Perception & Psychophysics
2003, 65 (3), 367-378

Translation and competition among internal representations in a reverse Stroop effect

FRANK H. DURGIN
Swarthmore College, Swarthmore, Pennsylvania

In two experiments, the contributions of internal and external competition in a task known to produce a reverse Stroop effect were investigated. In this paradigm, the verbal meaning of an incongruent Stroop stimulus is identified by pointing to a patch of matching color — a task made difficult by the incongruent print color. The experiments showed that there was somewhat more interference when the (irrelevant) color in which the target word was printed was available as an alternate response patch than when it was not. However, significantly more interference was determined by whether the (irrelevant) print color belonged to the set of colors actually used as target words in the course of the experiment. Moreover, the same patterns of interference emerged when the task was altered so that the mere presence of the correct color had to be indicated with a keypress, rather than by pointing to the color's location. Consistent with translation models of Stroop interference, these results demonstrate the theoretical importance of competition among internal representations, rather than among stimuli or responses.

In the classic Stroop effect (Stroop, 1935), two incompatible stimuli are presented: a color word and the (incongruent) color in which it is printed. A single response is required, which is to name the color in which the color word is printed. Performing this task with a series of such stimuli reveals, even subjectively, that it is extremely difficult to simply ignore the printed text. One finds oneself slowed by suppressing the word or, worse, saying it inadvertently. A crucial and somewhat confusing aspect of the Stroop effect is that this apparent failure of selective attention is asymmetrical. It is the verbal information that interferes with the sensory. It is further important to appreciate that this holds true even when the response is an arbitrarily assigned keypress. The present investigations concern an important exception to this general rule. Durgin (2000; see also Sugg & McDonald, 1994) has shown that the asymmetry of Stroop interference can be reversed, so that sensory color information interferes with verbal information, rather than vice versa. In the research to be presented here, this new reverse Stroop task (RST) was used to try to clarify the nature of Stroop and Stroop-like interference. In particular, it will be argued that theories couched exclusively in terms of stimuli and responses will probably need modification to include internal representations.

Durgin (2000) produced a reverse Stroop effect by having participants identify colors by pointing to one of several available color patches, using a mouse cursor. A version of the stimulus display is shown in Figure 1. When the task was to respond to the sensory color of the Stroop word stimulus, the irrelevant verbal information produced almost no interference (incongruent trials were about 10 msec slower than neutral trials). Conversely, when the task was to identify the word, the irrelevant sensory color information produced strong interference (incongruent trials were about 80 msec slower than neutral trials). Durgin argued that this result was due to a compatibility between sensory color information (whether as target or distractor) and the task of pointing to a sensory color. At any rate, the asymmetry of Stroop interference could not be understood in terms that demanded that verbal information always have precedence, because it clearly did not in this instance. Note that the relevance of the RST results to normal Stroop derives from the fact that (1) the target stimuli used are the same as those in a normal Stroop task (a single incongruently colored color word) and (2) the task is one of color identification (albeit by a different means).

Reverse Stroop effects are well documented in the literature. However, their interpretation has been controversial. When a Stroop stimulus was accompanied either by a rectangle of color or by an uncolored word, the task of deciding whether this accompanying stimulus was the *same* as one or another aspect of the Stroop stimulus showed symmetrical interference (Treisman & Fearnley, 1969). Specifically, the task of indicating whether the verbal aspect of a Stroop stimulus matched the color of an accom-

These data were previously presented at the 41st Annual Meeting of the Psychonomic Society, New Orleans, November 17, 2000. This work was supported by a Swarthmore Faculty Research Grant and by the Howard Hughes Medical Institute. Many thanks to John Christie and Tram Neill for valuable comments and suggestions. Thanks to Michelle Park and Laura Fox for conducting the experiments and to Michael Waddington for contributions to the design of Experiment 1. Correspondence concerning this article should be addressed to F. H. Durgin, Department of Psychology, Swarthmore College, 500 College Avenue, Swarthmore, PA 19081 (e-mail: fdurgin1@swarthmore.edu).

 Copyright 2003 Psychonomic Society, Inc.

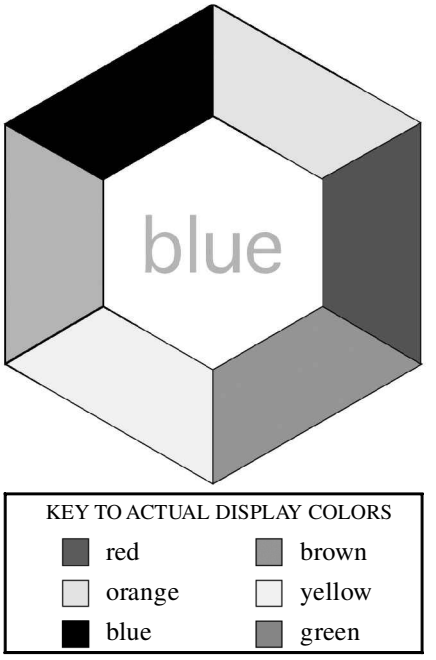


Figure 1. Hexagonal reverse Stroop display used for Experiments 1 and 2. The reverse Stroop task (Durgin, 2000) is to point to the trapezoidal color patch corresponding to the identity of the (incongruently colored) color word.

panying rectangle showed interference from incongruent print color. These tasks are known as matching tasks.

The symmetry of interference in matching tasks has been questioned, however, and has been attributed to strategic scanning or coding (Morton & Chambers, 1973). If only two stimuli are presented for comparison, the direction of comparison is up to the participant. In the RST paradigm, on the other hand, it is clearly advantageous to decode the target information first before attempting to figure out which of the surrounding patches match it, if any. For this reason, the RST is rather different from a classic matching task.

Moreover, some of these matching-task reverse Stroop effects could have been due to response-end competition from the evident difference between the sensory colors present. Indeed Egeth, Blecker, and Kamlet (1969) showed that in a color-matching task in which the words SAME and DIFF were used in place of color words, verbal interference was reestablished. Similarly, nonlinguistic recognition of color differences might have impeded *same* responses, and vice versa. The RST paradigm offers a number of advantages over traditional matching paradigms because, although entailing a final comparison stage, it more clearly involves identification of the target item, rather than a binary response (see also Compton & Flowers, 1977). In addition, the effect requires word identification, so coarse scanning strategies (e.g., Uleman & Reeves, 1971) are not a concern.

The classic Stroop effect has been treated, variously, as evidence of the automaticity of semantic processing of words, as a result of a biased race between verbal and sensory information, and as a failure of selective attention. Within the last decade, five distinct approaches to understanding Stroop phenomena have emerged. Three of these approaches were concerned primarily with explaining the asymmetry, and two were concerned primarily with a formal characterization of Stroop interference. Clearly, a reverse Stroop effect is most relevant to the approaches concerned with asymmetry. However, the present experiments are also relevant to the formal characterization of Stroop and Stroop-like interference, because they suggest a more complex information-processing architecture than is sometimes supposed.

