Depth cues do not underlie attentional modulations of the Stroop effect

PETER WÜHR and MARTINA WELTLE

Friedrich-Alexander Universität, Erlangen, Germany

The well-known Stroop effect is usually attributed to the automaticity of word reading. Recently, Wühr and Waszak (2003) had participants name the color of one of two rectangles and found that words in the relevant object produced larger Stroop effects than did words in the irrelevant object or in the background. They attributed this difference to an object-based mechanism of attentional selection that amplifies processing of all the features of an attended object. However, in the displays used by Wühr and Waszak, occlusion suggested the presence of different depth planes. Hence, the increased Stroop effect could have resulted from perceiving the words to be in the same depth plane as the relevant object and not from perceiving the words to be parts of the relevant object. Two experiments tested between these accounts by using displays without monocular depth cues. The results of both experiments replicate those of Wühr and Waszak, supporting their object-based account.

The Stroop effect (Stroop, 1935) is one of the best-known phenomena in experimental psychology. In the usual Stroop task, participants are presented with color words in different colors (e.g., the word RED in green ink). When participants have to report the ink color, irrelevant color words produce interference in comparison with a neutral condition (e.g., the letter string XXX in green ink); congruent color words improve performance (see MacLeod, 1991, for a review). The most popular explanation for the Stroop effect holds that word reading is automatic and therefore hard to avoid (see Brown, Gore, & Carr, 2002; MacLeod & MacDonald, 2000; Posner & Snyder, 1975; but see Melara & Algom, 2003, for a different account).

However, the results of a recent study by Wühr and Waszak (2003) suggest that the automaticity of word reading is not the only reason for the robustness of the Stroop effect, at least in integrated Stroop tasks. Rather, it appears that word shape is not only processed because word reading *per se* is automatic, but also because the shape belongs to the attended perceptual object (see Kahneman & Henik, 1981). Participants in Wühr and Waszak's study had to report the color of a relevant object that partially occluded an irrelevant object (see Figure 1). To provoke Stroop effects, displays contained congruent, neutral, or incongruent color words. Most importantly, these color words were presented as part of the relevant object (Figure 1A), as part of the irrelevant object (Figure 1B), or in

the background (Figure 1C). Because the spatial distance between the fixation point and the color words was identical in each condition, a spotlight-like mechanism of selective attention (see Posner, 1980) would not predict any differences between these conditions. Yet Wühr and Waszak observed greater facilitation from congruent words and greater interference from incongruent words when the words were part of the relevant object than when they were part of the irrelevant object or appeared in the background. The latter two conditions did not differ. The same result was obtained when participants had to report the color of the occluded rectangle.

Wühr and Waszak (2003) explained these results in terms of an object-based mechanism of selective attention, as suggested by Duncan (1984) and Kahneman (e.g., Kahneman & Henik, 1981; Kahneman & Treisman, 1984). According to this account, preattentive processes segment the visual field into candidate objects (figures) and background, and attention selects among the candidate objects. Importantly, selection of a particular object amplifies processing of all of its features, regardless of their task relevance. The results of the Wühr and Waszak study imply that there are not only quantitative but also qualitative differences between processing in separated and processing in integrated Stroop tasks.

Yet, there is an alternative explanation for the results of Wühr and Waszak (2003). The displays in that study contained a monocular depth cue—occlusion (see Howard, 2002). One rectangle occluded a second rectangle, which occluded the background. These displays may have produced the impression in the observer of at least three depth planes. Hence, it is possible that observers did not perceive the words as being part of an object, but only as being located in the same depth plane. If we further assume that observers can focus spatial attention upon a particular depth plane and that distractors located in the

This research was supported by a grant from the German Science Foundation to P.W. (WU-357/1-1). We thank Benedikt Viedenz for collecting the data. Moreover, we are grateful to Daniel Algom, Derek Besner, and Charles Folk for helpful comments on a previous draft. Correspondence concerning this article should be addressed to P. Wühr, Institut für Psychologie I, Friedrich-Alexander Universität Erlangen-Nürnberg, Kochstrasse 4, 91054 Erlangen, Germany (e-mail: prwuehr@phil.uni-erlangen.de).

