Object-based attentional selection can modulate the Stroop effect

PETER WÜHR
Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

and

FLORIAN WASZAK
Max Planck Institute for Psychological Research, Munich, Germany

The Stroop (1935) effect is the inability to ignore a color word when the task is to report the ink color of that word (i.e., to say "green" to the word RED in green ink). The present study investigated whether object-based processing contributes to the Stroop effect. According to this view, observers are unable to ignore irrelevant features of an attended object (Kahneman & Henik, 1981). In three experiments, participants had to name the color of one of two superimposed rectangles and to ignore words that appeared in the relevant object, in the irrelevant object, or in the background. The words were congruent, neutral, or incongruent with respect to the correct color response. Words in the irrelevant object and in the background produced significant Stroop effects, consistent with earlier findings. Importantly, however, words in the relevant object produced larger Stroop effects than did the other conditions, suggesting amplified processing of all the features of an attended object. Thus, object-based processing can modulate the Stroop effect.

In every moment, one's environment contains many objects, each of which has several features (e.g., color, shape). For most of one's actions, however, only one of these objects is of relevance. Moreover, in many cases, only a few object features are useful for guiding one's behavior (Shalev & Algom, 2000). This fact implies that the cognitive system has to solve several problems before one is ready to act upon an object. The first problem is to identify an object that has the action-relevant features. Irrelevant objects that are simultaneously present should be ignored. This task has been labeled *input selection* (Treisman, 1969). The second problem is to identify the action-relevant feature on the already selected object. Irrelevant features should be ignored. This task has been labeled *analyzer selection* (Treisman, 1969), or *dimensional selection* (Allport, 1971; Shalev & Algom, 2000). The third problem, of course, consists of selecting the appropriate response.

In the present study, we investigated the role of (object-based) input selection in the famous Stroop task (Stroop, 1935). In a variant of this task, participants have to name the ink color in which a single word is printed. The critical manipulation affects the congruency between word meaning and word color. Word meaning can be congruent (e.g., the word RED in red ink), neutral (e.g., the letter string XXX in red ink), or incongruent (e.g., the word BLUE in red ink) with color. Usually, congruent stimuli produce facilitation (i.e., shorter response times [RTs] than for neutral stimuli), whereas incongruent stimuli produce interference (i.e., longer RTs than for neutral stimuli; for a review, see MacLeod, 1991). A fundamental feature of this so-called Stroop effect is its asymmetry. Whereas irrelevant words exert a strong impact upon color naming, irrelevant ink colors generally have no impact upon word reading (e.g., Stroop, 1935). Moreover, the Stroop effect emerges not only when ink color and word shape are integrated within the same object, but also when color and shape are spatially separated (Dyer, 1973a; MacLeod, 1998).

Obviously, the Stroop task contains each of the three problems described above (Van der Heijden, 1992). The problem of input selection is to first localize the response-relevant stimulus, or object. This task is certainly much easier when the stimuli are presented singly than when a list of stimuli is presented simultaneously, as in Stroop's (1935) original study (cf. Dyer, 1973b). Next, the problem of dimensional selection is to determine the relevant feature dimension on a multidimensional stimulus and to identify the particular stimulus value on that dimension. Finally, after having determined both the relevant stimulus and the relevant stimulus feature, an appropriate response has to be selected and executed.

Despite the vast number of publications on the Stroop effect, only a few studies have been concerned with the issues of input selection and dimensional selection in the

Stroop task. One possible reason for the neglect of these issues is the widely held belief that the main difficulties in the Stroop task arise at the level of response selection (e.g., Dyer, 1973b; Lu & Proctor, 1995; Virzi & Egeth, 1985). For this reason, many models proposed to account for the Stroop effect have addressed neither input selection nor dimensional selection. This fact applies particularly to automaticity models (e.g., Posner & Snyder, 1975). Such models propose that word reading is more highly practiced and, therefore, more automatic than color naming. An automatic process is defined as being obligatory—that is, triggered whenever an appropriate stimulus is encountered, regardless of whether the observer focuses attention on that stimulus or intends to process it (cf. Brown, Gore, & Carr, 2002; Posner & Snyder, 1975). As a result, irrelevant words are unavoidably read, and therefore, word reading can interfere with the less practiced task of color naming.

The idea that word reading reflects an automatic process and, therefore, that the Stroop effect is unavoidable has been challenged. For example, Bauer and Besner (1997) were able to modulate the Stroop effect by manipulating task instructions. In the classify condition of their experiment, participants had to report the presentation of several colors by pressing a particular key (i.e., green, left key; red, right key). In contrast, in the detect condition, participants had to report the presence of a particular color by pressing one key and report the absence of the same color by pressing the alternative key (i.e., green present, left key; green absent, right key). Bauer and Besner (1997) observed a Stroop effect in the classify condition, but not in the detect condition. This result indicates that the Stroop effect is not unavoidable and challenges the view that word reading is (at least strongly) automatic.

In the following, we review empirical studies on whether the mechanisms of input selection are able to affect the Stroop effect. Mechanisms of dimensional selection, as well as their influence on the Stroop effect, are outside the scope of the present article. These mechanisms are thoroughly discussed by Dishon-Berkovits and Algom (2000), Melara and Mounts (1993), and Shalev and Algom (2000).

The purpose of input selection is to segregate the visual input into well-processed and less well processed parts. Two mechanisms have been proposed to achieve this purpose: the selection of locations (space-based selection) and the selection of perceptual objects (object-based selection). The notion of space-based selection rests on the idea that preattentive processes provide an empty representation of the spatial layout surrounding us and visual attention can move within that representation. Attentional selection is accomplished by focusing attention upon a particular spatial location. Selected stimuli are processed more strongly than stimuli that fall outside the attended region, which are effectively ignored. The best-known metaphor for a space-based mechanism of selective attention is the *spotlight* (Posner, Snyder, & Davidson, 1980). Later space-based models have retained the idea that attention is focused within a spatiotopic coordinate system, while providing the selection mechanism with more flexibility. This flexibility is expressed in such metaphors as attentional gradients (LaBerge & Brown, 1989) or zoom lenses (Eriksen & St. James, 1986).

Some studies have been performed to investigate whether the allocation of spatial attention influences the Stroop effect (e.g., Brown et al., 2002; Gatti & Egeth, 1978). In most of these studies, the separated version of the Stroop task has been used—that is, a color patch and a word were presented at different display locations, and researchers varied the spatial distance between the patch and the word, as well as their congruency. The participants' task was to name the color of the color patch and to ignore the word.