Response Competition

Fundamental to the Stroop effect is the notion of competition, but this is often regarded initially as a competition between the color and the word stimuli (e.g., speed of processing). In his 1991 review, MacLeod made a strong case for the idea of response competition as an important component of Stroop interference. The idea is that the decision point at which competition occurs is at the response end of the pathway. This kind of notion is strongly supported by response set effects in Stroop experiments: If a limited set of sensory colors is used in an experiment, color words from outside that set will interfere less than those from within the set (Klein, 1964; Proctor, 1978). This general notion is also consistent with stimulus–response compatibility effects. If one set of stimuli (e.g., the distractor information) is more compatible with the specific response required, it could carry more weight in the competition among responses (e.g., Flowers, Warner, & Polansky, 1979). Thus, changing the nature of the response can change the balance of the competition.

Stimulus Salience

Melara and Mounts (1993) argued that the asymmetry found in Stroop interference could be eliminated if the relative saliences of the color and the verbal information were matched, where relative salience was defined in terms of Garner interference. Garner interference is defined as the cost to making discriminations on one dimension when there is variation in another (irrelevant) dimension. Reducing the text size was sufficient to shift the balance both of Garner interference and of Stroop interference in a buttonpressing task using Stroop-like stimuli. Dyer and Severance (1972), improving on a study by Gumenik and Glass (1970), had similarly shown a *reversed Stroop* effect when word legibility was reduced by a mask.

Clearly, stimulus discriminability can contribute to processing speed, but the RST paradigm does not involve any reduction in word legibility and is, therefore, probably not a stimulus-end effect. Therefore, a concept of response-end compatibility still seems to be required to account for changes in the direction of asymmetry in the RST. A more complex architecture will be suggested later.

Pathway Strength

In an attempt to ground, in connectionist terms, the notion of automaticity commonly attributed to verbal processing in Stroop interference, Cohen, Dunbar, and McClelland (1990) developed a connectionist model of the Stroop effect. They modeled a two-color system and showed that by simply strengthening the pathways between the verbal stimuli and the correct output responses, they could produce patterns of performance that approximated human reaction time (RT) data in a Stroop task. Several critiques of the details of the model have been made (e.g., Kanne, Balota, Spieler, & Faust, 1998; Mewhort, Braun, & Heathcote, 1992), but the basic idea is clearly on track: Irrelevant information traveling along stronger pathways can interfere with target information traveling along weaker ones.

A limit of this model, however, is its inflexibility with respect to asymmetry. If verbal information is said to be "better connected," in order to explain the asymmetry of Stroop interference, how can reverse asymmetry, as in a reverse Stroop effect, be accounted for without invoking a whole new pathway? That is, Cohen et al.'s (1990) model describes the Stroop asymmetry and suggests how asymmetry can arise from pathway strength, but it does not explain the difference between the normal Stroop task and the RST paradigm. Adding a second network for handling the RST paradigm would seem the simplest option, but this seems unprincipled (see Sugg & McDonald, 1994). Cohen et al.'s model is not an explanation of the direction of asymmetry, because it determines this merely by positing an asymmetry in pathway strength.

Making use of H. A. Simon's (1981) notion of explanation in hierarchical systems, I would argue that the kind of explanation that Cohen et al. (1990) provide is cast at a lower level of explanation than that which I seek to provide in this article. What I will present later is a (simplistic) high-level architecture that does explain Stroop-like asymmetries. I will assume that Cohen et al.'s model is a kind of existence proof that precise algorithms can be developed that will fill in the role of *strong* versus *weak* pathways within that architecture (between a given set of pieces of information and a given set of activations). The important point is that explanations can be cast at multiple levels simultaneously. Rather than supposing that there is a flat architecture linking stimuli directly to responses, as Cohen et al. have done, I will argue that mediating representations of different kinds are valuable explanatory principles.

The Dimensional Overlap Model

Kornblum's (1992) dimensional overlap model of the Stroop effect recognizes two sources of interference. On the one hand, there is competition between the target stimulus and the distracting stimulus. In the classic Stroop task, the distracting stimulus is a color word, whereas the target stimulus is a sensory color. On the other hand, it is widely understood that there are also response compatibility effects, such that the response itself—naming the target color—is compatible with the distracting information. Kornblum has referred to these instances of competition and compatibility as varying in *dimensional overlap*. As Kornblum has pointed out, there are some tasks in which dimensional overlap is limited primarily to that between the distracting information and the response. In the oral Stroop task, there is dimensional overlap not only between the target and the distractor stimuli, but also between each of these stimuli and the set of responses.

At first glance, it would seem that a reverse Stroop effect could be accounted for by emphasizing the importance of stimulus–response compatibility. However, since the conflict stimulus (an incongruently colored color word) is the same in both classic and reverse Stroop tasks, this would be tantamount to completely discounting the importance of stimulus–stimulus competition in explaining Stroop asymmetry. This, in turn, would seem to undermine the motivation for applying the dimensional overlap model to Stroop interference. Rather than collapsing the dimensional overlap model to one of stimulus–response compatibility, it may be advisable to expand the model to include a notion of internal representation that mediates between stimulus and response.

The Translation Model

As originally proposed by Virzi and Egeth (1985; see also Compton & Flowers, 1977), the translation model suggests that, along the pathway from stimulus to response, the form of internal representation employed by cognition may require a change. If the task requires lexical information, as in a color-naming task, lexical stimulus information will have an advantage, because it requires little or no translation during processing. If the task requires sensory (imagistic) color information, however, that information will require translation (into lexical form) and, therefore, will be delayed.

This sort of model is consistent with response compatibility effects emphasized by Flowers et al. (1979) and Treisman and Fearnley (1969). More recent translation models explicitly model stimulus–response compatibility by going beyond stimulus and response and postulating the existence of internal representations of different forms. A translation model has been developed by Glaser and Glaser (1989) to account for picture/word interference and has been further refined by Sugg and McDonald (1994), who developed a reverse Stroop effect formally similar to Durgin's (2000).

The experimental technique of Sugg and McDonald (1994) used only two colors and two response buttons. The crucial manipulation involved how the response buttons were labeled. The buttons were actually virtual buttons on a touch screen and could be labeled either with verbal information or with sensory colors. The type of label was varied between blocks of trials. The specific labels varied randomly from trial to trial. Although their participants were pressing virtual buttons, the actual paradigm is quite similar to Durgin's pointing task, in which the virtual buttons are labeled with colors. As Sugg and

McDonald had predicted, the standard Stroop asymmetry of interference was found when the virtual buttons were labeled with words, but the opposite asymmetry (a reverse Stroop effect) was found when the buttons were labeled with sensory colors.

