On the role of stimulus–response and stimulus–stimulus compatibility in the Stroop effect

JAN DE HOUWER
Ghent University, Ghent, Belgium

Stroop effects might be due to differences in stimulus–response compatibility (SRC) and/or to differences in stimulus–stimulus compatibility (SSC). Recent evidence for the role of SSC is inconclusive, because there were no controls for effects of SRC that are based on short-term associations between stimuli and responses (i.e., associations set up as the result of task instructions). In two experiments, SRC effects were controlled for. Regardless of whether the irrelevant and the relevant stimulus features were separated (Experiment 1) or integrated in one stimulus (Experiment 2), the results revealed an effect of SSC and an effect of SRC that was based on short-term associations. The results thus confirm that both processes at the level of encoding and processes at the level of response selection contribute to the Stroop effect.

In the color–word Stroop task (Stroop, 1935), participants see words that refer to colors printed in a corresponding color (e.g., the word *blue* written in blue ink; congruent trials) or in a noncorresponding color (e.g., the word *green* written in blue ink; incongruent trials) and are asked to name the ink color while ignoring the meaning of the word. Performance in this task and other Stroop tasks is consistently better on congruent trials than on incongruent trials (see MacLeod, 1991, for a review).

Several researchers have pointed out that the difference in performance on congruent and incongruent trials can result either from differences in stimulus–response compatibility (SRC) or from differences in stimulus–stimulus compatibility (SSC; e.g., De Houwer, 2003; Kornblum, Hasbroucq, & Osman, 1990). On congruent trials, the meaning of the irrelevant word (e.g., *blue*) is compatible with both the ink color (e.g., blue) and the response (e.g., say "blue"), whereas on incongruent trials, word meaning is incompatible with the ink color and the response. The difference in SRC is assumed to have an effect on the response selection stage. That is, presentation of the irrelevant word automatically activates the responses with a similar meaning. This facilitates response selection on congruent trials but interferes with response selection on incongruent trials (e.g., Cohen, Dunbar, & Mc-Clelland, 1990). However, Stroop effects could also be due to differences in SSC that might have an impact on the stimulus-encoding stage. For instance, the presentation of the irrelevant word could facilitate the encoding

or identification of the relevant ink color on congruent trials and interfere with this encoding on incongruent trials (e.g., Hock & Egeth, 1970).

In an important recent article, Zhang and Kornblum (1998) examined the relative contributions of SRC and SSC to the Stroop effect. Two of their findings supported the hypothesis that differences in SSC do contribute to the Stroop effect. First, they asked participants to perform a Stroop task in which SRC effects could supposedly not operate. In one version of the task, three color words were presented on each trial, only the middle one being relevant. The participants were asked to say one of four digit names out loud on the basis of the identity of the middle color word. For instance, some participants were asked to say "two" when *blue* was the middle word, "four" when *green* was the middle word, "six" when *red* was the middle word, and "eight" when *yellow* was the middle word. Because the stimuli (i.e., color words) and responses (i.e., digit names) were unrelated (i.e., no overlap; see Kornblum et al., 1990), standard SRC effects could not arise. Despite this, performance was better when the three words were identical (e.g., *blue–blue–blue*; congruent trials) than when they were different (e.g., *green–blue–green*; incongruent trials). Hence, Zhang and Kornblum concluded that the meaning of the irrelevant words influenced the identification of the middle word and, thus, that SSC effects do contribute to the Stroop effect.

Second, other participants were asked to respond with color names on the basis of the identity of the middle color word. However, the color words were assigned to incompatible responses (e.g., if the middle word is *blue*, say "green"). The results again showed that performance was better on congruent trials (e.g., *blue–blue–blue*) than on incongruent trials (*green–blue–green*). This effect

could not have been due to standard SRC effects, because on congruent trials, the irrelevant words would have activated an incongruent response (e.g., *blue* activates the incorrect response "blue"), whereas on incongruent trials, the irrelevant words would have activated the correct response (e.g., *green* activates the correct response "green"). Zhang and Kornblum (1998) therefore concluded that Stroop effects are partly due to differences in SSC.