Gatti and Egeth (1978) presented a color patch at screen center and two color words 1°, 3°, or 5° above and below the patch. Stroop interference was found to decrease with increasing distance, but interference was still significant at a distance of 5°. The authors concluded that spatial attention could not be limited to a region smaller than 5°. The decrease of interference with increasing distance was attributed to a decrease in perceptual acuity. Consistent with this interpretation, Merikle and Gorewich (1979) observed no decrease of Stroop interference with increasing distance (from 0.5° to 2.5°) when letter size was increased to compensate for acuity loss. Similarly, Hagenaar and Van der Heijden (1986) observed almost identical Stroop effects when a color patch and a word were presented close (1.1°) or far (1.9°) and concluded that spatial attention does not affect processing of Stroop words.

In a recent study, Brown et al. (2002) presented color patches and words at different locations. Briefly before these stimuli, an uninformative location cue appeared either at the position of the color patch or at the position of the word. The authors expected the abrupt-onset cue to involuntarily draw attention to its location (Yantis & Jonides, 1984). In several experiments, Brown et al. observed that manipulating the locus of spatial attention could modulate, but not eliminate, the Stroop effect. That is, a Stroop effect occurred even when attention was drawn to the position of the color patch and the word was positioned 13° away. The authors interpreted these findings as evidence for their notion that word recognition can proceed without spatial attention.

In contrast to the studies reviewed so far, in which the distance between a color patch and a word was varied in a separated Stroop task, Shalev and Algom (2000) used integrated Stroop stimuli. In addition, they presented a location cue either at the position of the Stroop stimulus (valid condition) or at an empty position (invalid condition). Shalev and Algom found comparable Stroop effects in valid and invalid conditions and, therefore, concluded that semantic processing of Stroop words is not affected by spatial attention at all.

In summary, the reviewed studies obtained very little evidence for the idea that spatial attention can modulate the Stroop effect. As long as the distance between a relevant color patch and an irrelevant color word is less than 5°, distance variations do not seem to affect the Stroop effect

(e.g., Hagenaar & Van der Heijden, 1986). Only when the distance exceeds 5° can increasing distance reduce, but not eliminate, the Stroop effect (Brown et al., 2002).

In contrast to space-based selection, object-based input selection rests on the idea that preattentive processes segment the visual scene into figures and ground and attention selects one of these figures (i.e., candidate objects) for further processing (Duncan, 1984; Kahneman & Henik, 1981). The attentional selection of a candidate object was proposed to amplify processing of each of the features of the selected object, whereas unattended objects are ignored. Although the exact mechanisms of object-based input selection are still unclear (Vecera & Farah, 1994; Weber, Kramer, & Miller, 1997), there seems to be agreement about the existence of a mechanism that primarily selects perceptual objects.

Yet only a few studies have addressed the role of object-based input selection in the Stroop task. The first study was conducted by Kahneman and Henik (1981), who reasoned that the difficulty of dealing with the integrated Stroop task might be related to object-based attentional selection. To test their hypothesis, these authors presented a circle and a square on opposite sides of fixation, with a horizontal distance of approximately 9°. Both shapes contained a colored word. The participants' task was to report the color of the word in the circle. One of the two words was always neutral (e.g., the word *cute*) with respect to the correct response. The second (critical) word was congruent, neutral, or incongruent. The most important manipulation was whether the critical word appeared in the circle (relevant-object condition) or in the square (irrelevant-object condition). Kahneman and Henik observed strong Stroop interference (202 msec) in the relevant-object condition and relatively weak interference (50 msec) in the irrelevant-object condition. Their interpretation of these results was that object-based selection modulates the Stroop effect.

In a follow-up study, Van der Heijden, Hagenaar, and Bloem (1984) replicated Kahneman and Henik's (1981) results with much smaller distances of only 1° between stimuli. In another experiment, Van der Heijden et al. observed interference in the relevant-object condition, but not in the irrelevant-object condition, when they validly precued the position of the relevant word with a bar marker. Therefore, these authors viewed their results as consistent with Kahneman and Henik's claims that attention selects objects and that object-based attentional selection modulates the Stroop effect.

In summary, the results of Kahneman and Henik (1981) and of Van der Heijden et al. (1984) suggest that object-based input selection can modulate the Stroop effect, even when the spatial distance between the relevant and the irrelevant object is small (1° in Van der Heijden et al.'s study). Hence, these results are in apparent conflict with the results of the studies on the impact of space-based selection on the Stroop effect. The reason for the discrepant results is unclear. A possible explanation may be that most of the studies on space-based attention (with Shalev & Algom,

2000, being the only exception) compared Stroop effects that were obtained in two separated Stroop tasks. In contrast, the studies on object-based attention compared the Stroop effect obtained in an integrated task with the effect obtained in a separated task. Therefore, the discrepant results may point toward a qualitative difference between integrated and separated versions of the Stroop task.

The most important point, however, is that the results of Kahneman and Henik (1981) and Van der Heijden et al. (1984) cannot unequivocally be explained by object-based selection. The reason is that, in these studies, the manipulation of whether the relevant (color) and the irrelevant (word) information belonged to the same object was confounded with a simultaneous variation of the spatial distance between relevant and irrelevant information. If the color and the word belonged to the same object, the spatial distance between the color and the word was small (i.e., minimal); if the color and the word belonged to different objects, the spatial distance between them was large. Thus, the observed decreases in the Stroop effect in the irrelevant-object conditions may, at least in part, be related to effective space-based selection. To conclude, at present, there is no clear evidence for a role of object-based attentional selection in the Stroop task (cf. Shalev & Algom, 2000).

The aim of the present study was to investigate the role of space-based attentional selection in the Stroop task. The question was whether object-based selection can modulate the seemingly ubiquitous Stroop effect. To investigate this question, we developed displays in which the affiliation of colors and words to the same or to different objects was manipulated without simultaneously varying the spatial distance between a color and a word. In particular, participants named the color of one of two superimposed rectangles (i.e. objects; see Figure 1). Congruent or incongruent color words were presented as parts of the relevant object, as parts of the irrelevant object, or in the background. The main question of interest was whether color words would produce larger Stroop effects (i.e., larger facilitation by congruent words and larger interference by incongruent words) when they appeared as parts of the relevant object, as compared with the remaining conditions.