Although Durgin (2000) suggested that the reverse Stroop effect found with the RST was due to a special compatibility between manual pointing and sensory information, the results of Sugg and McDonald (1994) indicate that pointing (using a touch screen) can support normal Stroop interference when the locations are labeled with words. Rather than reflecting stimulus–response compatibility, the bottleneck lies in the translation between stimulus information and an internal representation suitable for guiding the response. The inclusion of this internal representation in a proper translation model can help to explain how manual buttonpress tasks tend to show normal Stroop interference. Verbal interference may occur because the internal codes used to label the various button responses are verbal, rather than sensory, in structure (see Hock & Egeth, 1970, for another discussion of implicit codes in Stroop interference).

Straw versions of the translation model that lack the theoretical construct of internal representations can hardly work. For example, Logan and Zbrodoff (1998) tried to argue against translation models by showing that a typing response (typing the color name) actually shows equal or greater interference than does a vocal response to a printed Stroop stimulus. Their reasoning was based on the assumption that "manual" responding is somehow nonlinguistic, and they explicitly argued that translation theories concern direct translation between stimulus and response. But this is false. Translation theories, such as Glaser and Glaser's (1989) model, explicitly depict mediating stages of encoding, including orthographic representations. Although it is true that typing does not engage syntax (Salthouse, 1986), it certainly seems to engage orthography. Considering that the interpretation of the printed word and the production of the typed response probably both depend on an (internal) orthographic representation, this heightened Stroop effect with "manual" responding by means of typing is actually a strikingly confirmed prediction of a properly construed translation model. Indeed, orthographic codes are explicitly represented in Glaser and Glaser's model.

Similarly, it is easy to understand reverse Stroop interference in the RST by appealing to an internal sensory (or image-like) representation that can guide visual search for sensory colors. Such a representation would be informed, most readily, by sensory color information from the target stimulus.

Refining the RST Paradigm

The reverse Stroop effect that Durgin (2000) found with the original RST paradigm is consistent with the latter sort of translation explanation. It is reasonable to believe that the reversed asymmetry of interference, in which sensory color interferes with the identification of the printed word, results from the necessity of translating the word into a sensory representation of color for purposes of searching for the appropriate response location. On this account, the direction of interference derives from the fact that the internal representation employed for the search task is imagistic, rather than lexical, in structure.

However, apparent reverse Stroop interference could also have been produced in the RST by peripheral competition from spatial location information, rather than from color identity. That is, the salient (although irrelevant) color match between the sensory color of the text and the sensory color of a (wrong) response location might have produced a conflicting spatial signal that hindered speedy response to the correct location—a kind of Simon effect (J. R. Simon, 1969). The latter account avoids appealing to any sort of internal sensory representation of color and would render the RST a mere curiosity, irrelevant to understanding Stroop-like competition.

One clear way to discriminate between these accounts is to introduce trials in which the target word is printed in a conflicting color that does not appear in a response location on that trial. For example, the word *red* might appear in blue on a trial in which blue does not appear in one of the response locations. If the peripheral account is correct, reverse Stroop interference should disappear in this condition, whereas the translation account predicts that the effect should remain. Of course, it is possible that both peripheral matches (location competition) and internal representations (coding competition) can contribute to the effect, but the contributions of the latter ought to be more important.

An alternative way of testing whether reverse Stroop interference results from competition at an internal coding stage is to look for *set* effects. As was reviewed above, Stroop effects are known to be less strong when the distracting information indicates a response never given (Klein, 1964; Proctor, 1978). For example, if the color blue never appears as the target color but the color red does, then identifying the color green will suffer more interference when presented as the word *red* than as the word *blue*. Although this typically has been understood to reflect a preactivation of the set of *responses*, it is equally consistent with the preactivation of internal (e.g., lexical) *codes* for color names necessary to responding. In the RST, one could use a similar manipulation to test for evidence of preactivation of sensory color codes corresponding to target words used. Since the overt responses are not the colors themselves (but rather a spatial pointing task), set effects in RST clearly would reflect internal codes.

Because the manipulations of set and of match are orthogonal to each other, they can be combined in a single experiment. For example, if eight colors are available as (distractor) text colors, but only four of these are ever used in the set of target color words, the others can still be used as the distractor on any given trial. Moreover, if only a subset of the eight colors is made available as responses (color patches) on any given trial, the distractor color used

need not have a match among the available responses. This means that in an RST trial, the color in which the target word is printed (the distracting color) can either be from the set of four target colors used in the experiment or not. Moreover, whether it is from the target set or not, that distracting color either can have a match among the responses available or can be unmatched by any of the visually available response options.

In the first experiment presented here, the results of such a manipulation are reported, demonstrating that peripheral matches, alone, cannot explain interference effects in the RST, whereas set effects are quite strong. Experiment 2 extended this investigation by changing the manner of response. The primary purposes of these experiments were to show that Stroop-like competition can occur among internal representations, rather than among stimuli or responses, and that the presence of a conflicting matching location is not necessary to generating a reverse Stroop effect when the RST paradigm is used.

EXPERIMENT 1
Set Effects and Match Effects

In the present experiment, two questions were simultaneously investigated. First, can reverse Stroop interference in the RST paradigm be explained by peripheral competition from the presence of a sensory match to the distracting color information? That is, is a peripheral sensory match to the distractor color necessary? Second, independent of the presence of a peripheral match to the distracting color, does the identity of the distracting color as a previously coded target influence the amount of interference?

In the design of the experiment, there were no neutral trials, so the theoretical comparisons are among trials that were all assumed to include some amount of reverse Stroop interference. Previous estimates had placed the amount of reverse Stroop interference (relative to neutral trials) at about 80 msec in a similar paradigm. On the assumption that peripheral matches account for only a minor amount of interference, it was predicted that the absence of a peripheral match to the distracting color would produce only a minor reduction in interference. Similarly, the set effect represented a manipulation of the quantity of interference expected and was not expected to eliminate all interference, but only to reduce it for distracting colors that were not from the set of targets used in the experiment.

The two effects might well be expected to be independent. One is supposed to occur near the level of explicit response encoding (representation of location), whereas the other is hypothesized to occur at the level of internal coding of the target stimulus (representation of sensory color).

Method

Participants. Twenty-seven introductory psychology students at Swarthmore College participated, in partial fulfillment of a course requirement.

Task. On each trial, a single color word, printed in an incongruent color, appeared on the display screen. This target word was surrounded by a hexagonal array of six trapezoidal color patches, as is shown in Figure 1. The observer's task was to use the mouse cursor to point to the color patch corresponding to the color named by the target word, ignoring the color in which the word was printed.

Design. All the trials were incongruent trials (the target word was presented in an incongruent color). However, four kinds of trials were defined by two aspects of the distracting color used. First, the color in which the word was presented either was from among the set of four color words used in the experiment or was one of the four colors never presented as a target word. (The four color words used out of the eight colors available in the experiment were chosen at random for each participant.) Thus, the distracting color could either be a color that had been a target on some previous trial (in set) or not (out of set). Moreover, because only six color choices were presented on each trial (three were always in-set colors, and the other three were always out-of-set colors), a color of each type was absent on each trial. Thus, the distracting color (in which the target word was printed) could be either present among the response colors for that trial (matched) or absent (unmatched).

