Psychonomic Bulletin & Review
2006, 13 (2), 310-315

Filling a gap in the semantic gradient: Color associates and response set effects in the Stroop task

EVAN F. RISKO
University of Waterloo, Waterloo, Ontario, Canada

JAMES R. SCHMIDT
University of Saskatchewan, Saskatoon, Saskatchewan, Canada

and

DEREK BESNER
University of Waterloo, Waterloo, Ontario, Canada

In the Stroop task, incongruent color associates (e.g., LAKE) interfere more with color identification than neutral words do (e.g., SEAT). However, color associates have historically been related to colors in the response set. Response set membership is an important factor in Stroop interference, because color words in the response set interfere more than color words not in the response set. It has not been established whether response set membership plays a role in the ability of a color *associate* to interfere with color identification. This issue was addressed in two experiments (one using vocal responses and one using manual responses) by comparing the magnitude of interference caused by color associates related to colors in the response set with that of interference caused by color associates unrelated to colors in the response set. The results of both experiments show that color associates unrelated to colors in the response set interfered with color identification more than neutral words did. However, the amount of interference was less than that from color associates that were related to colors in the response set. In addition, this pattern was consistent across response modalities. These results are discussed with respect to various theoretical accounts of Stroop interference.

The Stroop task (Stroop, 1935) and its many variants are a fixture in the cognitive psychology literature. This task typically involves the identification of the display color of an incongruent color word (e.g., the word BLUE displayed in red), which leads to slower responding than does the identification of the display color of a neutral stimulus such as a color patch or a neutral word (e.g., the word SEAT in red). The Stroop effect is robust and well documented (see MacLeod, 1991, for a review). In the present investigation, we examined one variant of the Stroop effect wherein words associated to color words (e.g., LAKE) interfere with color identification.

Klein (1964) measured the interference caused by different types of stimuli in a Stroop color-naming task. He reported a *semantic gradient* in which interference increased as a function of the relation between the word and color: (1) Color associates (e.g., LAKE) produced more interference than did neutral words (e.g., SEAT); (2) color words not in the response set (e.g., BROWN when the display colors were red/green/blue/yellow) produced more interference than did color associates; and (3) color words in the response set produced more interference than did color words not in the response set. Thus, as the semantic relationship between the irrelevant word and the display color increased, so did the magnitude of Stroop interference.

The properties of this semantic gradient are of major theoretical interest (MacLeod, 1991), and they have been investigated systematically (Fox, Shor, & Steinman, 1971; Glaser & Glaser, 1989; Klein, 1964; Proctor, 1978; Scheibe, Shaver, & Carrier, 1967; Sharma & McKenna, 1998). However, there remains an empirical gap in the research on this semantic gradient. Specifically, the role of response set membership vis-à-vis the color associate effect has yet to be systematically investigated. Klein's results, along with others' (Proctor, 1978; Sharma & McKenna, 1998), clearly established a role for response set membership with *color* words. However, no such test has been conducted with color word *associates*. Indeed, the research to date has established only that a color associate related to a color *in* the response set will interfere, in comparison with a neutral word (Klein, 1964; Scheibe et al., 1967; Sharma & McKenna, 1998). It has yet to be established whether a color associate unrelated to a color in the response set will interfere in comparison with a neutral word (e.g., will the word LAKE cause more interference

This work was supported by a Natural Sciences and Engineering Research Council of Canada Graduate Scholarship to E.F.R. and Grant A0998 from the Natural Sciences and Engineering Research Council of Canada to D.B. Address correspondence to E. F. Risko, Psychology Department, University of Waterloo, Waterloo, ON, N2L 3G1 Canada (e-mail: efrisko@watarts.uwaterloo.ca).

than that caused by the word SEAT if the display color blue is *not* in the response set?).