EXPERIMENT 1

In Experiment 1, we asked whether an irrelevant color word affects color-naming performance more strongly when the word and the to-be-named color belong to the same object, as compared with a situation in which the word and the color belong to different objects. The participants were presented with two rectangles, one partially occluding the other in the form of a plus sign, or cross (see Figure 1). The front rectangle could have one of four different colors (blue, red, green, or yellow), whereas the rear rectangle always appeared in light gray. The participant's task was to report aloud the color of the front rectangle as quickly as possible. Thus, in Experiment 1, the front rec-
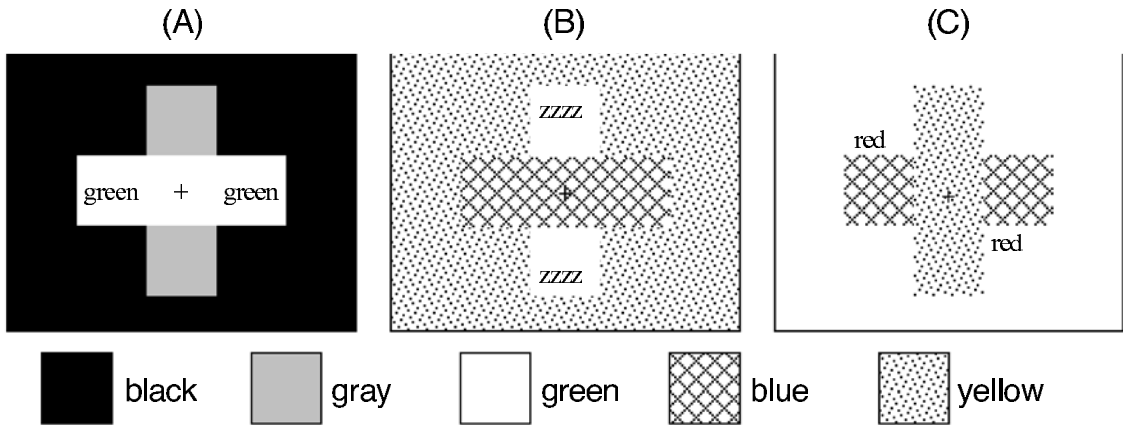
Figure 1. Examples of displays used in Experiments 1–3: (A) congruent color words in relevant object from Experiment 1, (B) neutral words in irrelevant object from Experiment 2, and (C) incongruent color words in background from Experiments 2 and 3. Note that, for matters of illustration, the depicted sizes of the stimuli are not proportional to the actual sizes used in the experiment (for actual sizes, refer to the description of methods).

tangle was the relevant object, whereas the rear rectangle was the irrelevant object. The critical manipulation was the presentation of irrelevant words at different screen locations. In the relevant-object condition, two identical word stimuli appeared at both ends of the relevant object. In the irrelevant-object condition, the word stimuli appeared at both ends of the irrelevant object. Moreover, the word stimuli were congruent, incongruent, or neutral with respect to the color of the relevant object (see Figure 1 for examples).

We expected that congruent color words would facilitate color naming and incongruent color words would interfere with color naming, independent of whether the words appeared in the relevant or the irrelevant object. This observation would be consistent with the results of several previous studies (e.g., Gatti & Egeth, 1978). However, we also expected facilitation from congruent words and interference from incongruent words to be larger when the words were part of the relevant object than when they were part of the irrelevant object. This observation would indicate the presence of object-based selection.

It is important to note that the spatial distances of the words from the fixation point were identical in the relevant-object and the irrelevant-object conditions. Thus, in contrast to Kahneman and Henik's (1981) study, our manipulation of whether the relevant color and the irrelevant word(s) belonged to the same object or to different objects was not confounded with a variation of the spatial distance between the position of the words and the focus of spatial attention. Moreover, the displays were presented very briefly (100 msec), to prevent eye movements from the fixation point to other parts of the display.

Method

Participants. Twenty-one volunteers (14 female), with a mean age of 23 years, participated in a single-session experiment. The experiment lasted about 30 min, and the participants received DM 12 for attendance. All the participants in this and the following experi-

ments were native German speakers, who classed themselves as having normal color vision as well as normal (or corrected-to-normal) visual acuity.

Apparatus and Stimuli. The experiment took place in a dimly lit and soundproof chamber. The participants sat in front of a 17-in. EIZO Flexscan 9080i monitor with an unconstrained viewing distance of approximately 80 cm. The participants responded by speaking into a microphone, which triggered a voice key measuring RTs to the nearest millisecond. An IBM-compatible computer controlled the presentation of stimuli and collected vocal RTs. Visual stimuli were shown on a black background. The fixation cross was a small "+" sign in font size 11, subtending 0.2° of visual angle. Each stimulus display consisted of two rectangles, which were superimposed upon each other and formed a cross, centered on the fixation point (see Figure 1). The short side of the rectangles subtended 1.1°; the long side subtended 3.2°. The orientation of the front and the rear rectangles (horizontal or vertical, respectively) was balanced across the experimental conditions.

The rear rectangle was always presented in light gray. By contrast, the color of the front rectangle was blue, green, red, or yellow. In each stimulus display, two identical words appeared at both short sides of one of the rectangles (see below). When the rectangle containing the words was oriented horizontally, one word appeared to the left and the other one to the right of fixation. When the rectangle containing the words was oriented vertically, one word appeared above and the other one below fixation. The words were presented in font size 11 and subtended 0.6° (*rot* [German for *red*]) and 0.8° (*grün* [German for *green*]). The spatial distance between the fixation point and the center of a word was 1.1° of visual angle in each condition.

Procedure. At the beginning of the session, the participants were familiarized with the task and were encouraged to respond loudly. They were also instructed to respond as fast and as accurately as possible and to fixate the fixation point. Moreover, the experimenter pointed out that the words presented in the display were irrelevant with respect to the task; the participants were told to ignore them.

After 20 practice trials, the experimental phase began with the presentation of the instructions on the screen. The experiment was run in blocks of 10 trials. At the beginning of each block, the string "Press a key" appeared. The keypress started a block of 10 trials, each of which contained the following events. First, the fixation cross appeared on the screen for 800 msec. Next, the stimulus display was presented for 100 msec. Then there was a blank interval

until the participant's response triggered the next trial, beginning with the presentation of the fixation cross. After the participant had responded to the last stimulus display of a block, the string "Press a key" appeared, and the participant could take a rest.

The participant's performance was monitored on line by the experimenter, who sat outside the experimental chamber. The experimenter heard the participant's responses via earphones and compared them with the correct answers that were shown on a second monitor. Each error was recorded in a list.

Design. The experiment used a 3×2 within-subjects design. The first factor was congruency. The irrelevant words were congruent, neutral, or incongruent with respect to the correct color-naming response. A congruent word denoted the color of the relevant object; an incongruent word denoted a different color than that of the relevant object (but one of the remaining three colors), whereas a neutral word was one of three meaningless letter strings (xxxx, yyyy, or zzzz). Figure 1 shows examples of the stimuli.

The second factor was object condition. The words were presented either as parts of the relevant object or as parts of the irrelevant object. The two experimental factors, as well as the color and orientation of the front rectangle, varied randomly from trial to trial. The participants were presented with 48 repetitions for each of the six experimental conditions (3 congruency $\times$ 2 object condition), resulting in a total of 288 trials. Within the two object conditions, each of the 4 congruent displays was repeated 12 times, whereas

each of the 12 neutral and incongruent displays was repeated 4 times. The last block contained 8 trials instead of 10.

Results

Response times. For each participant, we first removed all vocal RTs exceeding three standard deviations from the mean. Across participants, RTs < 111 msec (0.04%) and RTs $> 1,044$ msec (1.44%) were excluded from further analyses. The mean RT values (upper panel) and error percentages (lower panel) for each of the six conditions are presented in Figure 2.