The participants completed six blocks of 36 trials each. There were twice as many matched trials as unmatched, so that the 36 trials consisted of 12 in-set matched trials (ISMTs), 6 in-set unmatched trials (ISUTs), 12 out-of-set matched trials (OSMTs), and 6 out-of-set unmatched trials (OSUTs). Target locations were balanced within trial types for each block. Specific color selections were randomized. Trial order was randomized within each block. The first block was considered practice and was not analyzed. That left 180 trials for analysis, 60 of each of the matched trial types and 30 of each of the unmatched trial types.

Stimulus details. The words were presented in 64-point Geneva lowercase letters against a black background at the center of a high-resolution Macintosh display (28 pixels/cm) viewed, without restraint, at a distance of about 50 cm. The eight colors used in the experiment, specified as 8-bit RGB values, were red (255, 0, 0), green (0, 170, 51), blue (68, 68, 255), brown (132, 102, 17), yellow (255, 255, 0), purple (170, 0, 170), orange (255, 153, 0), and gray (170, 170, 170).

Each trial began with a white square fixation mark, 20 pixels across, at the center of the otherwise black screen. The student initiated the trial by clicking the mouse on the fixation square. This action caused the fixation square and the mouse cursor to disappear, so as not to mask the word. (The mouse cursor would then reappear as soon as the student moved the mouse outside the region defined by the fixation square.) After a delay of 500 msec, the stimulus word appeared simultaneously with the six response patches and remained on the screen until response. The trapezoidal response patches were 100 pixels (4 arc deg) from the center of the display and were 100 pixels thick.

RT was defined as the time at which the mouse cursor's tip entered one of the six colored response regions. A ballistic motion that passed the cursor through the color patch was, therefore, sufficient. A physical movement of the hand-held mouse by about 2 cm along the table surface sufficed to reach any patch. A beep for incorrect responses provided feedback.

Results and Discussion

Median RTs on correct trials for each of the four kinds of trials were computed for each participant. The means of these are shown in Figure 2. A 2×2 repeated measures analysis of variance (ANOVA) revealed main effects of both set and match but no interactions between them. Of principal interest is the finding that RTs when the distractor color was an out-of-set color (708 msec) were faster than when the distractor color was from within the

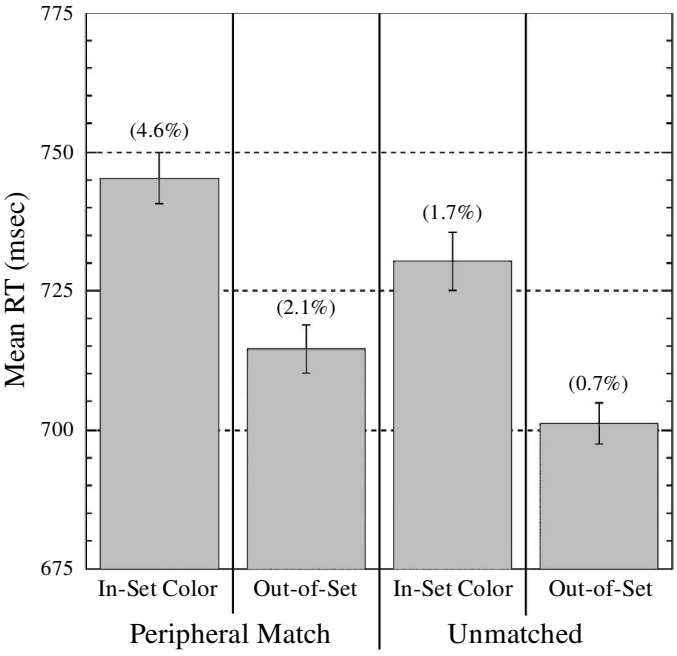


Figure 2. The results of Experiment 1. Average median reaction time (RT) is plotted as a function of whether the distracting color was within the set used as target words or not and whether or not it was matched by one of the response colors presented in the periphery. Percentage of errors is indicated for each trial type. Within-subjects standard errors are shown.

target set [738 msec; $F(1,26) = 36.2$, $p < .01$]. This is consistent with the hypothesis that in-set colors promote greater competition in the representation of the target color and, thus, more interference.

This interference was independent of whether or not the sensory color in which the word was presented was present among the available responses. That is, there was no interaction between distractor color and distractor match [$F(1,26) < 1$, n.s.]. Nonetheless, responses were faster overall on unmatched trials (716 msec) than on matched trials [730 msec; $F(1,26) = 9.37$, $p < .01$], consistent with the idea that peripheral color matches to the distractor color directly signal a possible response location (e.g., to spatial attention).

The amount of interference due to sensory color matches was compared with that due to target set membership by calculating difference scores for each individual and performing *t* tests on those scores. In fact, the amount of interference engendered by distractor color (out-of-set − in-set: 30 msec) was reliably greater than that engendered by the actual presence of a response patch matching the distractor [matched − unmatched: 14 msec; $t(26) = 2.11$, $p < .05$]. In other words, RT was more dependent on whether the (always incongruent) color in which the target word was printed had appeared as one of the target colors than whether it was physically available in a potential response location on the present trial. Thus, the most powerful interfering factor in the present situation was competition among representations of colors belonging to the target set, rather than competition among overt responses.

It must be emphasized that the size of the effects noted here may seem small relative to normal Stroop interference or even to reverse Stroop interference. This is because the effects we are measuring are between incongruent trials. Much larger effects have been demonstrated when incongruent trials were contrasted with neutral trials (Durgin, 2000).

Sugg and McDonald (1994), who developed the touchscreen reverse Stroop paradigm with only two virtual response buttons, undertook some closely related manipulations in their Experiments 2 and 3. However, a limitation of their experiments was that only a single color (blue) was added to the two colors already used in their design (red and green). In their Experiment 2, blue appeared only in OSUT trials—that is, the rectangle of (irrelevant) color was blue, whereas the response buttons were red and green. The results in this condition did not differ from those in the "neutral" condition, in which the rectangle remained gray—but, of course, gray is also a color. In their Experiment 3, they explored ISUT and ISMT trials by including blue as a possible target word. Perhaps due to the complexity of their design, which included seven different stimulus onset asynchronies (SOAs), they failed to detect any difference between matched and unmatched trials. Their graph, however, suggests that such

differences may have existed for an SOA of 0 msec and also when the distractor color preceded the target word by just 100 msec (see Sugg & McDonald, 1994, Figure 7, p. 671).