In the present experiments, we assessed the relative amounts of interference for color associates related to a color in or out of the response set. Vocal responses were used in Experiment 1, and manual responses were used in Experiment 2. A number of predictions can be derived from extant accounts of Stroop interference with respect to the effect of response set membership and response modality on the color associate effect.

Roelofs (2003)

A response competition model, the WEAVER++ model (Roelofs, 2003), is able to simulate both the previously established color associate effect and the response set membership effect with color words. A color associate can cause response competition by activating a color concept of a potential response via the conceptual network (e.g., the word LAKE is semantically associated with the color concept [blue], so if "blue" is a potential response then response competition would result).

Response set membership and color associates. According to Roelofs's (2003) account, color associates related to a response should produce more interference than should color associates unrelated to a response, given that the former activates the response-related color concept directly. For example, if "blue" is a potential response and "green" is not, the color associate LAKE would activate the concept [blue] and the response "blue," whereas the color associate FROG would activate the concept [green], which in turn would have to activate other color concepts (e.g., [red], [blue]) in order to produce response competition. Thus, this model predicts a response set membership effect for color associates (i.e., color associates related to a response should interfere more than color associates unrelated to a response). Further, color associates unrelated to a response should produce more interference than should neutral words, because neutral words (e.g., SEAT) do not activate color concepts.

Vocal versus manual responding. Roelofs's (2003) model predicts the same outcome for vocal and manual responses, provided one assumes that the latter are lexically mediated: "Stroop interference lies within the language production system. Interference should remain if lexical entries are needed to mediate a button press response" (p. 115). This assumption implies that manipulations of response type, when no effort has been made to rule out lexical mediation, should produce the same qualitative pattern of interference. Critically, this is not always true. For example, vocal but not manual responses yield a "lexical" effect (i.e., neutral words interfere more than consonant letter strings; Sharma & McKenna, 1998). In addition, manual responses, but not vocal responses, yield a reverse Stroop effect (i.e., the display color interferes with word identification; Blais & Besner, in press). Thus, vocal and manual responses do not always produce the same pattern of interference. Given these results, it is important to note that if lexical mediation is not assumed, then Roelofs's (2003) model predicts a response set membership effect for color associates when vocal responses are used, but no color associate effect when manual responses are used.

Sharma and McKenna (1998)

Sharma and McKenna (1998) have proposed two stages at which Stroop interference could arise: a lexical stage and a response selection stage. In their account, which is largely based on the Glaser and Glaser (1989) and Sugg and McDonald (1994) models, the color associate effect is due to interference at the lexical stage. Color associates produce more competition in the lexicon than neutral words do, because the former receive activation from both the direct perception of the word and its semantic association with the display color whereas the latter receive activation only from the direct perception of the word. In addition, Sharma and McKenna claim that the response set membership effect with color words is due to interference at the response selection stage. Color words in the response set can activate competing responses via an identity code.

Response set membership and color associates. According to Sharma and McKenna's (1998) account, the color associate effect should not be modulated by the color associate's status with respect to the response set. Color associates, at least here, are never identical to a response (i.e., the response is never "lake") and therefore should be unable to produce response competition via an identity code. Therefore, the amount of interference from color associates related or unrelated to a potential response should not differ, but both should interfere more than neutral words.

Vocal versus manual responding. In addition, Sharma and McKenna's (1998) account explicitly assumes that manual responses do not have "privileged" access to the lexical stage. Thus, effects claimed to be due to interference at the lexical stage (e.g., the color associate effect) should be present with vocal but not manual responses (but see Brown & Besner, 2001).

Schmidt and Cheesman (2005)

Another multiple-stage account (Schmidt & Cheesman, 2005) claims that Stroop interference is due to both stimulus conflict and response conflict (see also De Houwer, 2003; Zhang & Kornblum, 1998; Zhang, Zhang, & Kornblum, 1999). Stimulus conflict is interference that occurs during stimulus processing (see Klopfer, 1996; Seymour, 1977) and response conflict occurs during response selection (see Cohen, Dunbar, & McClelland, 1990; Roelofs, 2003). Schmidt and Cheesman concluded that the color associate effect was due entirely to stimulus conflict and not response conflict (e.g., the word LAKE activates the color concept [blue] and this interferes with the conceptual encoding of the correct display color on incongruent trials).