RTs were entered into a 3 (congruency) $\times$ 2 (object condition) repeated measures analysis of variance (ANOVA). RTs were shortest for congruent words (530 msec), intermediate for neutral words (561 msec), and longest for incongruent words [622 msec; $F(2,40) = 54.77$, $MS_e = 1,685.9$, $p < .001$]. Moreover, RTs were longer when words appeared in the relevant object (579 msec) than when they appeared in the irrelevant object [563 msec; $F(1,20) = 13.15$, $MS_e = 635.4$, $p < .01$]. Similarly, neutral stimuli in the relevant object (568 msec) caused longer RTs than did neutral stimuli in the irrelevant object [554 msec; $t(20) =$
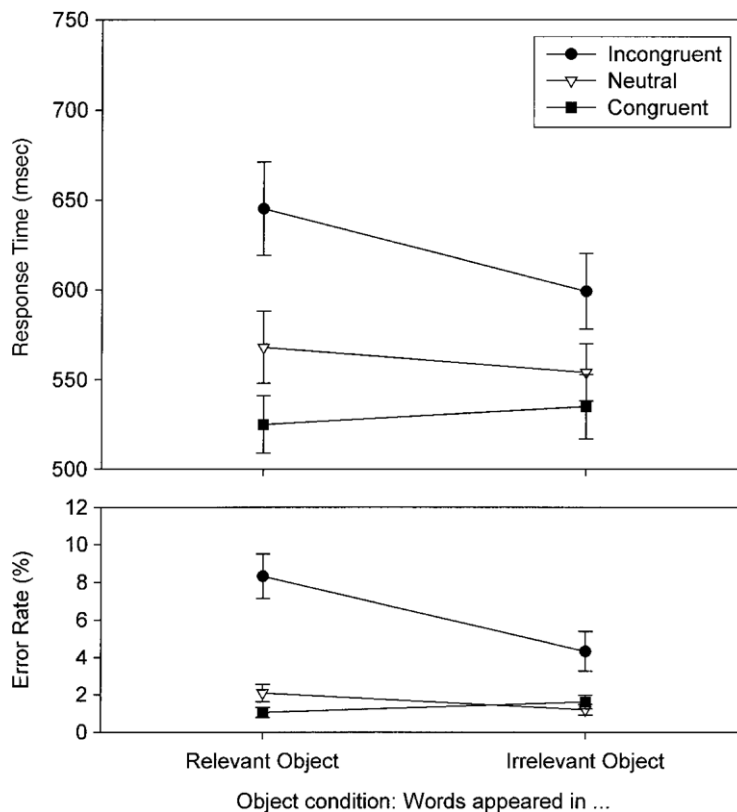

Figure 2. Experiment 1: mean color-naming response times (upper panel) and mean percentages of errors (lower panel) as a function of distractor type and object condition ($N = 21$; error bars are standard errors between the participants).

2.20, $p < .05$]. Most important, however, the interaction between congruency and object condition was also significant [$F(1,20) = 18.95, MS_e = 427.8, p < .001$].

To clarify the two-way interaction, facilitation scores (congruent minus neutral) and interference scores (incongruent minus neutral) were computed for each of the two object conditions. Then the facilitation scores from the two object conditions and the interference scores from the two object conditions were compared. For all planned pairwise comparisons reported in this article, Bonferroni-adjusted two-tailed t tests were used. Congruent words in the relevant object produced more facilitation (-43 msec) than did congruent words in the irrelevant object (-19 msec; $p < .01$). Similarly, incongruent words in the relevant object (77 msec) produced more interference than did incongruent words in the irrelevant object (46 msec; $p < .01$).

Errors. Error percentages (see the lower panel of Figure 2) were entered into a 3 (congruency) × 2 (object condition) repeated measures ANOVA. Results were similar to those for RTs. First, error percentages were lowest for congruent words (1.3%), intermediate for neutral words (1.6%), and highest for incongruent words [6.3%; $F(2,40) = 32.32, MS_e = 10.2, p < .001$]. Moreover, error percentages were higher when the words appeared in the relevant object (3.8%) than when they appeared in the irrelevant object [2.4%; $F(1,20) = 9.66, MS_e = 6.8, p < .01$]. Finally, the interaction between congruency and object condition was also significant [$F(1,20) = 6.80, MS_e = 8.3, p < .01$].

Facilitation scores from the two object conditions and interference scores from the two object conditions were compared, respectively. As for RTs, congruent words in the relevant object (-1.1%) produced more facilitation than did congruent words in the irrelevant object (0.4%; $p < .05$). Similarly, incongruent words in the relevant object (6.2%) produced more interference than did incongruent words in the irrelevant object (3.1%; $p < .01$).

Discussion

In Experiment 1, the participants named the color of a rectangular object that partially occluded a second, gray object. At the same time, the participants had to ignore color words that appeared either in the relevant colored object or in the irrelevant gray object. Two major results emerged. First, color words affected performance whether they appeared in the relevant or the irrelevant object and despite the fact that the words appeared approximately 1.1° from fixation. Congruent words facilitated and incongruent words impaired color naming. This observation is consistent with earlier findings (e.g., Brown et al., 2002) and confirms that color words are hard to ignore in the Stroop task. Second and more interesting, the irrelevant words produced larger Stroop effects, in terms of both facilitation and interference, when they occurred in the relevant object than when they occurred in the irrelevant object. Note that the spatial distance of the words from the fixation point was identical in both conditions.

Thus, these results support the view that object-based input selection is able to modulate the Stroop effect.

Unfortunately, Experiment 1 has two shortcomings. First, there is an additional difference between the relevant-object condition and the irrelevant-object condition of Experiment 1, which might be important. Whereas the color of the relevant object changed from trial to trial, the color of the irrelevant object (light gray) remained the same throughout the experiment. Therefore, the observed differences might have resulted from trials in which the contrast between the black words and the color of the relevant object (e.g., yellow) was higher than the contrast between the black words and the gray color of the irrelevant object, or it might have been easier to ignore the constantly colored irrelevant object. Second, Experiment 1 contained no condition in which the color words appeared in the background, and not in an object. Without such a comparison, it is unclear whether the effects found in Experiment 1 occurred because the processing of words as parts of the relevant object was amplified or because the processing of words as parts of the irrelevant object was attenuated. These two issues were addressed in Experiment 2.

EXPERIMENT 2

The purpose of Experiment 2 was to provide converging evidence for the results of Experiment 1, which indicated that object-based attention might be able to modulate the Stroop effect. There were two major changes in the procedure of Experiment 2, as compared with Experiment 1. First, in Experiment 2, the color of the relevant object, the color of the irrelevant object, and the color of the background changed from trial to trial. The task, however, remained the same: The participants still had to name the color of the occluding rectangle as quickly as possible and to ignore the words and the other colors present in the display. If the constant gray color of the irrelevant objects in Experiment 1 had affected the results of that experiment, Experiment 2 should reveal different results.

