrelevant code being available to create conflict will increase as the SOA increases and, as a result, the average amount of interference over all the trials in the condition will be higher. This increase is the product of a greater proportion of trials in which the irrelevant cue has time to reach the decision stage before the relevant code.

Alternatively, in a continuous flow model (Eriksen & Schultz, 1979; McClelland, 1979), processing at one stage is not necessarily complete before partial information is passed on to subsequent stages. If we relate this to the task employed in the current study, we arrive at the conclusion that processing and translation of the irrelevant pre-cue begins with its onset, just as in the discrete stage view. However, the mechanism leading to increased interference at longer SOAs is different. Rather than alterations in the proportion of trials in which the irrelevant code arrives prior to the critical code, a con-

tinuous flow model would vary the amount of evidence arriving at the decision stage over time. As the precue is in view for longer periods of time, conflicting trials would be hypothesized to exert greater effect by virtue of the salience of the irrelevant code at the decision stage involved in the processing. As more evidence contrary to the proper response accumulates at the longer SOAs, the overall reaction time on each trial should be slowed by some amount. The current study was not designed to distinguish between these two alternatives, but future research might profitably be focused on this distinction. Methods for distinguishing between continuous-flow and discrete-stage models have been discussed by Meyer, Yantis, Osman, and Smith (1984) and by Miller (1982). Also relevant here are Schweikert's (1978, 1983) efforts to model processes organized in networks. (Schweikert has analyzed Stroop-like tasks in which subjects make speeded reactions to both dimensions of the stimulus. His analysis is provocative, but it is not clear that his conclusions are directly relevant to standard Stroop tasks.)

One unexpected finding in our Experiment 4 was the large reaction time difference between compatible and conflicting trials at all SOAs for the Man-P group. On the basis of previous results, the difference at the shorter SOAs was expected to be closer to zero. The effect may have arisen from the subjects' strong tendency to try to group the two responses despite instructions that it was unnecessary to do so. (Voc-M subjects indicated they had no trouble keeping the vocal responses separated temporally, and a similar effect failed to obtain in the Voc-M condition.) Both Man-P responses may have been facilitated when the same overt response was required for each and slowed when the primary response differed from the secondary response. Still, superimposed on this overall tendency for compatible trials to be responded to faster than conflicting trials was an increased interference effect at longer SOAs, as is consistent with the translational model.

The other puzzling finding was the comparative speed of vocal responses in the 17-msec SOA. Concurrent with faster reaction times in this condition were lower error rates, eliminating a speed-accuracy trade-off explanation. Interestingly, a similar performance gain is shown in the error data for the Man-P group at the 17-msec preexposure duration. We are unable to interpret this finding and are interested in seeing whether it can be replicated.

It is also worth mentioning that reaction times in the primary tasks were not as fast as they were for the comparable conditions of Experiment 3. Even if we consider only the compatible trials (as there were no predictions regarding a change for these trials), we see that there is approximately a 200-msec increase. We suspect that this increase resulted from an overall increase in the complexity of the task required of subjects. It is simply harder to plan and execute two things than one. It is not a damaging result and is concordant with the increase in error rates for this experiment. The critical finding is that the difference between compatible and conflicting trials was larger at the longer SOAs.

Finally, it seems prudent to point to some shortcomings in the designs of Experiments 3 and 4. First, in Experiment 3, we only tested conditions that did not yield interferences in Experiment 2; precuing failed to produce any interference. It is possible that precuing itself eliminated the Stroop effect, and that, if we had included the conditions that yielded interference in Experiment 2, then even those conditions would have failed to show interference when the precue was used in Experiment 3. Although conceivable, this possibility seems unlikely considering the fact that precuing of the irrelevant dimension has not eliminated interference in ordinary Stroop tasks (e.g., Dyer, 1971; Glaser & Glaser, 1982).

A second problem is that SOA was confounded with the exposure duration of the precue. Consequently, it is possible that at short SOAs, subjects may have been unable to identify the prime. This confound makes it all the more impressive that interference did not vary with SOA in Experiment 3. In Experiment 4, this problem is more serious because the interesting finding is an increase in interference with SOA. In particular, it might seem that at the 17-msec exposure/SOA, it would be difficult to identify the prime. However, the interference effect increases as the exposure/SOA increases beyond 50 msec, even for the bar precue which ought to be easy to identify. Both experiments also lacked a condition in which an uninformative warning signal was used at each SOA to control for preparedness. Although this may have been worthwhile, it does not seem crucial in the present context, because the variable of chief interest is essentially a difference score between conflicting and compatible reaction times at each SOA, and both conflicting and compatible trials would probably be affected similarly by preparedness. Indeed, the fact that the focus of interest is on the change in the difference in reaction time between conflicting and compatible trials as a function of SOA helps allay other concerns as well (such as the possibility that the observed effect is due not to translation but to the memory or response requirements associated with the dual-response task used in Experiment 4).