Response set membership and color associates. Color associates will activate color concepts in semantics regardless of whether they are related or unrelated to potential responses. Thus, the stimulus conflict account of the color associate effect predicts interference, relative to neutral trials that do not activate color concepts, in both these conditions. This account makes no explicit prediction about the amount of interference for color associates related or unrelated to a response. It only predicts that both such associates should interfere more than neutral words.

Vocal versus manual responding. Stimulus conflict occurs at an early processing stage (semantics) for both vocal and manual responses. Both De Houwer (2003) and Schmidt and Cheesman (2005) have demonstrated stimulus conflict effects with manual responses, the latter with color associates, and Zhang and Kornblum (1998) have demonstrated stimulus conflict effects with vocal responses. Therefore, any observed effects should be independent of response modality.

Summary

Roelofs's (2003) model predicts that color associates related to a potential response will interfere more than color associates unrelated to a potential response, and that the latter will interfere more than neutral words. This response set membership effect with color associates should also be independent of response modality, as long as one assumes that manual responses are lexically mediated.

Sharma and McKenna's (1998) two-stage model of vocal and manual Stroop interference predicts that whether a color associate is related to a response or not should not matter; both should produce interference relative to a neutral word. In addition, there should be no color associate effect with manual responses.

Last, Schmidt and Cheesman's (2005) stimulus conflict account predicts that color associates related and unrelated to potential responses should interfere more than neutral items. This effect should also be independent of response modality.

METHOD

Participants

One hundred forty-two (30 in Experiment 1, and 112 in Experiment 2) University of Waterloo undergraduates participated in exchange for $4 each. All participants reported normal or corrected-to-normal vision and spoke English as a first language.

Apparatus

Stimuli were presented on a 17-in. ADI color monitor. Stimulus presentation and response collection were controlled by E-Prime software (Psychology Software Tools, 2002). Vocal responses were collected by a headset microphone. Manual responses were made on a standard QWERTY keyboard.

Stimuli

The fixation marker, either a "+" or a "−," was presented in white and subtended 2º of visual angle horizontally and vertically. Eight different display colors were used (red/green/blue values in E-Prime): white (255, 255, 255), orange (255, 153, 0), blue (0, 0, 255), green (0, 255, 0), brown (123, 71, 20), yellow (255, 255, 0), gray (155, 155, 155), and red (255, 0, 0). The display colors were separated into two sets of four (white/orange/blue/green and brown/yellow/gray/red). Half the participants received the first set, and the remaining half received the second set.

Sixteen uppercase words in Arial font were used, presented at the center of a black screen. The words were 4 or 7 letters in length, subtending 3º and 5º of visual angle horizontally, respectively, and 1º of visual angle vertically. Eight color-associated words (selected mainly from previous studies), one for each display color, and eight non-color-associated neutral words were selected. Neutral words were matched for length, number of syllables, and approximate frequency with one of the color associates. Color associates and neutral words did not share their first letter with any of the display colors (see the Appendix for the stimulus set).

Each participant saw all of the words. Because each participant received only four of the eight display colors, half the color associates were related to a color in the response set and half the color associates were unrelated to a color in the response set.

Color associates unrelated to a color in the response set, by definition, cannot be displayed in a congruent display color. We therefore eliminated congruent trials from the design. Each color associate related to a color in the response set and its matched neutral word was paired with a color associate unrelated to a color in the response set and its matched neutral word (e.g., FROG–KITE and LIPS–FOOT; see Appendix), and all four of these stimuli appeared in the remaining three display colors. Apart from these restrictions, all stimuli appeared an equal number of times in each of the display colors.