It is important to note that the conclusion reached by Zhang and Kornblum (1998) rests on the assumption that SRC effects are determined exclusively by long-term associations between stimuli and responses. Long-term associations refer to associations that have been overlearned during the lifetime of an individual. For instance, one can assume that there is a long-term association between the word *blue* and the response "blue." However, research has demonstrated that short-term associations can also produce SRC effects. Consider B. A. Eriksen and C. W. Eriksen's (1974) flanker task. Here, participants are instructed to give a certain response (e.g., to press a left key) upon presentation of certain letters (e.g., the letters F and G) and a different response (e.g., to press a right key) upon presentation of other letters (e.g., the letters R and S). During the test phase, three letters are presented side by side in the middle of the screen. The participants are asked to give the response that is associated with the middle letter while ignoring the flanker letters. Responses to the middle letter are slower when it is flanked by letters to which a different response was assigned (e.g., "G R G") than when the same response was assigned to the flanker letters and the middle letter (e.g., "S R S"). In this task, stimuli and responses are associated only as the result of task instructions. Moreover, same- and different-response trials differ only with regard to whether the irrelevant flankers are assigned to the same or to a different response than the relevant middle letter. The fact that performance is different on these types of trials therefore demonstrates that SRC effects can be based on short-term associations that are set up as the result of task instructions (see C. W. Eriksen, 1995, for a review).

Other evidence for the role of short-term stimulus–response associations comes from studies that show that spatial Simon effects can be reversed when an incompatible location-relevant task is presented before or is intermixed with a Simon task (e.g., Proctor & Vu, 2002; Tagliabue, Zorzi, Umiltà, & Bassignani, 2000). In a typical spatial Simon task, participants are asked to press a left or a right key on the basis of the color of stimuli (e.g., squares) that are presented on the left or the right side of a screen. It has repeatedly been demonstrated that left responses are faster to stimuli on the left and that right responses are faster to stimuli on the right, even though the location of the stimuli is irrelevant for the task (see Lu & Proctor, 1995, and Simon, 1990, for reviews). However, when participants also perform an incompatible location-relevant task for which they need to press a left key for uncolored stimuli presented on the right and a right key for uncolored stimuli presented on the left, the results on the trials with colored stimuli (location-irrelevant trials) change dramatically. In such cases, left responses are faster to right colored stimuli than to left colored stimuli, and right responses are faster to left colored stimuli than to right colored stimuli. These results can be explained in the following manner. To perform the incompatible location-relevant task, participants need to set up short-term associations between the representation of a left location and a right response and between the representation of a right location and a left response. When a colored stimulus is presented, it will activate the response that is associated through short-term associations with its (irrelevant) position, rather than the response that has a long-term association with its position (e.g., a colored stimulus on the left side will activate the right response rather than the left response).

Given that short-term stimulus–response associations can produce SRC effects, it is prudent to consider whether the SSC effects that Zhang and Kornblum (1998) observed could actually have been SRC effects that were based on such associations. Take their first task, in which participants responded by saying the name of a digit in response to the identity of the middle word. When the word *blue* is assigned to, for instance, the response "four" and the word *green* is assigned to the response "two," corresponding short-term associations are set up. As a result, those words will activate their responses also when they are presented as to-be-ignored flankers. Therefore, responses will be faster on congruent trials (e.g., *blue–blue–blue*) than on incongruent trials (e.g., *blue–green–blue*). Hence, the difference in performance on congruent and incongruent trials that Zhang and Kornblum observed could have been due to SRC effects, rather than to SSC effects. Likewise, in the second task of Zhang and Kornblum, participants needed to create short-term associations between words and incompatible responses (e.g., between *blue* and the response "green"). Results such as those of Proctor and Vu (2002) suggest that words will, therefore, automatically activate those incompatible responses even when they are presented as irrelevant flankers. This could have produced the difference between congruent and incongruent trials.