Second, Experiment 2 contained an additional condition in which the color words appeared in the empty background (see Figure 1C). Three patterns of results are possible with respect to the effects of color words in the background, in comparison with the effects of color words in the relevant- and the irrelevant-object conditions. First, color words in the background could have no effects at all. This outcome is unlikely, because of the results of earlier studies (e.g., Brown et al., 2002; Gatti & Egeth, 1978). Second, color words in the background could produce Stroop effects similar to those of color words in the irrelevant object, with color words in the relevant object producing even stronger effects. This outcome would suggest an object-based amplification in the processing of all the features of an attended perceptual object, as has been proposed by several authors (e.g., Duncan, 1984; Kahneman & Henik, 1981). Third, color words in the background could produce effects similar to those of color words in the rele-

vant object, with color words in the irrelevant object producing smaller effects. Such an outcome might also be taken as evidence for object-based attentional selection. In this case, however, object-based selection would become evident in the attenuation of the processing of the features of an irrelevant, potentially distracting object.

Method

Participants. Twenty-two new volunteers (14 female), with a mean age of 24 years, participated in a single-session experiment. The experiment lasted approximately 45 min, and the participants received DM 12 for attendance.

Apparatus and Stimuli. The apparatus was the same as that in Experiment 1. The stimuli were like those in Experiment 1, except for the following changes. First, in addition to the color of the relevant front rectangle, both the color of the irrelevant rear rectangle and the color of the background also changed from trial to trial. There were again four possible colors (blue, green, red, or yellow). However, in each display, the rectangles and the background had different colors. Thus, the colors of the irrelevant object and the background were always incongruent with respect to the to-be-named color of the relevant object. Moreover, an incongruent color word was incongruent with each of the three colors present in the display, whereas a congruent color word was congruent with the color of the relevant object but incongruent with the colors of the irrelevant object and the background.

Second, in addition to the relevant-object and the irrelevant-object conditions, there was a third condition in which two (identical)

words appeared in the background. In this case, the words appeared to the lower left and the upper right or to the upper left and the lower right of the two rectangles, respectively (see Figure 1C for an example). As in Experiment 1, the distance of the words from the fixation point was the same across the object conditions. Because of adding the words-in-the-background condition, the sizes of the rectangles and the distances between the words and the fixation point had to be increased, in comparison with Experiment 1. Thus, viewed from a distance of 80 cm, the short side of the rectangles subtended 1.4°, and the long side subtended 3.9°. The distance between the fixation point and the center of a word was 1.4° in each condition.

Procedure and Design. The procedure was the same as that in Experiment 1. The experiment was again run in participant-paced blocks of 10 trials each. The design was also the same as that in Experiment 1, except that there was an additional level of the factor object condition (words appearing in the background). This modification led to a 3 (object condition) $\times$ 3 (congruency) within-subjects design. The levels of these two factors, as well as the color and the orientation of the front rectangle, varied randomly from trial to trial. The participants were presented with 48 trials for each of the nine experimental conditions, resulting in a total of 432 trials. The last block contained 12 trials instead of 10.

Results

Response times. For each participant, we first removed all vocal RTs exceeding three standard deviations from the mean. Across participants, RTs < 163 msec (0.08%) and RTs $> 1,053$ msec (1.44%) were excluded from further
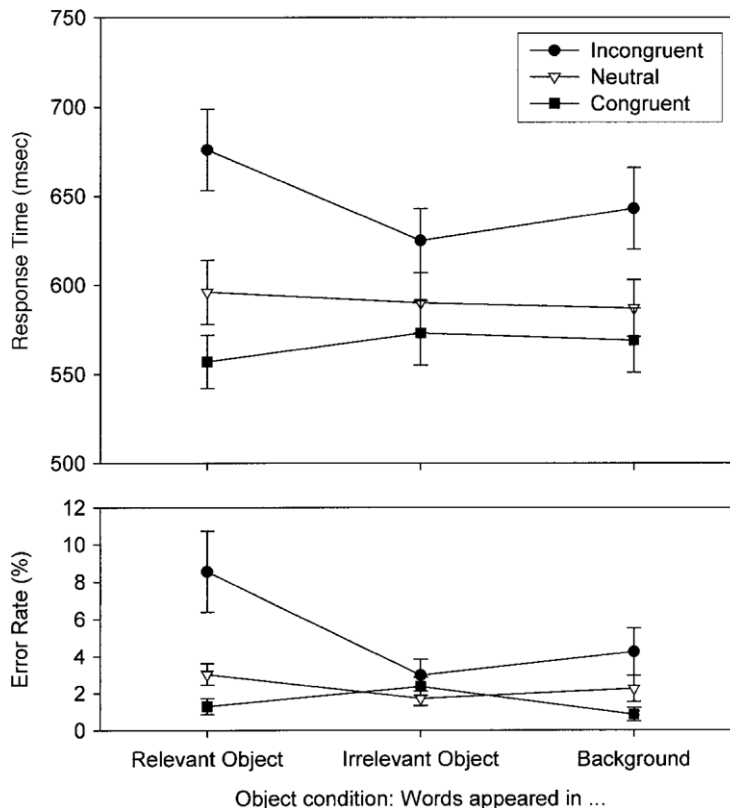

Figure 3. Experiment 2: mean color-naming response times (upper panel) and mean percentages of errors (lower panel) as a function of distractor type and object condition ($N = 22$; error bars are standard errors between the participants).

analyses. The mean RT values (upper panel) and error percentages (lower panel) for each of the nine conditions are presented in Figure 3.

RTs were entered into a 3 (congruency) × 3 (object condition) ANOVA. RTs were shortest for congruent words (566 msec), intermediate for neutral words (591 msec), and longest for incongruent words [648 msec; $F(2,42) = 81.24$, $MS_e = 1,415.7$, $p < .001$]. Moreover, RTs were longer when the words appeared in the relevant object (610 msec) than when they appeared in the irrelevant object (596 msec) or in the background [600 msec; $F(2,42) = 10.50$, $MS_e = 308.7$, $p < .001$]. An additional one-way ANOVA on the neutral conditions showed no effect of where the letter strings had appeared [$F(2,42) = 2.06$, $MS_e = 222.1$, $p = .14$]. Most important, however, the interaction between congruency and object condition was significant [$F(4,84) = 18.95$, $MS_e = 470.7$, $p < .001$].

Facilitation and interference scores in RTs were computed for the three object conditions and were compared. Congruent words in the relevant object produced more facilitation (-39 msec) than did congruent words in the irrelevant object (-17 msec; $p < .05$) and congruent words in the background (-18 msec; $p < .05$). The latter two conditions did not differ ($p > .90$). Similarly, incongruent words in the relevant object (80 msec) produced more interference than did incongruent words in the irrelevant object (35 msec; $p < .001$) and incongruent words in the background (56 msec; $p < .05$). The latter two conditions did not differ ($p > .10$).