Design

A 3 (trial type: color associate related to a color in the response set, color associate unrelated to a color in the response set, neutral) single-factor within-subjects design was used in both experiments.

Procedure

Each trial began with the presentation of a fixation marker ("+" or "−") in the center of the screen. Participants initiated the trial by pressing the space bar. A blank screen was then presented for 500 msec, after which the colored word appeared at fixation.

In Experiment 1, participants were asked to name the display color of the word aloud. After the participant's response, the experimenter keyed in its accuracy.

In Experiment 2, participants were asked to press a key associated with the display color. The "S," "D," "K," and "L" keys were used as responses, and each key was assigned to each color response an equal number of times across participants. The keys were not labeled.

After either a response or 2,000 msec, a blank screen was presented for 1,000 msec, followed by the next fixation. The fixation was a "+" if the response on the previous trial was correct, and a "−" otherwise. Participants performed one block of 48 practice trials and eight blocks of 48 experimental trials.

RESULTS

Spoiled trials (microphone errors, time-outs, and responses <200 msec; 3.5% in Experiment 1 and 0.6% in Experiment 2) and errors (0.1% in Experiment 1 and 3.6% in Experiment 2) were removed before response time analysis. The remaining data were subjected to a recursive trimming procedure that removed outliers (1.4% of the raw data in Experiment 1 and 2.7% in Experiment 2) on the basis of a criterion cutoff set independently for each participant in each condition with reference to the sample size and the standard deviation in that condition (Van Selst & Jolicœur, 1994). The results from Experiments 1 and 2 are presented in Figure 1.

Experiment 1: Vocal

The main effect of trial type was significant [$F(2,58) = 16.90$, $MS_e = 164.75$, $p < .001$]. Responses to color associates related to a color in the response set (651 msec)

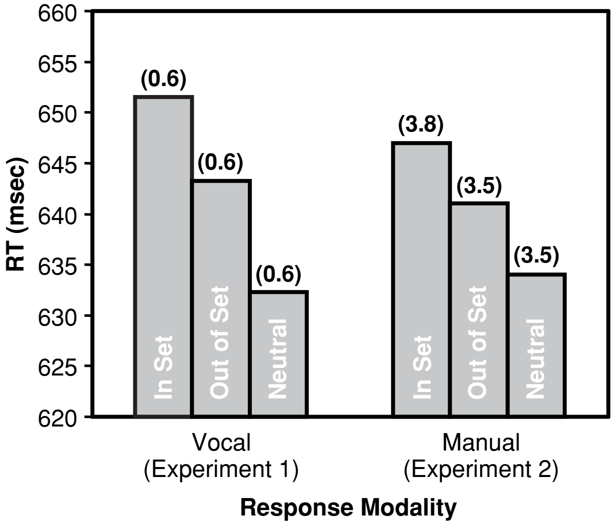

Figure 1. Mean response times (RTs) and percentage errors (in parentheses) for the three trial types: (1) color associates related to a color in the response set (In Set); (2) color associates unrelated to a color in the response set (Out of Set); and (3) neutral (Neutral), as a function of vocal and manual responses.

were slower than responses to color associates unrelated to a color in the response set (643 msec) [$t(29) = 2.33$, $SEM = 3.52$, $p < .05$]. In addition, responses to color associates unrelated to a color in the response set were slower than responses to neutral words (632 msec) [$t(29) = 4.04$, $SEM = 2.72$, $p < .05$]. Errors were committed on less than 1% of the trials, so no error analysis was conducted.

Experiment 2: Manual

The main effect of trial type was significant [$F(2,222) = 8.32$, $MS_e = 537.63$, $p < .001$]. The results were consistent with those of Experiment 1 in that responses to color associates related to a color in the response set (647 msec) were slower than responses to color associates unrelated to a color in the response set (641 msec) [$t(111) = 1.74$, $SEM = 3.38$, $p < .05$, one tailed], and responses to color associates unrelated to a color in the response set were slower than responses to neutral words (634 msec) [$t(111) = 2.39$, $SEM = 2.82$, $p < .05$]. Nothing in the error data contradicted the interpretation of the response times.