In an effort to rule out this alternative explanation of the results of Zhang and Kornblum (1998), the present study used the first task of Zhang and Kornblum. Importantly, control conditions were added that are necessary for an examination of whether SSC plays a role in Stroop effects over and above the effects of SRC. On each trial, three words were presented, and participants were asked to press one of two keys on the basis of the identity of the middle word. Word triads were used in Experiment 1, rather than standard Stroop color–word stimuli, in order to make the present study as similar as possible to those of Zhang and Kornblum. Two words (e.g., *blue* and *purple*) were assigned to the left response, and two words (e.g., *green* and *orange*) were assigned to

the right response. There were three types of trials. First, all three words could be identical (identity trials; e.g., *blue–blue–blue*). Second, the irrelevant flanker words could differ from the middle word but be assigned to the same response (same-response trials; e.g., *purple–blue–purple*). Third, the flanker words could differ from the middle word and be assigned to a different response (different-response trials; e.g., *green–blue–green*).

Same-response and different-response trials differed only with regard to whether the irrelevant words and the relevant word were associated with the same response through task instructions. The flanker and the middle words always differed on both types of trials, so any difference in performance could not be due to SSC effects. Likewise, none of the stimuli had long-term associations with the responses; therefore, standard SRC effects could not occur. Hence, any difference in performance on same-response and different-response trials can be attributed to the difference in SRC as defined by short-term associations. Faster responses on identity trials than on same-response trials, on the other hand, would support the hypothesis that SSC effects do contribute to Stroop effects. This is because the irrelevant words and the relevant word are associated with the same response on both types of trials but refer to the same stimulus only on identity trials. To summarize, better performance on same-response trials than on different-response trials would provide evidence for SRC effects that are based on short-term associations. Better performance on identity trials than on same-response trials would provide evidence for SSC effects.

EXPERIMENT 1

Method

Participants. Fifty-four psychology undergraduates at Ghent University participated for partial fulfillment of course requirements or received 7 Euros for their help in this and unrelated experiments. All had normal or corrected-to-normal vision.

Stimuli and Apparatus. A custom-made Turbo Pascal program was used to present the stimuli and record the responses. This program was implemented on an IBM-compatible 486 computer with a 14-in. screen. The Dutch words *blauw* (blue), *groen* (green), *oranje* (orange), and *paars* (purple) were presented in white letters. On the screen, each letter was 7 mm high and 6 mm wide. The three words on each trial were presented one above the other in the middle of the screen. Responses were made by pressing one of two keys on a keypress device that was connected to the game port of the computer. Presentation duration and reaction times were determined by a highly accurate Turbo Pascal timer (Bovens & Brysbaert, 1990).

Procedure. Instructions were presented on the computer screen. The participants were informed that three words would be presented on each trial and were asked to press a left or a right key on the basis of the identity of the middle word. Two words were assigned to the left response and two to the right response. Which word was assigned to which response was counterbalanced across participants. The instructions further informed the participants that only the middle word was important and that they should try to ignore the top and bottom words of each triad. They were told that there would be one practice block of 32 trials, followed by two test blocks of 32 trials each. The participants were asked to respond as quickly but also as accurately as possible.

On any given trial, the top and the bottom words were identical. There were thus 16 possible triads (4 possible middle words × 4 possible flanking words). Each of these 16 triads was presented twice during each block. The order of the trials was randomized for each block and participant separately, with the following restrictions: (1) a triad could not be repeated until all other triads had been presented once, (2) the flanking words could not be associated with the same response on more than 4 consecutive trials, and (3) the middle word could not be associated with the same response on 5 or more consecutive trials. The participants could take a self-terminated break after each block of 32 trials.

Each trial started with a fixation cross presented at the screen center for 500 msec, followed by a blank screen for 500 msec and a triad that was presented at the screen center until a response was registered or 3,000 msec had elapsed. An incorrect response was indicated by a 200-Hz tone presented for 750 msec. The intertrial interval was 1,500 msec.

Results

The mean reaction time and percentage of errors on identity trials, same-response trials, and different-response trials for each participant and each block of 32 experimental trials were computed separately. Trials on which the reaction time was shorter than 150 msec or longer than 1,500 msec were excluded. Reaction times on trials with an incorrect response were also discarded. For one of the participants, 25% of the reaction times on experimental trials were shorter than 150 msec or longer than 1,500 msec, which was more than three SDs higher than the mean percentage of outliers for the total group. Another participant gave an incorrect response on 15% of all the valid trials. This was more than three SDs higher than the mean percentage of incorrect responses for the total group. The data for these 2 participants were excluded from the analyses. For the remaining participants, 1.27% of all the experimental trials were discarded because the reaction time was shorter than 150 msec or longer than 1,500 msec. An incorrect response was given on 3.51% of the remaining trials.