Errors. Error percentages (see lower panel of Figure 3) were also entered into a 3 (congruency) × 3 (object condition) ANOVA. The results were similar to those of the RT analysis. Error percentages were lowest for congruent words (1.5%), intermediate for neutral words (2.3%), and highest for incongruent words [5.3%; $F(2,42) = 6.57$, $MS_e = 39.1$, $p < .01$]. Moreover, error percentages were higher when the words appeared in the relevant object (4.3%) than when they appeared in the irrelevant object (2.4%) or in the background [2.5%; $F(2,42) = 8.03$, $MS_e = 9.8$, $p < .01$]. Most important, the interaction between congruency and object condition was also significant [$F(4,84) = 7.15$, $MS_e = 9.3$, $p < .001$].

Congruent words in the relevant object produced more facilitation (-1.7%) than did congruent words in the irrelevant object (0.7%; $p < .05$), but not more than did congruent words in the background (-1.4%; $p > .90$). The latter two conditions did not differ ($p > .10$). Incongruent words in the relevant object (5.5%) failed to produce more interference than did incongruent words in the irrelevant object (1.3%; $p > .10$), but the former condition produced more interference than did incongruent words in the background (2.0%; $p < .05$). Again, interference scores for the irrelevant object and the background did not differ ($p > .90$).

Discussion

Experiment 2 revealed three major results. The first result was that irrelevant color words affected the ease of naming the color of a rectangular object whether the words appeared in the relevant object, the irrelevant object, or the background. In each of these conditions, there was both significant facilitation from congruent words and significant interference from incongruent words. This result is consistent with previous studies and confirms the difficulty of ignoring color words in the Stroop task.

The second major result of Experiment 2 was that both Stroop facilitation and Stroop interference were significantly larger, at least in RTs, when the congruent or the incongruent color words appeared in the relevant object than when they appeared in the irrelevant object. Thus, the results of Experiment 2 perfectly replicated the results of Experiment 1, despite several methodological differences. One such difference was that the irrelevant object changed its color from trial to trial in Experiment 2, but not in Experiment 1. This result confirms our hypothesis that object-based selection is effective in the Stroop task.

The third major result of Experiment 2 was that color words in the relevant object produced more facilitation and more interference, at least in RTs, than did color words that appeared in the background. Moreover, color words in the irrelevant object and color words in the background produced similar Stroop effects. These observations suggest that object-based attentional selection amplifies the processing of all the features of the relevant object. However, before these conclusions can be accepted, an alternative explanation for the larger Stroop effects in the relevant-object conditions of Experiments 1 and 2 must be tested.

EXPERIMENT 3

The aim of Experiment 3 was to differentiate between two explanations for the main result of Experiments 1 and 2—the larger Stroop effects in the relevant-object condition than in the irrelevant-object condition. Note that in Experiments 1 and 2, the relevant object was also the object that appeared closer in depth to the observer than did the irrelevant object, because the former object partially occluded the latter object. Therefore, there are two explanations for the larger Stroop effects in the relevant-object condition. According to the relevant-object account, color words are more effective when they are part of the relevant object, because attentional selection of the relevant object amplifies processing of all its features. In contrast, according to the nearer-object account, color words are more effective when they are part of the object that appears to be closer to the observer, because of an attentional gradient in depth (Downing & Pinker, 1985).

Several studies have revealed evidence that participants can focus attention on a particular depth plane when depth is defined by binocular disparity (e.g., Andersen, 1990; Theeuwes, Atchley, & Kramer, 1998). In Theeuwes et al.'s study, participants were faster to detect and process a target element among distractor stimuli when the target element appeared in the validly cued depth plane than when it appeared in the invalidly cued depth plane. Moreover,

Andersen observed that response-incompatible stimuli presented at a different depth plane than the target still interfered with target processing. Most important for the present study, however, was Andersen's observation that response-incompatible distractors produced less interference when they appeared closer to the observer than the target, as compared with distractors that appeared at the same depth plane as the target.

In Experiment 3, the same displays were used as in Experiment 2, but the instruction was changed. The participants had to name the color of the occluded object as quickly as possible. Thus, in contrast to Experiments 1 and 2, the relevant object appeared to be farther away than the irrelevant object. The predictions are clear. If the larger effects in the relevant-object conditions of Experiments 1 and 2 were due to object relevance, interference should also be largest in the relevant-object condition of Experiment 3. If, however, the larger effects were due to perceived distance, interference should be largest in the irrelevant-object condition of Experiment 3.

Another change in Experiment 3 concerned the distractor stimuli. In Experiments 1 and 2, the color words were congruent with the color of the relevant object in one third of the trials. There is evidence that the presence of congruent trials in the Stroop task might lead participants to divide their attention over the two stimulus dimensions (color and word meaning), at least on some trials (e.g., Logan, Zbrodoff, & Williamson, 1984; Lowe & Mitterer, 1982). Because the task of naming the correct color in our experiments was difficult, it is possible that the participants sometimes attended to the color word, which provides the correct answer on one third of the trials and impairs performance only on one third of the trials. Such a strategy might be related to our observation of a larger effect of color words in the relevant object. To test whether the presence of congruent stimuli in Experiments 1 and 2 had an effect, there were no congruent trials in Experiment 3.

Method

Participants. Twenty-one new volunteers (13 female), with a mean age of 24 years, participated in a single-session experiment. The experiment lasted approximately 45 min, and the participants received DM 12 for attendance.

Apparatus and Stimuli. The apparatus and stimuli were the same as those in Experiment 2 (see Figure 1C for an example stimulus), except that the words were never congruent to the color of the relevant rectangle and there were twice as many incongruent trials as neutral trials.

Procedure. The procedure in Experiment 3 was the same as that in Experiment 2; only the instructions were changed. Instead of
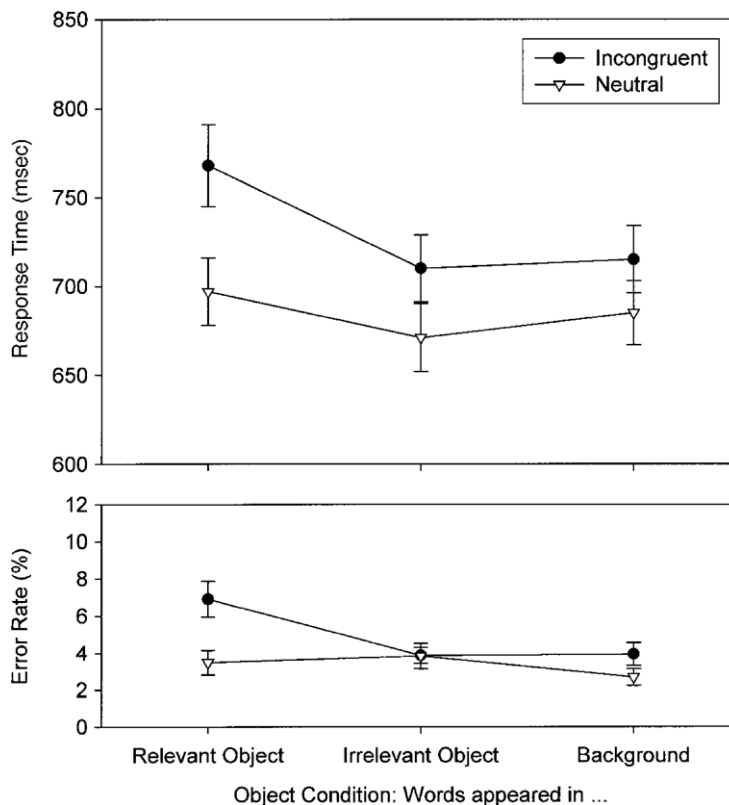