Figure 1. Panels A–C show examples of displays used in the experiments of Würh and Waszak (2003). Panels D–F show examples of displays used in the present experiments. Congruent or incongruent color words appeared in the relevant object (A, D), in the irrelevant objects (B, E), or in the background (C, F).

same depth plane as a target produce the most interference, we get an alternative explanation for the results.

The main purpose of the present study was to replicate the findings of Würh and Waszak (2003) while testing the alternative depth explanation for their results. Therefore, we used displays in which the impression of the objects being located in different depth planes was eliminated (see Figures 1D–1F). In the new displays, two irrelevant objects appear in the same depth plane as the relevant object. If the alternative explanation is correct, words in the relevant and irrelevant objects should produce comparable Stroop effects in the new displays, whereas words in the background should still produce much smaller Stroop effects.

EXPERIMENT 1

Experiment 1 tested the object-based explanation for the results of Würh and Waszak (2003) against a depth-based explanation. In order to compare these accounts, we developed a new set of displays (see Figures 1D and 1E for examples). These displays contained three rectangles that did not occlude each other. The central rectangle was the relevant object, the color of which had to be named. The two flanking rectangles were irrelevant objects and were presented in a different color than the relevant object. Importantly, the three objects in the new displays appeared to be located in the same depth plane because they did not occlude each other. Accordingly,

the depth-based account predicts that color words in the relevant object and in the irrelevant objects should produce comparable Stroop effects, which should be larger than the effects produced by words in the background. In contrast, the object-based account predicts that color words in the relevant object should produce larger Stroop effects than words in the irrelevant objects and in the background. After the experiment, participants were asked about their perception of the displays.

There were two further differences between the present experiments and those of Wühr and Waszak (2003). First, in order to measure the processing of irrelevant words, only congruent and incongruent conditions were compared in the present experiments. Second, in order to discourage participants from deliberately attending to the color words, the words and the colors of the relevant object were not correlated in the present experiments¹ (see Melara & Algom, 2003, for a discussion of this issue).

Method

Participants. Sixteen volunteers (7 female) with a mean age of 24 years participated in a single-session experiment. The experiment lasted about 45 min, and the participants received €6 for their attendance. All participants in the present study were native German speakers who classed themselves as having normal (or corrected-to-normal) visual acuity. The participants' color vision was tested with the Ishihara Tests for Colour Blindness (15th ed.; Ishihara, 1960).

Apparatus and Stimuli. The participants sat in front of a 17-in. color monitor. Viewing distance was constrained to be approximately 50 cm by a chin- and headrest. The participants responded by speaking into a microphone, which triggered a voice key measuring response times (RTs) to the nearest millisecond. An IBM-compatible computer controlled the presentation of stimuli and collected vocal RTs.

The fixation cross was a small "+" sign subtending 0.5° of visual angle. Each stimulus display consisted of three rectangles (see, e.g., Figure 1D). A large rectangle (2.2° × 7.8°), which was centered on the fixation point, was flanked by two small rectangles (2.2° × 2.4°). The orientation of the large rectangle was balanced across the experimental conditions. The small rectangles appeared to the left and right of fixation when the large rectangle was oriented vertically or above and below fixation when the large rectangle was oriented horizontally. The distance between the edges of the large and the small rectangles was 0.2°.

The background, the two small rectangles, and the large rectangle could each have one of four different colors (blue, green, red, or yellow). In a particular trial, the background and the large rectangle appeared in two different colors, and the two small rectangles appeared in a third color. Finally, in each stimulus display, two identical color words (the German words for blue [*blau*], green [*grün*], red [*rot*], or yellow [*gelb*]) appeared in black in either the large rectangle, the small rectangles, or the background. In each case, the two words were located on an imaginary circle around the fixation point with a radius of 3° (see Figure 1). Thus, in the large rectangle, the words were located at the short edges. In the small rectangles, one word was centered in each rectangle. In the background, the two words appeared either to the lower left and upper right or to the upper left and lower right of fixation. The color words subtended from 1.1° (*rot*) to 1.4° (*gelb*).