The reaction time and error data were first analyzed using block (first or second) × trial type (identity, same response, different response) analyses of variance (ANOVAs). Greenhouse–Geisser corrections were applied for effects involving the variable trial type, because this within-subjects variable consists of more than two levels. A priori paired samples t tests were used to examine the effect of trial type in more detail. The relevant means are displayed in Table 1.

The ANOVA for the mean reaction times revealed a main effect of trial type [$F(1.87, 95.51) = 15.25$, $MS_e =$

Table 1
Experiment 1: Mean Reaction Time (RT) and Percentage of Errors (PE) as a Function of Block and Trial Type

| | Block 1 | | | | Block 2 | | | |
| | RT | | PE | | RT | | PE | |
Trial Type	M	SD	M	SD	M	SD	M	SD
Identity	545	108	2.20	4.86	561	97	3.64	6.76
Same response	561	106	2.92	6.89	580	117	2.44	5.05
Different response	582	99	5.12	6.03	599	101	3.29	4.64

38,659.51, $p < .001$]. A priori t tests showed that reaction times were shorter on same-response trials ($M = 571$ msec) than on different-response trials [$M = 591$ msec; $t(51) = 2.81, p = .007$], on identity trials ($M = 553$ msec) than on different-response trials [$t(51) = 5.07, p < .001$], and on identity trials than on same-response trials [$t(51) = 3.03, p = .004$]. The main effect of block was also significant [$F(1,51) = 7.04, MS_e = 23,734.27, p = .01$], showing that reaction times were shorter in the first block ($M = 563$ msec) than in the second experimental block ($M = 580$ msec). This effect could have been due to fatigue or boredom. The interaction between both variables was not significant ($F < 1$).

In the ANOVA for the error data, the main effect of trial type was only marginally significant [$F(1.92, 98.02) = 2.60, MS_e = 72.64, p = .08$]. The a priori t tests showed that fewer errors were made on same-response trials ($M = 2.68\%$) than on different-response trials [$M = 4.21\%; t(51) = 2.35, p = .02$] and, marginally, on identity trials ($M = 2.92\%$) than on different-response trials [$t(51) = 1.77, p = .08$]. There was no difference between identity and same-response trials ($t < 1$). Although the effect of trial type tended to be more pronounced in the first block than in the second (see Table 1), the interaction between block and trial type was not significant [$F(1.91, 97.34) = 2.09, MS_e = 73.86, p = .13$]. The main effect of block was also not significant ($F < 1$).

Discussion

The fact that reaction times were significantly shorter on identity trials than on same-response trials supports the hypothesis that Stroop effects are due in part to differences in SSC. Importantly, this result cannot be explained on the basis of SRC effects that are due to short-term associations, because the flanker words were associated with the correct response on both the same-response trials and the identity trials. There are, however, a number of caveats. First, both in the present experiment and in the experiments of Zhang and Kornblum (1998), a rather unusual variant of the Stroop task was used in which the irrelevant and the relevant stimulus features were part of different stimuli. In the original color–word Stroop task and most subsequent variants of the task, the irrelevant and the relevant stimulus features are part of the same stimulus (e.g., a color word written in a particular ink color). It is possible that SSC effects are specific to the type of task that was used in the present experiment and in those of Zhang and Kornblum. On identity trials, the same word is presented three times, whereas on same-response trials, the flanker words and the middle word differ. As a result, the visual display is less complex and might be encoded more quickly on identity trials than on same-response trials. Such differences in encoding are less likely to arise in a traditional Stroop color–word task, because the irrelevant word and the relevant color are similar only with regard to meaning, and not with regard to form. The present experiment and those of Zhang and Kornblum cannot inform us

about whether SSC effects will also be found when only more central processes can produce these effects. In Experiment 2, I therefore implemented the same design as in Experiment 1 but used color–word stimuli rather than word triads.