GENERAL DISCUSSION

The main conclusion reached is that the translational model provides a useful model of dual attribute processing. In reviewing the literature, shortcomings of simple speed-of-processing accounts of the Stroop phenomenon were pointed out. The most crucial evidence supporting the translational model comes from studies in which changes in the response mode reverse the direction of interference effects. The reversed-Stroop effect was clearly demonstrated for the standard color/word form of the task in Experiment 1 and for the location/word form of the task in Experiment 2.

One might argue that, in Experiment 1, the response mode was not altered because subjects, regardless of what condition they were in, responded by placing cards in bins. This is true, but misses the point of the experiment. The bin-label manipulation was introduced to provide

provide what nature has neglected to give to humans: a means of outputting a color code without recourse to linguistic processing. In Experiment 1, the color patch labels provided an external color code to which the internal color code could be matched. Linguistic processing was not needed to perform the task, and so processing could be contained within the color system.

The results were consistent with this hypothesis. Color-to-color matches were not slowed in the presence of a conflicting word. Matching a word to a color patch was shown to be adversely affected by a conflicting ink color. The mechanism underlying this interference was hypothesized to be competition between the translated word code and the ink code at the color-decision stage. To reduce this result to the simplest terms, the subject had difficulty deciding which of the two color codes that were active should be matched to the color patch above the bin.

The precise way to resolve the conflict arising in tasks that require a translation has been left unspecified. Klein (1964) advanced the restimulation hypothesis that may be interpreted as an active check on the source of the codes. Consider a two-step decision mechanism in which the first step can rapidly identify the codes that are active at the decision stage. For compatible trials, these two codes will indicate the same value, so the first step may pass either code, and the response will be correct. If the two codes indicate different values, as they do on conflicting trials, the second stage of the decision stage would be used to determine the source of the two codes. Through this second, time-consuming process, the correct code will be chosen, but only at the expense of added time and effort. This explanation accounts for the difference in time required for compatible and conflicting trials.

While this is a plausible mechanism, it has by no means been firmly established. As mentioned in the discussion of Experiment 4, one may postulate a continuous flow model in which the mechanism for conflict resolution is based on the accumulation of evidence for one response or another over time. As an example, take the case in which the task is to respond manually to word meaning, and the stimulus word *RIGHT* appears to the left of fixation. Information from both dimensions of the stimulus is hypothesized to be accumulating over time at the spatial decision stage. Eventually, a threshold criterion for one of the values will be exceeded, and that response will be initiated. However, the two types of input are presumed to inhibit each other, resulting in a longer sampling time than for a compatible stimulus. For this latter case, both dimensions of the stimulus will indicate the same value and will not compete with one another. The type of mechanism has been modeled as a random walk process (Pachella, 1974; see Logan, 1980, and Logan & Zbrodoff, 1979, for comments relating to random walk processes in the Stroop task). Evidence for and against the possible responses accumulates until one of the response thresholds is exceeded, thus initiating that response. Either of these mechanisms could be incorporated into the translational model.

To explain the results of the color/word form of the task (Experiment 1) and the position/word form of the task (Experiments 2, 3, and 4), three cognitive systems were implicated. For McClain's (1983) task involving auditorily presented words that varied in pitch, we must include a tone-processing system to enable the model to account for her findings. Finally, the Flowers, Warner, and Polansky (1979) results, showing reversed interference effects for numerosity and digit-identity-based judgments when subjects changed from vocal naming to a tapping response, suggests a fifth system for encoding and processing numerosity. The judgments based on digit identity could be carried out within the linguistic system. The digit "6," for example, does not intrinsically represent six items any more than the word *RED* represents the color red.