Procedure. At the beginning of the session, the participants were shown written instructions on the screen. They were instructed to fixate the fixation point and to verbally report the color of the

large rectangle as quickly and accurately as possible. Moreover, the instructions pointed out that the words presented in the display were irrelevant with respect to the task; the participants were told to ignore them.

The experiment was run in 20 blocks. At the beginning of each block, the prompt *Press a key* appeared. The keypress started a block of trials, each of which contained the following events: First, the fixation cross appeared on the screen for 500 msec. After a blank period of 50 msec, the stimulus display was presented for 200 msec. Presentation time was short in order to prevent eye movements during the presentation. Then, a blank interval occurred until the participant's response triggered the next trial. After the participant had responded to the last display of a block, *Press a key* appeared once again, and the participant could take a rest. The participants' performance was monitored on line by the experimenter, who sat in an adjacent room. The experimenter heard the participants' responses via earphones and compared them with the correct answers that were shown on a list. Each error was recorded in that list.

At the end of the experimental session, each participant was shown a sample display (without color words) and asked two questions regarding perception of the display. The first question concerned how many objects the participant saw in the display, and the second was whether the objects appeared to be at the same distance or at different distances.

Design. The experiment was based on a 2 × 3 within-subjects design. The first factor was congruency: The irrelevant words were congruent or incongruent with respect to the correct color naming response. A congruent word denoted the color of the relevant object, whereas an incongruent word denoted a color that was not present in the display (but was a member of the response set). There were 25% congruent trials and 75% incongruent trials; hence, the color words and the color of the large rectangle were uncorrelated (but see note 1). The second factor was object condition of the color words: The words were presented in the relevant object, in the two irrelevant objects, or in the background. The two experimental factors, the colors of the rectangles and the background, and the orientation of the rectangles varied randomly from trial to trial. Repetition of the same response more than four times consecutively was not permitted. There were 32 congruent and 96 incongruent trials for each object condition, resulting in a total of 384 experimental trials, plus 20 practice trials. The last block contained 24 trials instead of 20.

Results

Fourteen of the 16 participants reported seeing three objects at the same distance. Two participants reported seeing four objects (including the fixation point) at different distances. Excluding the latter 2 participants from the analyses did not alter the results.

Response times. To eliminate outliers, for each participant we removed all vocal RTs exceeding two standard deviations from the mean. Across participants, RTs < 344 msec (0.1%) or > 952 msec (4.1%) were excluded from further analyses. Means of RTs and error percentages are presented in Table 1.

RTs were entered into a 2 × 3 repeated measures analysis of variance (ANOVA). The significant main effect of congruency indicated a Stroop effect [$F(1,15) = 54.32$, $MS_e = 578.56$, $p < .001$]. RTs were shorter with congruent words (603 msec) than with incongruent words (639 msec). The significant main effect of object condition [$F(2,30) = 8.71$, $MS_e = 543.16$, $p < .01$] indicated words in the relevant object (635 msec) elicited longer

Table 1
Mean Color Naming Response Times (RT, in Milliseconds) and
Mean Percentages of Error (%E) Observed in Experiments 1
and 2, as Functions of Color–Word Correspondence and
Object Condition

	Words Appeared in . . .					
	Relevant Object		Irrelevant Objects		Background	
	RT	%E	RT	%E	RT	%E
Experiment 1						
Corresponding	605	0.2	598	1.4	606	0.8
Noncorresponding	664	1.9	623	0.8	631	1.6
Experiment 2						
Corresponding	625	0.8	623	2.4	620	1.3
Noncorresponding	693	2.1	644	1.6	643	1.4

RTs than did words in the irrelevant objects (611 msec) or in the background (619 msec). Most importantly, the two-way interaction was also significant [$F(2,30) = 7.44$, $MS_e = 397.21$, $p < .01$]; this interaction indicated that words in the relevant object produced larger Stroop effects ($D = 58$ msec) than did either words in the irrelevant objects [$D = 25$ msec; $t(15) = 3.41$, $p < .01$] or in the background [$D = 25$ msec; $t(15) = 3.21$, $p < .01$]. Stroop effects in the latter two conditions did not differ [$t(15) = 0.05$, $p = .97$].