A second concern that remained after Experiment 1 related to the reliability of SSC effects. Although the difference between identity and same-response trials was significant in the analysis of the mean reaction times, the percentage of errors did not differ on those two types of trials. A replication of the SSC effect thus seemed necessary before strong conclusions could be drawn.

EXPERIMENT 2

The main difference between Experiments 1 and 2 was that in Experiment 2, the participants were asked to respond on the basis of the ink color of Stroop color–word stimuli. As such, the irrelevant information (word meaning) and the relevant information (ink color) were part of the same stimulus. A second difference was that each participant performed the experiment twice, once without a secondary task and once while counting from 1 to 10 out loud throughout the experiment. In a Stroop color–word task, a possible advantage on identity trials, as compared with same-response trials, could be due to processes at the level of internal naming of the ink colors. That is, when participants are asked to press one of two keys on the basis of ink color, they might internally name the ink color before selecting a response. This step of internal naming might be facilitated by the presence of a word that corresponds to the name of the ink color, thus facilitating overt responses on identity trials. Some might argue that this effect on internal naming should be regarded as an SRC effect (i.e., the word facilitates an internal response; e.g., Dalrymple-Alford & Azkoul, 1972) rather than as an SSC effect (i.e., the word facilitates the identification of the ink color). Processes at the level of internal naming might thus lead to an overestimation of the contribution of SSC to the Stroop effect. Although I had no a priori position with regard to the validity of this argument, I decided to add a manipulation whose aim was to decrease the impact of processes at the level of internal naming. Asking the participants to count out loud from 1 to 10 put a high load on verbal memory (e.g., Baddeley & Hitch, 1974; Levy, 1971), thus making it less likely that the participants would internally name the ink color.

Method

Participants. Forty-eight undergraduates at the University of Southampton participated in the experiment. Twenty-four were psychology students who volunteered. The remainder were students from various faculties, who received £6 for their participation in this and an unrelated second experiment. All had normal or corrected-to-normal vision.

Stimuli, Apparatus, and Procedure. Experiment 2 was identical to Experiment 1, except on the following points. First, colored words were presented instead of word triads. Second, the English

words *blue*, *green*, *pink*, and *yellow* were used. Third, these words were colored in blue, green, pink, or yellow, using the standard settings of Turbo Pascal. Fourth, the participants were told that colored words would appear on the screen one by one and were asked to press a left or a right key on the basis of the ink color of the word while ignoring the word meaning. Which color was assigned to which response was counterbalanced across participants. Fifth, the experimenter explained that the entire task had to be performed twice. Half of the participants first performed the task (three blocks of 32 trials) without counting and afterward did the same task while counting. The order was reversed for the other participants. In the counting condition, the participants were asked to count as fast as possible from 1 to 10 over and over again throughout all three blocks. They were asked to speak at a volume that would allow the experimenter, who remained present during the entire experiment, to verify the numbers uttered by the participant.

Results

The means were calculated in the same way as in Experiment 1, except that they were now calculated separately for the task with counting and the task without counting. For one participant, 13% of the reaction times on experimental trials were longer than 1,500 msec (more than three *SD*s higher than the mean percentage of outliers for the total group), and the mean reaction time on the remaining valid trials was 866 msec (more than two and one half *SD*s higher than the mean reaction time for the total group). Two other participants gave an incorrect response on more than 25% of the valid trials that was more than three *SD*s higher than the mean percentage of errors for the total group. The data for these 3 participants were not included in the analyses, but this did not alter the conclusions. For the remaining 45 participants, 1.65% of all the experimental trials were discarded because the reaction time was shorter than 150 msec or longer than 1,500 msec. An incorrect response was given on 5.53% of the remaining trials. The data were analyzed using secondary task (present or absent) × block (first or second) × trial type (identity, same response, different response) ANOVAs. Greenhouse–Geisser corrections were applied for effects involving the variable trial type. A priori *t* tests were used to examine the effect of trial type in more detail. The relevant means are displayed in Table 2.