How many systems are needed to model all the results found in the literature? This question is open to debate, but one thing is clear: As the number of systems needed to account for the obtained results grow, the attractiveness and elegance of the translational model is diminished, unless converging evidence for the existence of the hypothesized systems can be found (see also Wickens, 1984). As it stands now, the model is largely a redescription of the results. Separate systems are assumed when a change in response mode leads to a change from an interference effect to no interference effect (or vice versa). It would clearly be an improvement if the plausibility of the hypothesized system could be evaluated with respect to other sorts of evidence. Some steps in this direction have been taken by Posner, Henik, and McLeod (1982) and Kinsbourne and Hicks (1978). These researchers have tried to link cognitive systems to available evidence about the neural substrate used to implement these systems.

In Experiments 3 and 4, the focus of the research was changed. Confirmation of the translational model was not the primary goal: An in-depth examination of the translational mechanism was attempted. Experiment 3 demonstrated that translation of the irrelevant dimension is not obligatory. Even when processing of this dimension began 500 msec before processing of the criterial dimension, there was no evidence of it being recoded into the system used to respond. Compatible and conflicting trial types were responded to with equal speed.

Experiment 4 demonstrated that when the irrelevant code must be translated due to task requirements, this code will interact with the code from the criterial dimension. Increased interference was evidenced as processing of the precue was allowed to proceed for longer periods of time before onset of the criterial dimension. Two mechanisms for this buildup of interference were discussed. In a discrete stage view, the proportion of trials in which interference occurs increases with time, and in a continuous flow view, the amount of interference on each trial increases as evidence about the irrelevant dimension accumulates.

To conclude, the translational model is favored for the depth and breadth of the experimental results it can account for. Those models, based on entry of codes into

a single, centralized channel, are unable to predict changes in the patterns of interference that result from changes in response mode. Examination of the specifics of the translational mechanism was carried out in Experiments 3 and 4. It was concluded that translation occurs only when it is required by the demands of the task. Finally, suggestions were made regarding the method future research might take to examine the mechanism for the resolution of conflict.

REFERENCES

- DERKS, P. L., & CALDER, E. S. (1969, April). *Information processing and verbal labels: The Stroop color-word test*. Paper presented at the meeting of the Eastern Psychological Association, Philadelphia, PA.
- DUNBAR, K., & MACLEOD, C. M. (1984). A horse race of a different color: Stroop interference patterns with transformed words. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 622-639.
- DYER, F. N. (1971). The duration of word meaning responses: Stroop interference for different preexposures of the word. *Psychonomic Science*, **25**, 229-231.
- ERIKSEN, C. W., & SCHULTZ, D. W. (1979). Information processing in visual search: Some theoretical considerations and experimental results. *Perception & Psychophysics*, **25**, 249-263.
- FLOWERS, J. H., WARNER, J. L., & POLANSKY, M. L. (1979). Response and encoding factors in "ignoring" irrelevant information. *Memory & Cognition*, **7**, 86-94.
- GLASER, M. O., & GLASER, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 875-894.
- GLASER, W. R., & DOLT, M. O. (1977). A functional model to localize the conflict underlying the Stroop phenomenon. *Psychology Research*, **39**, 287-310.
- GREENWALD, A. G. (1972). On doing two things at once: Timesharing as a function of ideo-motor compatibility. *Journal of Experimental Psychology*, **25**, 52-57.
- JAENSCH, E. R. (1929). *Grundformen menschlichen Seins. Mit Berücksichtigung ihrer Beziehungen zu Biologie und Medizin, zu Kulturphilosophie und Pädagogik*. Berlin: Otto Elsner.
- KINSBOURNE, M., & HICKS, R. E. (1978). Functional cerebral space: A model for overflow, transfer and interference effects in human performance: A tutorial review. In J. Requin (Ed.), *Attention and performance VII* (pp. 345-362). Hillsdale, NJ: Erlbaum.
- KLEIN, G. S. (1964). Semantic power measured through the interference of words with color-naming. *American Journal of Psychology*, **77**, 576-588.
- LOGAN, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, **12**, 523-553.
- LOGAN, G. D., & ZBRODOFF, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition*, **7**, 166-174.
- MCCLAINE, L. (1983). Stimulus-response compatibility affects auditory Stroop interference. *Perception & Psychophysics*, **33**, 266-270.
- MCCLELLAND, J. (1979). On the relations of mental processes: An examination of systems of processes of cascade. *Psychological Review*, **86**, 287-330.
- MEYER, D. E., YANTIS, S., OSMAN, A., & SMITH, J. E. K. (1984). Discrete versus continuous models of response preparation: A reaction time analysis. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes* (pp. 69-94). Hillsdale, NJ: Erlbaum.
- MILLER, J. (1982). Discrete versus continuous stage models of human information processing: In search of partial output. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 273-296.
- MORTON, J. (1969). Categories of interference: Verbal mediation and conflict in card sorting. *British Journal of Psychology*, **60**, 329-346.
- PACHELLA, R. G. (1974). The interpretation of reaction time in information processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, NJ: Erlbaum.
- PALEF, S. (1978). Judging pictorial and linguistic aspects of space. *Memory & Cognition*, **6**, 70-75.
- PALEF, S. R., & OLSON, D. R. (1975). Spatial and verbal rivalry in a Stroop-like task. *Canadian Journal of Psychology*, **29**, 201-209.
- POSNER, M. I., HENIK, A., & MCLEOD, P. (1982, November). *Patterns of cross-talk between sensory, cognitive and response systems*. Paper presented at the meeting of the Psychonomic Society, Minneapolis, MN.
- PRITCHATT, D. (1968). An investigation into some of the verbal processes that underly associative verbal processes of the Stroop color effect. *Quarterly Journal of Experimental Psychology*, **20**, 351-359.
- SCHWEIKERT, R. (1978). A critical path generalization of the additional factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology*, **18**, 105-139.
- SCHWEIKERT, R. (1983). Latest network theory: Scheduling of processes in sentence verification and the Stroop effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 353-383.
- SEYMOUR, P. H. K. (1974). Stroop interference with response, comparison and encoding stages in a sentence picture comparison task. *Memory & Cognition*, **2**, 19-26.
- STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.
- TAYLOR, D. A. (1977). Time course of context effects. *Journal of Experimental Psychology: General*, **106**, 404-426.
- TREISMAN, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, **76**, 282-289.
- TREISMAN, A. M., & FEARNLEY, S. (1969). The Stroop test: Selective attention to colors and words. *Nature*, **222**, 437-439.
- ULEMAN, J. S., & REEVES, J. (1971). A reversal of the Stroop interference effect, through scanning. *Perception & Psychophysics*, **9**, 293-295.
- WICKENS, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63-102). Orlando, FL: Academic Press.