Error percentages. The results of the ANOVA on error percentages paralleled the RT results. An almost significant main effect of congruency suggested a Stroop effect [$F(1,15) = 4.51$, $MS_e = 2.13$, $p = .051$]. There were fewer errors with congruent words (0.8%) than with incongruent words (1.4%). The main effect of object condition was not significant ($F < 1$), but the two-way interaction was [$F(2,30) = 5.16$, $MS_e = 2.13$, $p < .05$]. The interaction indicated that words in the relevant object produced a larger Stroop effect in errors ($D = 1.7\%$) than did words in the irrelevant objects [$D = -0.6\%$; $t(15) = 3.09$, $p < .01$] and a numerically larger Stroop effect than did words in the background [$D = 0.8\%$; $t(15) = 1.70$, $p = .11$]. Stroop effects in the latter two conditions did not differ [$t(15) = 1.65$, $p = .12$].

Discussion

The results of Experiment 1 almost perfectly replicate those of Wühr and Waszak (2003). Color words in the relevant object produced substantially larger Stroop effects than did words in the irrelevant objects or in the background, and the Stroop effects in the latter two conditions did not differ. This result disconfirms the depth-based account, according to which color words should have the largest effects when they are perceived in approximately the same depth plane as the target stimulus, which was true for the words in both the relevant and irrelevant objects. The results agree, however, with the object-based account, according to which color words should have the largest effects when they are perceived as being part of the relevant object (see Kahneman & Henik, 1981). Moreover, the agreement between the present results and those

of Wühr and Waszak suggests that correlation between the words and the colors does not affect the results.

EXPERIMENT 2

The goal of Experiment 2 was to replicate the results of Experiment 1 under improved conditions. Because the displays were presented very briefly and the irrelevant objects had the same color in Experiment 1, it was possible that participants (at least sometimes) overlooked the separation between the objects and perceived the relevant object as occluding an irrelevant object, as in the Wühr and Waszak (2003) study. Admittedly, after the experiment, most participants reported having seen three objects at the same distance. However, during the interview, participants could look at a sample display without temporal restrictions, which might have revealed a different percept than they experienced during the experiment. To avoid these problems, the two irrelevant objects appeared in different colors in Experiment 2 (see Figure 1F for an example). Therefore, a fifth color (brown) was included in the design of this experiment. Presenting the two irrelevant objects in different colors should prevent them from being perceived as parts of a single occluded rectangle.

Method

Participants. Sixteen new volunteers (6 female) with a mean age of 25 years participated in a single-session experiment.

Apparatus, Stimuli, and Procedure. The apparatus, stimuli, and procedure of Experiment 2 were the same as those used in Experiment 1, except for two differences. First, one additional color (brown) was used in Experiment 2, and second, the two irrelevant rectangles had different colors. The color words subtended between 1.1° (*rot*) and 1.7° (*braun*, German for brown).

Design. Experiment 2 used the same within-subjects design as Experiment 1. Repetitions of the same response for more than three consecutive trials were not permitted.

Results and Discussion

Again, 14 of the 16 participants reported seeing three objects at the same distance. Two participants reported seeing three objects at varying distances (e.g., “yellow objects appear closer”). Excluding the latter 2 participants from the analyses did not alter the results.

Response times. Across participants, RTs < 358 msec (0.2%) and > 990 msec (4.9%) were excluded from further analyses. Table 1 presents the mean RTs for each condition.

RTs were entered into a 2×3 repeated measures ANOVA. The significant main effect of congruency indicated a Stroop effect [$F(1,15) = 50.60$, $MS_e = 661.11$, $p < .001$]. RTs were shorter with congruent words (623 msec) than with incongruent words (660 msec). The significant main effect of object condition [$F(2,30) = 23.51$, $MS_e = 312.33$, $p < .001$] indicated longer RTs with words in the relevant object (659 msec) than with words in the irrelevant objects (633 msec) or in the background (632 msec). Most importantly, the interaction was again significant

[$F(2,30) = 16.36$, $MS_e = 354.71$, $p < .001$], indicating that words in the relevant object produced a larger Stroop effect ($D = 68$ msec) than did words in the irrelevant objects [$D = 20$ msec; $t(15) = 4.17$, $p < .01$] or in the background [$D = 23$ msec; $t(15) = 4.66$, $p < .001$]. Stroop effects in the latter two conditions did not differ [$t(15) = 0.52$, $p = .61$].