The ANOVA of the mean reaction times revealed a main effect of trial type [$F(1.82, 80.13) = 19.31$, $MS_e = 108,723.25$, $p < .001$]. A priori paired samples *t* tests showed that reaction times were shorter on same-response trials ($M = 538$ msec) than on different-response trials [$M = 565$ msec; $t(22) = 4.14$, $p < .001$], on identity trials ($M = 518$ msec) than on different-response trials [$t(22) = 5.63$, $p < .001$], and on identity trials than on same-response trials [$t(22) = 2.63$, $p = .01$]. The effect of trial type was not modulated by secondary task, by block, or by a combination of secondary task and block (all *F*s < 1). The only other effect that approached significance was the interaction between secondary task and block [$F(1,44) = 3.25$, $MS_e = 21,279.53$, $p = .08$], indicating that reaction times tended to be shorter in the first block than in the second, but only when there was no secondary task. All other effects were nonsignificant (*F*s < 1).

Table 2
Experiment 2:
Mean Reaction Time (RT) and Percentage of Errors (PE)
as a Function of Block, Secondary Task, and Trial Type

Trial Type	Block 1				Block 2			
	RT		PE		RT		PE	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
No Secondary Task								
Identity	527	121	6.32	9.00	508	107	4.44	8.06
Same response	552	142	3.14	6.19	539	127	5.08	8.67
Different response	582	155	5.00	5.88	560	135	5.47	5.62
Secondary Counting Task								
Identity	514	118	8.73	10.32	521	115	4.44	8.68
Same response	523	143	7.78	10.35	539	132	3.37	6.26
Different response	558	137	5.77	6.68	558	121	6.63	7.36

A similar analysis of the error data did not reveal a main effect of trial type [$F(1.79, 78.67) = 1.31$, $MS_e = 72.01$]. Likewise, a priori *t* tests showed that none of the pairwise comparisons reached significance (*t*s < 1.45). The effect of trial type did not interact with secondary task (*F* < 1), but the interaction between block and trial type was significant [$F(1.85, 81.41) = 3.98$, $MS_e = 17.09$, $p = .03$]. Table 2 shows that there tended to be more errors on identity trials than on other trials in the first block, but not in the second block. The three-way interaction did not reach conventional levels of significance [$F(1.70, 74.59) = 2.31$, $MS_e = 15.17$, $p = .12$]. Of the remaining, less interesting effects, the main effects of secondary task [$F(1,44) = 2.74$, $MS_e = 19.81$, $p = .11$] and block [$F(1,44) = 3.49$, $MS_e = 19.94$, $p = .07$] approached significance, and the interaction between secondary task and block was significant [$F(1,44) = 6.67$, $MS_e = 26.32$, $p = .01$]. Table 2 shows that fewer errors were made in the second block than in the first, but only when there was a secondary task.

GENERAL DISCUSSION

For many years, there has been a debate about whether Stroop effects are due entirely to processes at the response selection stage (i.e., differences in SRC) or also partly due to processes at the stimulus identification stage (i.e., differences in SSC; see MacLeod, 1991, for a review). Zhang and Kornblum (1998) recently reported evidence for the role of SSC in the Stroop effect. However, their conclusions might have been premature, because they did not take into account SRC effects that are based on short-term associations. In the present experiments, such SRC effects were controlled for. In both experiments, a significant effect of SSC was found. That is, the participants were faster to respond when the irrelevant and the relevant stimulus features referred to the same color (e.g., the word *blue* flanked by the word *blue* or written in blue ink; identity trials) than when they referred to different colors that were assigned to the same response (e.g., the word *blue* flanked by the word *green* or written in green ink when green and blue were as-

signed to the same response; same-response trials). This effect was found irrespective of whether the irrelevant and the relevant stimulus features were part of different stimuli (Experiment 1) or were integrated into one stimulus (Experiment 2).

In Experiment 2, I also examined whether the effect of SSC in a color–word Stroop task was dependent on internal naming of the color. When participants are asked to press one of two keys on the basis of ink color, they might first internally name the ink color. This internal naming response might be facilitated on identity trials, as compared with same-response trials, and this might have lead to a difference in performance between the two types of trials. Results showed, however, that the effect of trial type was not modulated by a secondary task condition. This suggests that the effect of SSC was not situated at a stage of internal naming. Note, though, that the difference between identity and same-response trials did tend to be smaller in the secondary counting task condition (13 msec) than in the condition without a secondary task (28 msec). Moreover, one could argue that the counting task did not disrupt internal naming entirely and that the statistical tests might have lacked sufficient power to detect an impact of the counting task on SSC effects. It might, thus, be worthwhile to examine this issue further.