Errors. Errors in the pointing task were computed as a percentage of each trial type. Note that pointing errors ought to reflect a strong influence of the presence of an alternative match to the distractor. When no alternative match is present, a correct response might be delayed, but an error response seems less likely. (One might waste time looking for *blue*, for example, but one can hardly point to it if it is not there.) In fact, there was a reliable interaction between set and match in the error analysis [$F(1,26) = 4.3, p < .05$], so separate paired comparisons were conducted among the four trial types. Paired t tests showed that error rates for ISMT (4.6%) were higher than those for ISUT [1.7%; $t(26) = 4.13, p < .01$]. Error rates were also higher for OSMT (2.1%) than for OSUT [0.7%; $t(26) = 2.70, p < .025$]. Within the matched trials, ISMT error rates were reliably higher than OSMT [$t(26) = 4.54, p < .01$]. However, the difference between ISUT errors and OSUT errors was not reliable [$t(26) = 1.62, p > .10$].

Overall, these patterns of errors support the RT data. Pointing errors are obviously more likely when there is a distractor match to point to, but the role of set is still clearly important in determining which matches are more likely to produce errors. Distractor match errors corresponding to within-set colors were more than twice as frequent as those corresponding to nonset colors.

Conclusions. This experiment has accomplished the goals set out in the introduction. The strong set effects support the claim that Stroop-like interference in this paradigm results from competition among internal representations, which can be made prepotent by their use. At the target identification stage, the color in which a word is printed can produce greater interference when it is a familiar target color than when it is a color that has never served as an actual target. Moreover, the reverse Stroop interference found with the RST paradigm cannot be attributed primarily to peripheral matching of the distracting color, because the match effect is only 14 msec. It seems likely that sensory color matches may affect spatial attention or otherwise directly affect the determination of the overt pointing response.

It is worth noting, once again, that all the trials here were conflict trials. Although the presence of a match was manipulated directly (present or absent), a conflicting text color was always present. The manipulation of set therefore determined the *extent* of the inner conflict, whereas the manipulation of match determined the *existence* of the outer conflict. It is, therefore, all the more striking that the full measure of the outer conflict (about 14 msec) was reliably less than that of the set effect (about 30 msec).

In sum, the present results indicate that a major portion of interference in RST occurs at a site that represents the stimulus identity for purposes of responding. Consistent with response compatibility and translation theories, we can conclude that the reversed direction of interference in this kind of paradigm is most likely determined by the form of the internal representation that can most efficiently guide responding. If the internal representation used were lexical, we would expect traditional Stroop interference. We would expect this if the response could be guided categorically by linguistic labels (which might be implicit in a keypress task). If a sensory (imagistic) representation was employed in the current task, that would be sufficient to explain why the direction of the asymmetry of interference was reversed.

EXPERIMENT 2
Representation Matters More Than Task

Durgin (2000) emphasized the importance of the pointing task for producing his reverse Stroop effect, suggesting a privileged perception–action connection. However, it now seems unlikely that the pointing task per se produces an advantage for sensory color. Rather, the results of Sugg and McDonald (1994) suggest that a reverse Stroop would obtain only when it is a color, rather than a word, that is the target of the pointing. Conversely, the use of a pointing response need not be crucial to the generation of reverse Stroop interference.

It was shown in Experiment 1 that RST interference does not depend on the presence of a conflicting color match. Although pointing errors were more likely if there was an alternative response available, RTs were more affected by the identity of the distracting color (relative to the set of targets). This suggests that interference arises in the identification of the target, rather than in the act of pointing itself.

One way of thinking about the RST paradigm is as a visual search task in which the search target is identified to the searcher by the Stroop word, although the search target itself is defined as a sensory color. In order to demonstrate that the reverse Stroop results from the RST are not dependent on the requirements of the pointing task itself, the present experiment introduced a nonpointing variant of the task. Here, the participant's task, as in some visual search paradigms, was simply to indicate whether the target color was present among the colors in the hexagonal array. In terms of stimulus information, this task was quite similar to the pointing task, except that mere presence, rather than location, was to be indicated. The requirements of the pointing task of Experiment 1 could be described as a three step process: R1, identify the target color (e.g., in a sensory code); R2, find the corresponding location (visual search); and R3, point to it.

Step R2 seems to require a sensory representation to be formed in Step R1, consistent with translation theories. Primary interference from the distracting sensory color, therefore, probably occurs in Step R1 (whereas the match effect might arise at Step R2, if attention is drawn to the patch matching the distracting color, or at Step R3).

From this perspective, R3, the overt response, might be completely past the site of reverse Stroop interference (although spatial interference could easily arise at this stage). Thus, substituting a new *go* response, G3, in place of R3, ought to produce patterns of data very similar to those found in Experiment 1.

Pilot testing showed that requiring participants to make a present/absent decision produced substantially longer RTs. Apparently, the complexity of making a response choice was quite costly. A go/no-go task was therefore chosen as more similar to the cognitive requirements of the simple pointing task of Experiment 1. Because some modification of the experimental design was required, including the inclusion of a ninth color (see below) and more trials, a full replication of the pointing task of Experiment 1 was conducted in conjunction with the go/no-go task. As will be shown below, RTs were quite comparable between the two tasks.

Method

Participants. The participants were 40 students at Swarthmore College, who received partial credit toward a Psychology 1 requirement. Twenty students completed the go/no-go task, and 20 completed the pointing task. None had participated in previous Stroop experiments. The data of 3 additional students were not included due to unusually high error rates or long RTs—beyond three standard deviations from the rest of the data.

Design. The task (point or go) was manipulated between subjects. In each case, there were 24 practice trials followed by five blocks of 48 trials each. In the pointing task, all the trials were positive trials, and there were equal numbers of ISMT, ISUT, OSMT, and OSUT trials. (See Experiment 1 for an explanation, but note that this distribution of trials is different from that of Experiment 1.) For the go task, half of the trials were catch trials in which the color that the word named was not presented. Both misses and false alarms were recorded for the go task, based on 30 trials per trial type. (Go/no-go trials terminated after 1,500 msec if no response had been made.) For the pointing task, errors were recorded from among the 60 trials per trial type. Response latencies were recorded for both tasks.

Because the catch trials were structured identically to the positive trials, it was necessary to introduce a ninth color, pink (255, 0, 170), so that there could be five, rather than four, colors in the target set. This enabled us to maintain the provision that three of the six colors presented in the response array on any given trial were from the target set, while also allowing both an (absent) target color and an unmatched distractor color from within the response set. Four colors were still sufficient for the out-of-set colors.

As in Experiment 1, the target colors were randomly selected for each participant. The target location was balanced within each trial type and block. The choice of colors was completely randomized from trial to trial within the formal constraints of the design. Minimum block size was therefore 48: 6 (location) × 2 (set) × 2 (match) × 2 (presence).

Dummy trials, selected randomly, were inserted after each error trial, so that all analyzed trials followed correct responses. The data from the dummy trials were not recorded.

Results and Discussion

Median RTs for both tasks were analyzed in a mixed within–between repeated measures ANOVA with task as a between-subjects factor and set and match as within-subjects factors. There was no main effect of task, nor did task interact with any other factors. Instead, just as in Experiment 1, there were main effects of set and of match and no reliable interaction between set and match. The RT data are shown in Figure 3.