These results perfectly replicate those of Experiment 1, despite the fact that the two irrelevant objects had different colors in Experiment 2. Presenting the irrelevant objects in different colors should have reduced the probability of perceiving them as parts of one object that was occluded by the relevant object. Hence, the results of Experiment 2 refute the depth-based account and are consistent with an account in terms of object-based attentional selection.

Error percentages. There were no significant effects (all F s < 1.5 , all p s $> .25$).

GENERAL DISCUSSION

The present study investigated possible explanations for a modulation of the Stroop effect reported by Wühr and Waszak (2003). In their study, words in a relevant object produced much larger Stroop effects than did words in an irrelevant object, which was partly occluded by the relevant object, or words in the background. According to these authors, the modulation arose from object-based attentional selection of the relevant object, which also boosted processing of irrelevant features of the relevant object (the object-based account). Alternatively, due to occlusion, the increased Stroop effect could have resulted from perceiving the words in the same depth plane as the relevant object and not from perceiving them to be part of the relevant object (the depth-based account). Two experiments tested between these accounts by using displays without monocular depth cues (e.g., Figures 1D–1F), and both experiments replicated the modulation of the Stroop effect. Hence, the results of the present study refute the depth-based theory and are compatible with an object-based account.

Our account holds that object-based attentional selection boosts processing of words that are part of the relevant object. Yet, alternative accounts are conceivable. For example, it is possible that word processing is unaffected by where the words are presented, but that presenting the words outside the relevant object helps keep the products of processing relevant and irrelevant information separate, and in so doing reduces the size of the Stroop effect (see Manwell, Roberts, & Besner, 2004). Moreover, the results of the present study are in accord with the tectonic theory of Stroop effects proposed by

Melara and Algom (2003). Their theory rests on the idea that an observer's efforts to attend selectively to relevant information are counteracted by the tendency to integrate new information. Irrelevant information will be processed if it is either salient and/or correlated with the relevant information. Words that are part of the relevant object may be more salient than words presented elsewhere, and this difference in salience may modulate the Stroop effect.

REFERENCES

- 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.
- DUNCAN, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, **113**, 501-517.
- HOWARD, I. P. (2002). Depth perception. In H. Pashler & S. Yantis (Eds.), *Stevens' Handbook of experimental psychology: 1. Sensation and perception* (3rd ed., pp. 77-120). New York: Wiley.
- ISHIHARA, S. (1960). *Tests for colour blindness* (15th ed.). Tokyo: Kanehara.
- KAHNEMAN, D., & HENIK, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 181-211). Hillsdale, NJ: Erlbaum.
- KAHNEMAN, D., & TREISMAN, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29-61). Orlando, FL: Academic Press.
- MACLEOD, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, **109**, 163-203.
- MACLEOD, C. M., & MACDONALD, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences*, **4**, 383-391.
- MANWELL, L. A., ROBERTS, M. A., & BESNER, D. (2004). Single letter coloring and spatial cuing eliminates a semantic contribution to the Stroop effect. *Psychonomic Bulletin & Review*, **11**, 458-462.
- MELARA, R. D., & ALGOM, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, **110**, 422-471.
- POSNER, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, **32**, 3-25.
- 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.
- STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.
- WÜHR, P., & WASZAK, F. (2003). Object-based attentional selection can modulate the Stroop effect. *Memory & Cognition*, **31**, 983-994.

NOTE

1. One of the reviewers (Daniel Algom) pointed out that the color-word correlation was zero across the whole experiment but nonzero in each particular block of 20 trials. This is correct. However, given the small number of trials per block and the fact that correlations changed from block to block, it should have been very difficult for participants to learn and to use these correlations.

(Manuscript received December 1, 2003;
revision accepted for publication July 25, 2004.)