NOTES

1. The task was actually reported by E. R. Jaensch in 1929. Stroop was the first to report its use in this country.
2. Throughout this paper, capitals will be used to indicate stimulus words and lowercase will be used to indicate other dimensions of the stimulus. Thus, LEFT to right is unambiguously the word LEFT appearing to the right of fixation.
3. To the extent that there is overlap in the distribution of arrival times, a stimulus dimension which is processed relatively slowly could interfere with the dimension that is processed relatively quickly. In the present studies, the slower dimensions do not interfere with the faster ones, so we shall ignore the overlap issue.
4. One might wish to argue that black is an ink color, and resultantly the control deck does not really differ from the other decks used in the experiment. This may be so, but the important distinction between this deck and other decks is that the ink color information is invariant over the deck of cards. It therefore represents a null value. The same comments hold for the rows of colored Xs, to be discussed shortly. What little linguistic information is represented by the Xs is held constant in the deck, and it too represents a null value.
5. The data tend to support the hypothesis that the irrelevant code is not translated, at least when the subject responds vocally to the meaning of the word. Position information is most likely abstracted before word meaning. This is evidenced by the subject's ability to respond almost 150 msec faster to the position using a manual response than any of the other conditions. Because position information is available first, if it were automatically translated into the linguistic system, it should interfere with responses to the word's meaning. Because it does not, we have indirect evidence that the translation is not obligatory. Experiment 3 tested this hypothesis directly.
6. We would have preferred to use the words LEFT and RIGHT;

however, because LEFT has an even number of letters, it could not be precisely centered on our CRT display. It would have extended one letter too far to either the left or the right of fixation. We feared that this would create a confounding spatial incompatibility effect.

7. The following argument is concerned chiefly with explaining the increase in Stroop interference with increasing SOA (i.e., with preview of irrelevant information). Strictly speaking, Stroop interference refers to the difference between conflicting and neutral trials, not to the differ-

ence between conflicting and compatible trials, as the latter differences may include both facilitatory as well as inhibitory effects (cf. Taylor, 1977; Glaser & Glaser, 1982). To the extent that this is the case, a separate explanation is required for facilitatory effects; these are ignored herein.

(Manuscript received May 21, 1984;
revision accepted for publication April 5, 1985.)