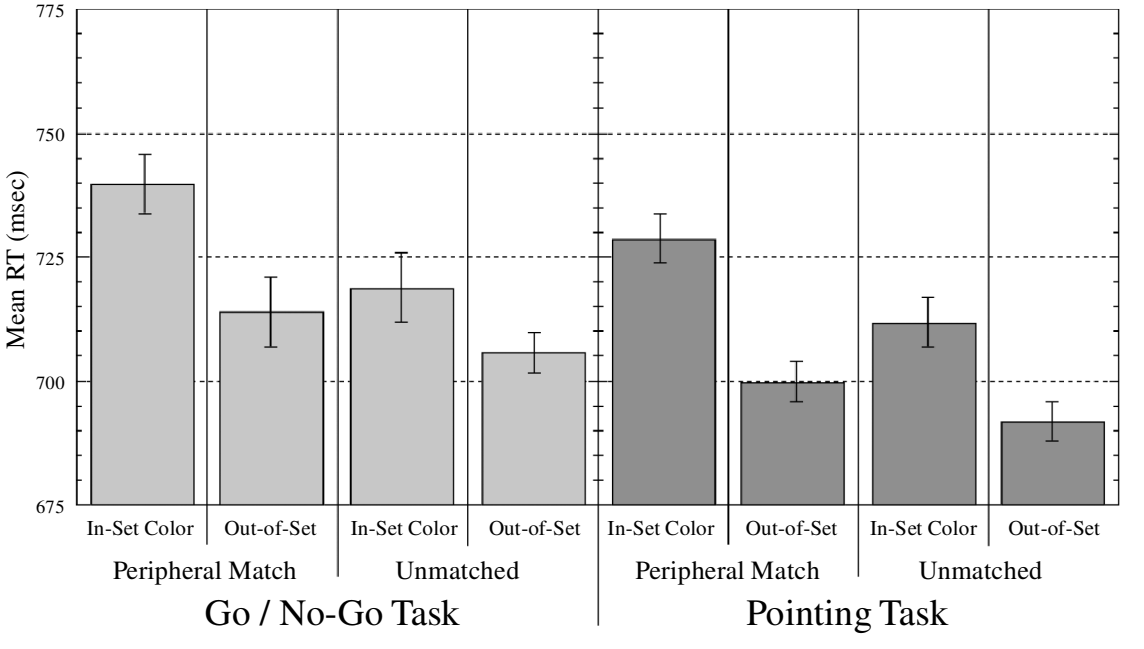


Figure 3. Reaction time (RT) data of Experiment 2. Average median RT is plotted as a function of whether the distracting color was within the set used as target words or not and whether or not it was matched by one of the response colors presented in the periphery. Within-subjects standard errors are shown.

Overall, RTs for trials with in-set distractors (725 msec) were longer than those for trials with out-of-set distractors [703 msec; $F(1,38) = 18.1, p < .01$]. RTs for matched trials (721 msec) were also longer than those for unmatched trials [707 msec; $F(1,38) = 9.95, p < .01$]. There was no reliable interaction between set and match [$F(1,38) = 2.38$, $p > .10$]. When separate analyses were conducted for the two tasks independently, the same patterns were obtained, except that the effect of match was only marginally reliable in the go/no-go task [$F(1,19) = 4.26, p = .053$].

Thus, the results of Experiment 1 were substantively replicated with the modified design for both the pointing task and the go/no-go task. Clearly, interference in the RST did not depend on the task of pointing per se. Indeed, the effect seems to be nearly identical in detail for the two kinds of tasks. On the assumption of the stage model presented in the introduction to this experiment, it would seem that *match* interference can be attributed to Stage R2 (search), rather than to R3 or G3 (response), because it is stable across response types. If match interference were due to spatial interference with the pointing response, it should not have influenced the go/no-go task.

Errors. Errors were defined differently for the two tasks and, so, were analyzed separately. For the pointing task, the pattern of errors replicated the results of Experiment 1. It appeared in the earlier data that pointing errors were uncommon when no distractor match was present but that pointing errors increased when the matched distractor was also from the target set. This pattern was expressed again in the present data as a marginal interaction between set and match [$F(1,19) = 4.51, p = .059$]. Separate t tests indicated that among matched-distractor trials, in-set distractors (4.1%) produced more errors than did out-of-set distractors [2.5%; $t(19) = 2.45, p < .05$]. Among unmatched trials, in-set errors (1.3%) did not differ from out-of-set errors [1.3%; $t(19) = 0$]. There was a trend for OSMT errors to outnumber OSUT errors [$t(19) = 1.8, p = .088$], but only among the in-set trials did matched-trial errors reliably outnumber unmatched-trial errors [$t(19) = 4.01, p < .01$]. Thus, only the ISMT trials differed reliably from all the others. These data are shown in Figure 4.

For the go/no-go task, the closest equivalent to a pointing error is a false alarm—a positive response when no color patch actually matches the target color. It is, therefore, unsurprising that the error patterns for false alarms were remarkably similar to those for the pointing task. As for pointing errors, there was a reliable interaction between set and match [$F(1,19) = 5.00, p < .05$]. When separate t tests were conducted, it appeared that here, too, only ISMT errors (4.8%) were reliably more frequent than ISUT errors [2%; $t(19) = 2.54, p < .05$], OSMT errors [2.5%; $t(19) = 3.20, p < .01$], or OSUT errors [1.3%; $t(19) = 2.67, p < .05$].

Misses from the go/no-go task were also analyzed. It might be expected that fewer misses would be likely for matched trials than for others, and the data plotted in Figure 4 support this conjecture, but because there were so few miss trials (1.9% overall), there were no reliable differences among these trials.

Conclusions. The patterns of RT data found with the go/no-go version of the RST paradigm replicated those found with the pointing task in this experiment and in Experiment 1.

The false alarm data may be interpreted as evidence of response competition in the go task, insofar as the presence of a sensory "match" seems to produce a tendency to respond, much as it produces a tendency to point to the incorrect color in the pointing task.