DISCUSSION

The present experiments produced three findings. First, color associates related to a color in the response set interfered more than color associates unrelated to a color in the response set. Second, color associates unrelated to a color in the response set produced more interference than did neutral words. Finally, these effects were observed for both vocal and manual responses. None of these effects have been reported previously. We turn now to a discussion of the theoretical implications of these results.

Roelofs (2003)

Roelofs's (2003) response competition model correctly predicted the ordering of the three conditions. In addition, Roelofs's model can also account for the consistency across response modalities, provided it is assumed that manual responses are lexically mediated. However, as noted earlier, there is evidence inconsistent with this assumption (Blais & Besner, in press; Sharma & McKenna, 1998).

Sharma and McKenna (1998)

The present results are inconsistent with the predictions from Sharma and McKenna's (1998) two-stage account of Stroop interference (see also Brown & Besner, 2001). First, their model predicts no response set membership effect with color associates because color associates do not share an identity code with a response. Importantly, the results of the present experiments demonstrate that a subset of the irrelevant stimuli need not be identical to a response in order to produce a response set effect (see also Durgin, 2003).

Sharma and McKenna (1998) also claimed that manual responses do not have access to the lexical stage at which the color associate effect is claimed to originate. The color associate effect obtained here with manual responses is inconsistent with this claim. Either manual responses have access to the lexical stage, or the lexical stage is not where the color associate effect is produced. Our own preference is for the idea that the color associate effect arises in semantics and that manual responses have access to semantic level processing (Brown & Besner, 2001).

Schmidt and Cheesman (2005)

Schmidt and Cheesman (2005) argued that the color associate effect is due entirely to stimulus conflict. This account correctly predicted that color associates related and unrelated to a response produce interference. The fact that these effects were observed across response modalities is also consistent with Schmidt and Cheesman's account. However, their stimulus conflict account makes no explicit prediction regarding the effect of response set membership on the color associate effect. Thus, the present results force a refinement of their account. That is, if the color associate Stroop effect is due to stimulus conflict, then the stage at which this conflict occurs (i.e., semantics) must be sensitive to response set membership. A "priming" process (see, e.g., Glaser & Glaser, 1989; cf. Cohen et al., 1990) may be used to account for response set effects with color associates in this context.

Thus, the present results are consistent with both the Roelofs (2003) and Schmidt and Cheesman (2005) accounts. It is important to note that the two accounts make very different claims regarding (1) the locus of Stroop interference and (2) how semantics produces the color associate effect. In Roelofs's model, interference occurs at a single response selection stage and associates act via a facilitatory mechanism in semantics. In Schmidt and Cheesman's account, interference is of two types, stimulus conflict and response conflict, and associates act via an inhibitory mechanism in semantics (i.e., a stimulus conflict effect). If both of these accounts provide an explanation for the present results, then one is tempted to prefer the single-stage account on grounds of parsimony. However, De Houwer (2003), using a variant of the standard Stroop paradigm, has provided some evidence for the existence of both stimulus and response conflict. In his paradigm, two ink colors are mapped to one response key (e.g., blue and red to one key and yellow and green to another key), producing three conditions: (1) identity trials, in which the irrelevant word is congruent with both the display color and the target response (e.g., the word BLUE in blue), (2) *same* response trials, in which the irrelevant word is incongruent with the display color but congruent with the target response (e.g., the word BLUE in red), and (3) *different* response trials, in which the irrelevant word is incongruent with both the display color and the target response (e.g., the word BLUE in green). The stimulus conflict effect was indexed by comparing identity trials with *same* response trials, and the response conflict effect was indexed by comparing *different* response trials with *same* response trials.