Figure 4. Experiment 3: mean color-naming response times (upper panel) and mean percentages of errors (lower panel) as a function of distractor type and object condition ($N = 21$; error bars are standard errors between the participants).

being required to name the color of the front rectangle, the participants were instructed to name the color of the rear rectangle.

Design. The design was the same as that in Experiment 2, except that there was no congruent condition, resulting in a 3 (object condition) × 2 (congruency) within-subjects design. The levels of these two factors, as well as the color and the orientation of the relevant rear rectangle, varied randomly from trial to trial. The participants were presented with 48 trials for each neutral condition and with 96 trials for each incongruent condition. The total number of trials was 432. The last block contained 12 trials instead of 10.

Results

Response times. Vocal RTs exceeding three standard deviations from the mean were removed. Across participants, RTs < 162 msec (0.08%) and RTs > 1,285 msec (1.34%) were excluded from further analyses. The mean RT values (upper panel) and error percentages (lower panel) for each of the six conditions are presented in Figure 4.

RTs were entered into a 2 (congruency) × 3 (object condition) ANOVA. RTs were longer for incongruent words (731 msec) than for neutral words [684 msec; $F(1,20) = 63.4, MS_e = 1,093.5, p < .001$]. Moreover, RTs were longer when the words appeared in the relevant object (734 msec) than when they appeared in the irrelevant object (691 msec) or in the background [700 msec; $F(2,40) = 28.47, MS_e = 722.4, p < .001$]. The affiliation of letter strings also had an effect on RTs in neutral conditions only [$F(2,40) = 6.20, MS_e = 579.8, p < .01$]. Most important, however, the interaction between congruency and object condition was significant [$F(2,40) = 10.23, MS_e = 489.2, p < .001$].

For each object condition, interference scores were computed from RTs. Incongruent words in the relevant object (72 msec) produced more interference than did incongruent words in the irrelevant object (40 msec; $p < .05$) and incongruent words in the background (30 msec; $p < .01$). The latter two conditions did not differ ($p > .30$).

Errors. Error percentages (see the lower panel of Figure 4) were also entered into a 2 (congruency) × 3 (object condition) ANOVA. The results were similar to those of the RT analysis. Error percentages were larger for incongruent words (4.9%) than for neutral words [3.3%; $F(1,20) = 14.15, MS_e = 5.5, p < .01$]. Moreover, error percentages were larger when the words appeared in the relevant object (5.2%) than when they appeared in the irrelevant object (3.9%) or in the background [3.3%; $F(2,40) = 8.33, MS_e = 4.8, p < .01$]. The interaction between congruency and object condition was also significant [$F(2,40) = 7.58, MS_e = 4.1, p < .01$].

For each object condition, interference scores were computed from error percentages. Incongruent words in the relevant object (3.4%) produced more interference than did incongruent words in the irrelevant object (0.3%; $p < .05$) and incongruent words in the background (1.2%; $p < .05$). The latter two conditions did not differ ($p > .40$).

Discussion

The results of Experiment 3 support the relevant-object account and refute the nearer-object account. Incongruent color words in the relevant object interfered more strongly with color-naming performance than did incongruent color words in the irrelevant object, although occlusion produced the impression that the relevant object was more distant from the observer than the irrelevant object. Moreover, incongruent color words in the background, which might be perceived as being most distant from the observer, produced as much interference as did incongruent color words in the irrelevant occluding object, which might be perceived as being closest to the observer. Thus, independent of their apparent position in depth, words that belong to the task-relevant object more strongly affect performance than do words that belong to an irrelevant object or words that appear in the background. Together, the results support the conclusion that attention selects the relevant object from the display and this selection amplifies processing of all the features of the attended object (Duncan, 1984; Kahneman & Henik, 1981).

In contrast to Experiments 1 and 2, Experiment 3 contained no congruent conditions, and the percentage of incongruent trials (67%) was twice as large as the percentage of neutral trials (33%). This was done to discourage the participants from paying attention to the color words, which might have happened in Experiments 1 and 2. Despite this difference, the results in Experiment 3 were very similar to the results from the incongruent condition in Experiment 2. This similarity suggests that amplified processing of the (incongruent) color words, when they are part of the attended relevant object, is hard to avoid.

GENERAL DISCUSSION

In this study, we investigated whether object-based mechanisms of input selection can modulate the Stroop effect. The participants saw two overlapping colored rectangles presented at screen center (see Figure 1). The task was to name the color of one of the rectangles—for example, the rectangle that was in front of the other—and to ignore everything else. Congruent, neutral, or incongruent color words appeared as parts of the relevant object, as parts of the irrelevant object, or in the background. However, the distance between the words and the fixation point was identical in each condition. The question was whether congruent and/or incongruent color words in the relevant object would produce larger Stroop effects than would color words in the irrelevant object or in the background.

Four major results were observed. First, color words produced Stroop effects in each condition—that is, regardless of whether they occurred as parts of the relevant object, as parts of the irrelevant object, or in the background. This result agrees with those of previous studies and demonstrates that color words are hard to ignore in the Stroop task. Second, congruent words produced more facilitation, and incongruent words produced more interference, when they appeared as parts of the relevant object. In fact, Stroop effects were approximately twice as large in this condition as in the remaining conditions. We interpret this result as a consequence of object-based input selection. Third, words in the irrelevant object and in the background

produced similar Stroop effects. This result suggests that the effect of object-based selection mainly is to amplify the processing of the features of the selected object, rather than to attenuate the processing of irrelevant objects. Fourth, the increased Stroop effects in the relevant-object condition occurred whether the relevant object occluded the irrelevant object or vice versa. Thus, these effects were determined by the task relevance of an object, and not by its apparent relative distance in depth.