Nonetheless, the set effects observed in these experiments remain independent of the peripheral match effects. It should be emphasized that these set effects

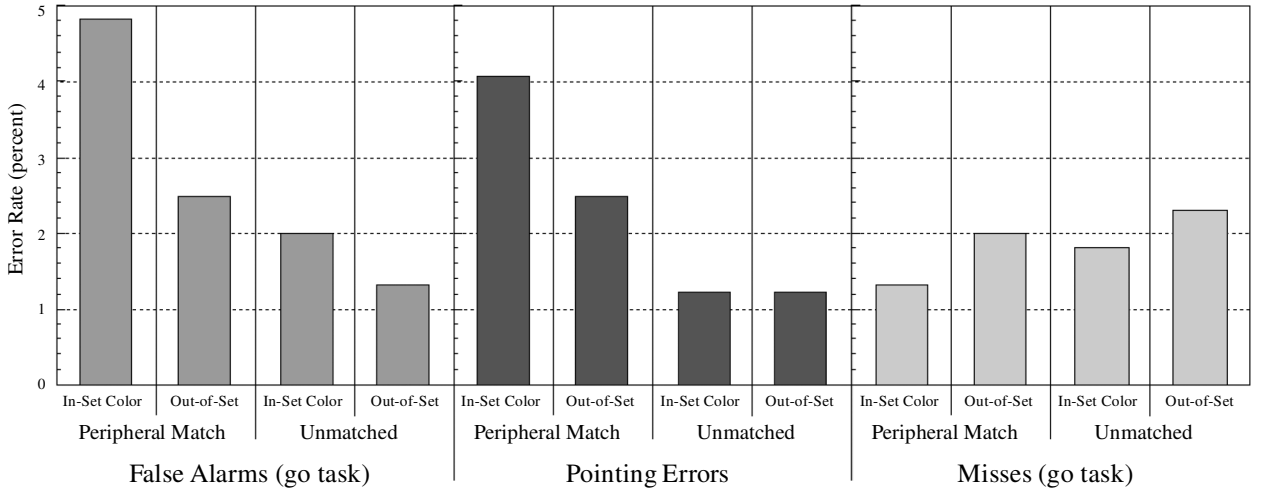


Figure 4. Error data of Experiment 2. Pointing errors, as well as misses and false alarms in the go/no-go task, are shown. Average error rates are plotted as a function of whether the distracting color was within the set used as target words or not and whether or not it was matched by one of the response colors presented in the periphery.

should be regarded not as evidence of *response* competition but, rather, as evidence of competition among *internal representations* of the stimuli.

GENERAL DISCUSSION

The two experiments reported here were designed to test three hypotheses. First, they tested whether the apparent reverse Stroop effect generated by Durgin's (2000) RST paradigm was caused by peripheral spatial distraction, rather than by true competition in the encoding of the Stroop target word. Both experiments showed that the presence of a false match among the peripheral response locations has only a limited cost on response latency. Second, the experiments tested whether the interference produced by the print color was modulated by whether it belonged to the response set. Both experiments confirmed that it was, and in neither experiment was there an interaction in RT between this set effect and the peripheral match effect. Finally, Experiment 2 showed that even when the response was changed to simple visual search ("Is the target color present?"), the patterns of interference were the same as those for the pointing task. This suggests that none of the interference in RST is related to the pointing response per se. It is concluded that the set effect reflects internal color-coding competition, whereas the match effect reflects spatial location competition (e.g., distraction of spatial attention) in the visual search stage of both tasks.

Treisman and Fearnley (1969) suggested that "if it were possible for subjects to generate colours," the direction of Stroop interference could be reversed. The results of the present experiments are consistent with this view. It seems that the task of locating the color that matches the one named requires generating an internal sensory representation of the appropriate color. The result is a reverse Stroop effect.

These experiments have further shown that set effects (often referred to as *response* set effects) can arise at an encoding stage that precedes identification of an overt response. Thus, interpreting set effects as *late* effects (see Sharma & McKenna, 1998) is not, a priori, warranted. Rather than response competition, the present results seem to reflect strong coding competition—in this case, at a stage of processing that encodes stimulus identity, probably in nonlexical codes. An outline of proposed processing steps includes (1) target identification, (2) visual search for target color, and (3) response (pointing to target color or merely indicating that the target was found). I have argued that *match* effects probably occur at Stage 2, whereas *set* effects, which are stronger, occur at Stage 1.

A schematic translation model is shown in Figure 5A (see also Compton & Flowers, 1977; Glaser & Glaser, 1989; Sugg & McDonald, 1994). In the terms of this diagram, set effects are hypothesized to result from competition among different color tokens within a single type of code (sensory or lexical) with the site of competition (and therefore, the direction of the asymmetrical advantage), depending on the task. Thus, target identification (Step 1, outlined above) would take place at one or the other of the diamonds representing internal codes.

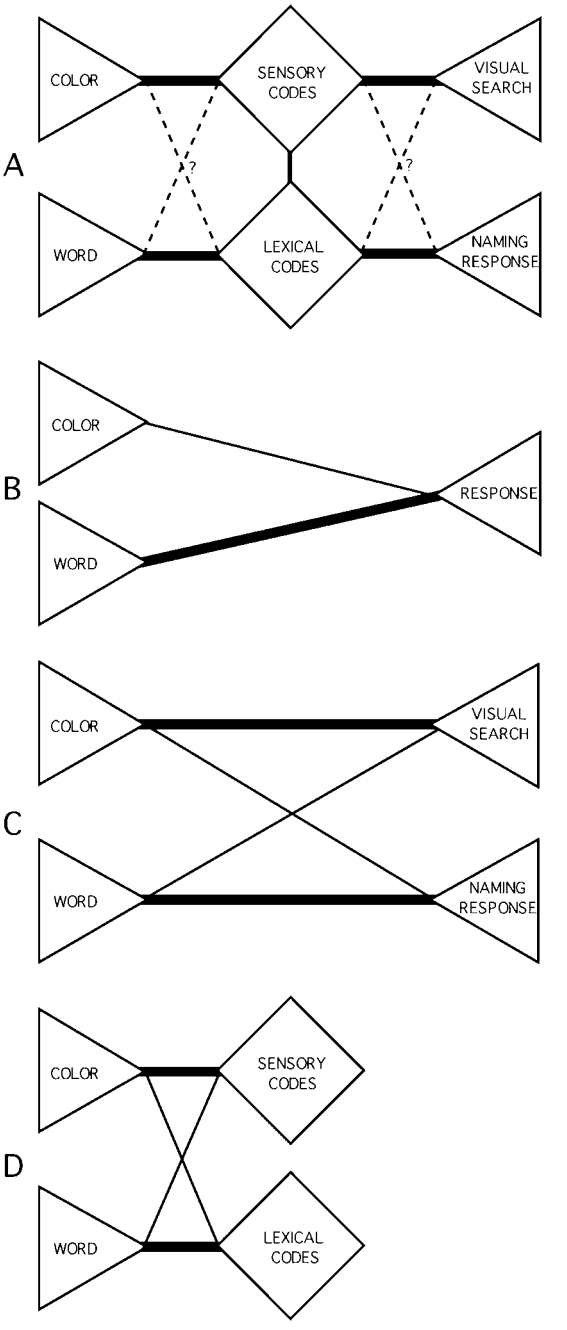


Figure 5. (A) A schematic illustration of a translation model, with processing from left to right. Dashed lines represent connections that might exist but are not necessary to the model and are presumed to be quite weak. (B) The general stimulus–response architecture of a pathway strength model of Stroop asymmetry (Cohen et al., 1990), shown for comparison. (C and D) Possible reinterpretations of the pathway strength model.

Note that, for simplicity and clarity of exposition, this schematic does not include a mechanism for spatial interference in the visual search stage that might be associated with the *match* effects of the present experiments.