The present results not only inform us about the role of SSC in Stroop effects, but they also provide further evidence for SRC effects that are based solely on short-term associations and show that such SRC effects can operate in Stroop tasks. On trials in which the irrelevant and the relevant stimulus features referred to different colors, responses were clearly faster when they were associated with the same response than when they were associated with different responses. This effect arose even though there was no similarity or overlap between the stimulus features and the responses (press a left or a right key). Results such as these demonstrate that automatic SRC effects occur not only when stimuli and responses are associated by long-term associations, but also when they are associated only by short-term associations that are set up as the result of instructions (see also De Houwer, in press).

Finally, the results have important implications for models of the Stroop effects. Most often, Stroop effects are attributed to SRC effects that are based on long-term stimulus–response associations (e.g., Cohen et al., 1990; Glaser & Glaser, 1989; Virzi & Egeth, 1985). That is, the irrelevant word is assumed to automatically activate the response alternative with which it has a long-term association (i.e., because the word and the response refer to the same color). This will facilitate response selection on congruent trials but will hinder response selection on incongruent trials.

Most models, however, do not include mechanisms that could produce SSC effects or SRC effects that are based on short-term associations. One possible exception is the connectionist model of Zhang, Zhang, and Kornblum (1999; Zhang & Kornblum, 1998). In this model, a distinction is made between input, intermediate, and output representations. Within each level of representation, nodes that represent different mutually exclusive features (e.g., different colors, meanings, or responses) are linked by inhibitory associations. Therefore, conflicts can arise at each of these three levels. Within the model, SSC effects can result from conflicts at the input or the intermediate level, whereas SRC effects can be due to conflicts at the output level. It is also assumed that intermediate and output nodes are linked not only by long-term associations, but also by short-term associations (which Zhang et al., 1999, called "control lines") that are set up as the result of task instructions. SRC effects that are based on short-term associations can, thus, be modeled by allowing activation to spread along short-term associations even when the activation originates from an irrelevant stimulus feature.

REFERENCES

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.

Bovens, N., & Brysbaert, M. (1990). IBM PC/XT/AT and PS/2 Turbo Pascal timing with extended resolution. *Behavior Research Methods, Instruments, & Computers*, **22**, 332-334.

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.

Dalrymple-Alford, E. C., & Azkoul, J. (1972). The locus of interference in the Stroop and related tasks. *Perception & Psychophysics*, **11**, 385-388.

De Houwer, J. (2003). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219-244). Mahwah, NJ: Erlbaum.

De Houwer, J. (in press). Spatial Simon effects with non-spatial responses. *Psychonomic Bulletin & Review*.

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

Eriksen, C. W. (1995). The flankers task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition*, **2**, 111-118.

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

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

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

Levy, B. A. (1971). The role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **10**, 123-132.

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.

Proctor, R. W., & Vu, K.-P. L. (2002). Eliminating, magnifying, and reversing spatial compatibility effects with mixed location-relevant and irrelevant trials. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention and performance XIX* (pp. 443-473). Oxford: Oxford University Press.

SIMON, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus–response compatibility: An integrated perspective* (pp. 31-86). Amsterdam: Elsevier, North-Holland.

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

TAGLIABUE, M., ZORZI, M., UMILTÀ, C., & BASSIGNANI, F. (2000). The role of long-term-memory and short-term-memory links in the Simon effect. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 648-670.

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

ZHANG, H., & KORNBLUM, S. (1998). The effects of stimulus–response mapping and irrelevant stimulus–response and stimulus–stimulus overlap in four-choice Stroop tasks with single-carrier stimuli. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 3-19.

ZHANG, H., ZHANG, J., & KORNBLUM, S. (1999). A parallel distributed processing model of stimulus–stimulus and stimulus–response compatibility. *Cognitive Psychology*, **38**, 386-432.