The most important result of the present study is that both Stroop facilitation and Stroop interference were significantly larger when the congruent or the incongruent words were part of the relevant object. This observation supports the conclusions (1) that object-based input selection can modulate the Stroop effect to a significant degree and (2) that object-based input selection is associated with amplification in the processing of all of the features of the selected object (Duncan, 1984; Kahneman & Henik, 1981). The amplification in the processing of irrelevant features of a selected object increases benefits if the irrelevant features are response congruent and increases costs if the irrelevant features are response incongruent.

The present study adds to existing knowledge about the conditions that will increase or decrease the Stroop effect. However, the study provides no clues about the sources of facilitation and interference and does also not tell us how the cognitive system solves the problems imposed by incongruent stimuli. Most probably, the main problems in the Stroop task do not arise at the level of input selection, but at the level of dimensional selection (cf. Melara & Mounts, 1993) and the level of response selection (cf. Lu & Proctor, 1995). However, mechanisms of input selection are clearly able to modulate the size of the problems that arise at the latter two stages of processing.

Integrated Versus Separated Stroop Tasks

An important question concerning the Stroop effect relates to the nature of the differences between the integrated and the separated versions of the Stroop task (MacLeod, 1991, 1998). To what extent does the degree of melding of the two dimensions influence the extent of facilitation and interference? This question must be addressed in order to be able to compare different versions of the Stroop task meaningfully, such as the (integrated) color–word task and the (separated) picture–word task (Glaser & Düngelhoff, 1984).

Previous studies have shown that integrated Stroop tasks produce more interference than do separated tasks (e.g., Kahneman & Henik, 1981; MacLeod, 1998). Unfortunately, in these studies, variations in task type were correlated with variations in the spatial distance between color and word. Because an increase in spatial distance per se can reduce the Stroop effect (e.g., Gatti & Egeth, 1978), it is not clear whether larger Stroop effects in integrated tasks, as compared with separated tasks, only reflect quantitative differences in spatial distance, which might be related to gradients in visual acuity or gradients of spatial attention. Alternatively, larger Stroop effects in integrated tasks, as compared with separated tasks, might reflect qualitatively different processing of features of a relevant object, as compared with processing of features of an irrelevant object.

The results of the present study seem to support the second hypothesis. Our condition in which words appeared in the relevant object resembles an integrated task, whereas our conditions in which words appeared in the irrelevant object or in the background resemble separated tasks. We found larger Stroop effects in our integrated condition than in our separated conditions, and we interpret this result as evidence for the hypothesis that processing in integrated Stroop tasks is qualitatively different from processing in separated tasks. In the case of our separated tasks, irrelevant words could not be ignored because they fell within the focus of spatial attention or because their recognition was highly practiced. In the case of our integrated task, processing of the words was amplified because they belonged to the task-relevant (i.e., selected) object.

In the introduction, we argued that many of the problems a perceiver-actor is faced with in our normal environment also occur in the Stroop task. These problems involve input selection, dimensional or analyzer selection, and response selection. Most of the previous studies have focused on the latter two problems, whereas input selection was considered to play a minor role in the Stroop task (e.g., Shalev & Algom, 2000). The results of the present study, however, suggest that object-based input selection plays an important role, at least in integrated Stroop tasks. This implies that, to develop a complete understanding of the Stroop effect, each of the three problems just described has to be considered.

REFERENCES

Allport, D. A. (1971). Parallel encoding within and between elementary stimulus dimensions. *Perception & Psychophysics*, **10**, 104-108.

Andersen, G. J. (1990). Focused attention in three-dimensional space. *Perception & Psychophysics*, **47**, 112-120.

Bauer, B., & Besner, D. (1997). Processing in the Stroop task: Mental set as a determinant of performance. *Canadian Journal of Experimental Psychology*, **51**, 61-68.

Brown, T. L., Gore, C. L., & Carr, T. H. (2002). Visual attention and word recognition in Stroop color naming: Is word recognition "automatic?" *Journal of Experimental Psychology: General*, **131**, 220-240.

Dishon-Berkovits, M., & Algom, D. (2000). The Stroop effect: It is not the robust phenomenon that you have thought it to be. *Memory & Cognition*, **28**, 1437-1449.

Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and Performance XI* (pp. 171-187). Hillsdale, NJ: Erlbaum.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, **113**, 501-517.

Dyer, F. N. (1973a). Interference and facilitation for color naming with separate bilateral presentations of the word and color. *Journal of Experimental Psychology*, **99**, 314-317.

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

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, **40**, 225-240.

GATTI, S. V., & EGETH, H. E. (1978). Failure of spatial selectivity in vision. *Bulletin of the Psychonomic Society*, **11**, 181-184.

GLASER, W. R., & DÜNGELHOFF, F.-J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 640-654.

HAGENAAR, R., & VAN DER HEIJDEN, A. H. (1986). Target–noise separation in visual selective attention. *Acta Psychologica*, **62**, 161-176.

KAHNEMAN, D., & HENIK, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 181-211). Hillsdale, NJ: Erlbaum.

LABERGE, D., & BROWN, V. (1989). Theory of attentional operations in shape identification. *Psychological Review*, **96**, 101-124.

LOGAN, G. D., ZBRODOFF, N. J., & WILLIAMSON, J. (1984). Strategies in the color–word Stroop task. *Bulletin of the Psychonomic Society*, **22**, 135-138.

LOWE, D. G., & MITTERER, J. O. (1982). Selective and divided attention in a Stroop task. *Canadian Journal of Psychology*, **36**, 684-700.

LU, C.-H., & PROCTOR, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review*, **2**, 174-207.

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

MACLEOD, C. M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. *Memory & Cognition*, **26**, 201-211.

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.

MERIKLE, P. M., & GOREWICH, N. J. (1979). Spatial selectivity in vision: Field size depends upon noise size. *Bulletin of the Psychonomic Society*, **14**, 343-346.

POSNER, M. I., & SNYDER, C. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55-85). Hillsdale, NJ: Erlbaum.

POSNER, M. I., SNYDER, C. R., & DAVIDSON, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, **109**, 160-174.

SHALEV, L., & ALGOM, D. (2000). Stroop and Garner effects in and out of Posner's beam: Reconciling two conceptions of selective attention. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 997-1017.

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

THEEUWES, J., ATCHLEY, P., & KRAMER, A. F. (1998). Attentional control within 3-D space. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1476-1485.

TREISMAN, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, **76**, 282-299.

VAN DER HEIJDEN, A. H. (1992). *Selective attention in vision*. London: Routledge.

VAN DER HEIJDEN, A. H., HAGENAAR, R., & BLOEM, W. (1984). Two stages in postcategorical filtering and selection. *Memory & Cognition*, **12**, 458-469.

VECERA, S. P., & FARAH, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, **123**, 146-160.

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

WEBER, T. A., KRAMER, A. F., & MILLER, G. A. (1997). Selective processing of superimposed objects: An electrophysiological analysis of object-based attentional selection. *Biological Psychology*, **45**, 159-182.

YANTIS, S., & JONIDES, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 601-621.