For comparison, the high-level architecture of a pathway strength model of Stroop asymmetry (Cohen et al., 1990) is shown in Figure 5B. Pathway strength (represented by line thickness) is used in such a model to explain asymmetry of competition. An extension of the pathway strength architecture to capture stimulus–response compatibility effects is shown in Figure 5C, but this kind of extension seems inadequate to account for set effects without further elaboration. Figure 5D illustrates how the front end of the translation model might be simulated (but not explained) by a network like that of Cohen et al. To the extent that connections are believed to exist between lexical and sensory representations of concepts, the additional translation connection of Figure 5A (the vertical connector) seems fairly uncontroversial. In the translation architecture of Figure 5A, the chain of connections (and the absence of a robust direct pathway) between word and sensory codes or between color and lexical codes might be considered a *weaker* pathway than that connecting color and sensory codes and that connecting words and lexical codes. This is a formal description only, however, and cannot replace the explanatory power of the translation model concerning asymmetries of interference.

A feature of the translation model is that the kind of response demanded places constraints on the type of internal representation that is most useful. The present results strongly support the intuitive notion that visual search for a sensory color is best conducted by using a sensory representation (image) of that color (see Pashler & Shiu, 1999). This, in turn, determines the pathway along which information must travel in Figure 5A and seems to impose a need for translation and a clear advantage for sensory information.

It is important to remember that a distracting color can delay processing in the RST even when no match to that color is available among the responses. This is strong evidence that the bulk of RST interference is accomplished in the process of generating an internal representation of the target stimulus, rather than in searching for a candidate response location. The RST paradigm cannot be properly described as a matching paradigm (e.g., Treisman & Fearnley, 1969), because the explicit task is not to compare two stimuli. When Egeth et al. (1969) used the words SAME and DIFF instead of color words in a perceptual color-matching task, interference was reestablished. This suggests that a processing bottleneck in matching tasks would lie in the activation of *same* or *different*, rather than in the activation of color identities per se. By contrast, the data of the present paper suggest that internal representations identifying specific colors are justifiably regarded as the primary site of competitive interference in RST. Indeed, this was shown by the finding of Experiment 2 that the presence of a distractor color on *present* trials produced *slower* responses in the search task, which is the opposite of what might be expected by pure response-end competition.

As was emphasized above, the present task is better characterized as a visual search task (requiring target identification) than as a matching task (requiring only comparison). Although a process of comparison must occur in RST, that comparison may be best understood as taking place between an internal representation and an external set of color patches (visual search), rather than between two external stimuli. Matching tasks may produce reverse Stroop effects when they, too, employ sensory representations. Conversely, a search task or matching task might produce normal Stroop interference if the search or match objects were printed words, rather than sensory colors. It is the form of representation required, not the kind of task per se, that seems to determine the direction of asymmetrical interference.

Although the invocation of internal representations may seem uncontroversial to some, the general implications for normal Stroop effects may seem unproven to others. For those, it may be useful to consider that the translation model reinterpretation of the work of Logan and Zbrodoff (1998), presented in the introduction, was made possible by directly applying Glaser and Glaser's (1989) model, which already includes orthographic representations. This shows that translation models are actually helpful for predicting the asymmetry of interference in normal Stroop and for predicting how that asymmetry interacts with the specific response required.

In summary, interpreting Stroop-like interference without reference to internal representations may be unwise. Models that limit their discussion to stimuli and responses may be overly simple. Dynamic models of activation (e.g., Cohen et al., 1990) may themselves require translation so as to be incorporated into a more complex architecture of the stages of information processing in Stroop-like tasks.

Do the results of this paper truly apply to normal Stroop tasks? Some might still suppose that the standard automaticity account and the explanations that refer exclusively to stimulus and response are sufficient. However, denying that there are internal representations involved in normal Stroop tasks seems rather extravagant. After all, no one believes that the act of word naming (aloud) is automatic. Thus, it is reasonable to accept that a standard Stroop task also involves internal (lexical) representations. Once this is granted, a plausible formal similarity between standard Stroop interference and that generated by RST is manifest. The arguments made here about translation theories as a basis for Stroop asymmetry may, therefore, be hypothesized to apply both to the classic Stroop task and to the RST, as is illustrated in Figure 5A. Theories, such as translation models, that take into account the structure of internal representations seem the most promising avenue for developing sound

models of the general architecture of the processing underlying the asymmetries in Stroop and Stroop-like interference.

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.

Compton, R. P., & Flowers, J. H. (1977). Pictorial interference with geometric shapes and achromatic shades. *Perception & Psychophysics*, **22**: 303-309.

Durgin, F. H. (2000). The reverse Stroop effect. *Psychonomic Bulletin & Review*, **7**, 121-125.

Dyer, F. N., & Severance, L. J. (1972). Effects of irrelevant colors on reading of color names: A controlled replication of the "reversed Stroop" effect. *Psychonomic Science*, **28**, 336-338.

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

Flowers, J. H., Warner, J. L., & Polansky, M. L. (1979). Response and encoding factors in "ignoring" irrelevant information. *Memory & Cognition*, **7**, 86-94.

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

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

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

Kanne, S. M., Balota, D. A., Spieler, D. H., & Faust, M. E. (1998). Explorations of Cohen, Dunbar, and McClelland's (1990) connectionist model of Stroop performance. *Psychological Review*, **105**, 174-187.

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

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 II* (pp. 743-777). Amsterdam: Elsevier.

Logan, G. D., & Zbrodoff, N. J. (1998). Stroop-type interference: Congruity effects in color naming with typewritten responses. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 978-992.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Review*, **109**, 163-203.

Melara, R. D., & Mounts, J. R. W. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition*, **21**, 627-645.

Mewhort, D. J., Braun, J. G., & Heathcote, A. (1992). Response time distributions and the Stroop task: A test of the Cohen, Dunbar, and McClelland (1990) model. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 872-882.

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

Pashler, H., & Shiu, L.-P. (1999). Do images involuntarily trigger search? A test of Pillsbury's hypothesis. *Psychonomic Bulletin & Review*, **6**, 445-448.

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

Salthouse, T. A. (1986). Perceptual, cognitive, and motoric aspects of transcription typing. *Psychological Bulletin*, **99**, 303-319.

Sharma, D., & McKenna, F. P. (1998). Differential components of the manual and vocal Stroop tasks. *Memory & Cognition*, **26**, 1033-1040.

Simon, H. A. (1981). *The sciences of the artificial* (2nd ed.). Cambridge, MA: MIT Press.

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, **81**, 174-176.

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

Sugg, M. J., & McDonald, J. E. (1994). Time course of inhibition in color-response and word-response versions of the Stroop task. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 647-675.

Treisman, A. M., & Fearnley, S. (1969). The Stroop test: Selective attention to colours 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.

Virzi, R. A., & Egeth, H. E. (1985). Toward a translational model of Stroop interference. *Memory & Cognition*, **13**, 304-319.

(Manuscript received April 3, 2001;
revision accepted for publication October 3, 2002.)