According to a single-stage response competition account like Roelofs's (2003) there should be no difference between identity and *same* response trials (i.e., both signal the same response). However, De Houwer (2003) found that color words produced both a stimulus conflict effect (identity trials were responded to faster than *same* response trials) and a response conflict effect (*same* response trials were responded to faster than *different* response trials). Schmidt and Cheesman (2005) replicated De Houwer's results and also demonstrated that color associates produced only stimulus conflict. It is unclear how a single-stage response competition model would account for these results. The results, however, are consistent with a multiple-locus account of Stroop interference. More generally, although single-stage response competition accounts have historically been favored in the Stroop literature (see MacLeod, 1991), a number of multiple-stage accounts of Stroop interference do exist (De Houwer, 2003; Klopfer, 1996; Seymour, 1977; Zhang & Kornblum, 1998; Zhang et al., 1999). However, to date they have failed to find a receptive audience.

Conclusion

The present investigation provides the first demonstration that associates unrelated to a color in the response set cause interference and also the first demonstration of a response set effect for color associates. Finally, these effects are independent of response modality. The present results thus add to the large body of empirical phenomena associated with the Stroop effect and provide additional constraints on evolving theories of Stroop interference.

REFERENCES

Blais, C., & Besner, D. (in press). Reverse Stroop effect with untranslated responses. *Journal of Experimental Psychology: Human Perception & Performance*.

Brown, M., & Besner, D. (2001). On a variant of Stroop's paradigm: Which cognitions press your buttons? *Memory & Cognition*, **29**, 903-904.

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.

De Houwer, J. (2003). On the role of stimulus–response and stimulus–stimulus compatibility in the Stroop effect. *Memory & Cognition*, **31**, 353-359.

Durgin, F. H. (2003). Translation and competition among internal representations in a reverse Stroop effect. *Perception & Psychophysics*, **65**, 367-378.

Fox, L. A., Shor, R. E., & Steinman, R. J. (1971). Semantic gradients and interference in naming color, spatial direction, and numerosity. *Journal of Experimental Psychology*, **91**, 59-65.

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

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

Klopfer, D. S. (1996). Stroop interference and color-word similarity. *Psychological Science*, **7**, 150-157.

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

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

Psychology Software Tools (2002). E-Prime [Computer software]. Pittsburgh, PA: Author.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, **110**, 88-125.

Scheibe, K., Shaver, P. R., & Carrier, S. C. (1967). Color association values and response interference on variants of the Stroop test. *Acta Psychologica*, **26**, 286-295.

Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus–stimulus and response–response effects in the Stroop task. *Canadian Journal of Experimental Psychology*, **59**, 132-138.

Seymour, P. H. (1977). Conceptual encoding and locus of the Stroop effect. *Quarterly Journal of Experimental Psychology*, **29**, 245-265.

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

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

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

Van Selst, M., & Jolicœur, P. (1994). A solution to the effect of sample size on outlier estimation. *Quarterly Journal of Experimental Psychology*, **47A**, 631-650.

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.

APPENDIX
Stimulus Set for Experiments 1 and 2

There were two display color groups (white/orange/blue/green and brown/yellow/gray/red), which corresponded to the two stimulus groups (SNOW/PUMPKIN/LAKE/FROG and DIRT/MUSTARD/IRON/LIPS). Each participant received one of the display color groups; the corresponding stimulus group made up the *in response set associates*, and the other stimulus group made up the *out of response set associates*. Matched pairs (1–4) never appeared in the display color congruent with the item acting as an in response set associate.

Table A1
Stimulus Set

	Stimulus Group 1		Stimulus Group 2	
Matched Pair	Associate	Neutral Match	Associate	Neutral Match
1	SNOW	MINE	DIRT	TOUR
2	PUMPKIN	INCENSE	MUSTARD	SHERIFF
3	LAKE	SEAT	IRON	COAT
4	FROG	KITE	LIPS	FOOT

(Manuscript received December 22, 2004;
revision accepted for publication August 9, 2005.)