

When the visual format of the color carrier word does and does not modulate the Stroop effect

CHRIS BLAIS and DEREK BESNER
University of Waterloo, Waterloo, Ontario, Canada

We investigated the impact of making the color carrier word visually unfamiliar via case and font mixing in the context of three Stroop experiments. Experiment 1 yielded an *increase* in the size of the Stroop effect when the color carrier words were visually unfamiliar relative to lowercase words that were case and font consistent. Experiments 2A and 2B showed that the modulation of the Stroop effect by visual familiarity observed in Experiment 1 was eliminated when there was no correlation between the color and the color carrier word. These results are considered in the light of four different theoretical accounts of the Stroop effect (strength of association [Cohen, Dunbar, & McClelland, 1990], instance [Logan, 1988], schema [MacLeod, 2000], and obligatory processing followed by deactivation [Coltheart, Woollams, Kinoshita, & Perry, 1999]). None of these accounts appear capable of explaining all the results.

Stroop (1935) was the first to demonstrate that the time required to identify the ink color of stimuli in a list was slowed by the presence of incongruent color words. Hundreds of papers have subsequently explored a number of variants of Stroop's seminal work. In the prototypical modern variant, a single item is displayed and a response to that stimulus is measured. The typical finding is that words displayed in an incongruent color (e.g., the word RED displayed in blue) produce slower reaction times (RTs) and more errors than do words displayed in a congruent color (e.g., the word RED displayed in red; see MacLeod's 1991 review). The present article investigates the consequences, in three Stroop experiments, of making the color carrier word visually unfamiliar (e.g., by case and font mixing). We also consider four extant theories of the processes underlying the Stroop effect, none of which are consistent with the data from all of these three experiments. We end with a discussion of how a subset of the four theories can be modified to account for the present results. We also offer a speculative proposal in which (1) identifying words and colors utilizes common attentional resources; (2) visually unfamiliar words attract more of this resource (by virtue of the word's novelty) than do familiar words; and (3) as a consequence, participants are more

likely to pick up on the contingency between the word and the color (when the congruency proportion is 50:50) and to use this information in the unfamiliar condition.

Automaticity

The received view of the Stroop effect is that it arises because the irrelevant word is processed "automatically." Automaticity is standardly taken to mean that processing (1) is stimulus driven, (2) occurs without intent, (3) is capacity free (in some accounts), and (4) cannot be derailed or interrupted (i.e., once initiated, it runs to completion). Those who discuss the Stroop effect (e.g., Brown, Gore, & Carr, 2002; Posner & Snyder, 1975) typically invoke one or more of these criteria.

Many scholars of adult human cognition believe that neurologically intact young adults who are highly literate have reached this goal of automated word recognition . . . [i]ndeed, it is fair to say that the assumption of automated word recognition in the mature reader is the 'standard' or 'received' view in cognitive science, in part because of the impact exerted by results from the Stroop task. (Brown et al., 2002, p. 220)

Automaticity is also widely believed to be on a continuum such that some processes are more automatic than others, rather than "all or none." MacLeod and Dunbar (1988) were one of the prime movers of the idea that "automaticity is on a continuum" in the context of the Stroop effect. When a participant views a compound stimulus (e.g., a word printed in a color), processing occurs on both the word dimension and the color dimension. The major claim is that the process that is more automatic interferes with the less automatic one, but not vice versa.

Evidence for the idea that heavily practiced stimulus-response (S-R) pairings are responsible for the Stroop effect comes from a training study in which participants labeled shapes with color names. MacLeod and Dunbar

Supported by Grant A0998 from the NSERC of Canada to D.B. Experiment 1 was presented at the 12th Annual Meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science (CSBBCS), Vancouver, Canada. Experiments 2A and 2B were presented at the 14th Annual CSBBCS meeting, St. John's, Canada. We thank Tram Neill, Jennifer Burt, Veronica Dark, and several anonymous reviewers for comments on earlier versions of the manuscript. We also thank Nadine Laudi, Kate Stewart, and Tanya Tolstunov for help in data collection. Correspondence concerning this article should be sent to either C. Blais or D. Besner, Psychology Department, University of Waterloo, Waterloo, ON, N2L 3G1 Canada (e-mail: cblais@watarts.uwaterloo.ca, dbesner@watarts.uwaterloo.ca).

(1988) showed that practice in naming the shapes produced results consistent with a continuum of automaticity framework. After 1 day of practice, familiar color names interfered with naming unfamiliar shapes, but not vice versa. After 20 days of practice, however, heavily practiced shape names interfered with naming familiar colors, but not vice versa. MacLeod and Dunbar argued that these results imply that the more practiced a person is at reading words *relative* to identifying colors, the more word reading interferes with making a color response (but see Roelofs, 2003, for an argument as to why *color-shape* interference is fundamentally different from *color-word* interference).

Context Dependency

In addition to being on a continuum, automaticity is also widely believed to be intimately context dependent. For example, MacLeod (2000) argued that “automaticity is a schema that develops with practice and is highly specific.” This statement was made in particular reference to Besner, Stolz, and Boutilier’s (1997) report that coloring only a single letter in a color word causes a *reduction* in the size of the Stroop effect, relative to a baseline condition in which all of the letters are colored. Besner et al. (1997) used this result in support of the argument that visual word recognition is not automatic in the sense of being obligatory, ballistic, and incapable of being interfered with. In contrast, MacLeod argued that a reduction in the size of the Stroop effect resulting from coloring only a single letter simply shows that the underlying schema(s) for recognizing words is (are) *highly specific*, rather than such data being evidence that visual word recognition is not automatic. In the experiments reported here, we investigate the effect of making words visually unfamiliar in another way, because a number of theories make a variety of predictions.

What Are the Consequences of Making the Word Visually Unfamiliar?

Our departure point for the present set of experiments is the observation that the time to read aloud a letter string or make a lexical decision about it is slowed by case mixing (see, e.g., Besner, 1983; Besner & McCann, 1987). This fact can be viewed, in the context of MacLeod’s context dependency framework, as evidence that the schema at work here is less “automatic” than the schema at work when stimuli are case consistent. A slightly different interpretation is that case mixing reduces processing efficiency. However, for our purposes, these two accounts appear functionally similar.

In the experiments reported here, we examine how variations in the visual familiarity of the color carrier word affect the size of the Stroop effect. A number of existing theoretical accounts appear to make quite different predictions about the effects of such manipulations.

Veridical Versus Abstract Representations

An important issue is the nature of the representation of the words in memory that serve as color carriers in

the experiment. If the representation contains information about visual details of the word, this can be likened to the idea that information is stored in memory as a collection of exemplars (instances) of the word. On the other hand, if the representation does not include visual details, this is analogous to the idea that information about the word is stored in an abstract “prototypical” form that is not tied to its visual properties (see, e.g., Besner, Coltheart, & Davelaar, 1984; Coltheart, 1981; Evett & Humphreys, 1981; Rayner, 1975, for the idea of “abstract letter identities”). In summary, veridical representations will contain information about the word’s visual properties, whereas abstract representations will not.

If the representation of the word in memory is veridical with respect to the word that is presented, then making the color carrier word visually unfamiliar should serve to reduce the strength of association between the presented stimulus and its representation in memory, because fewer examples would exist in memory that were similar to the stimulus in the display. However, if the representation of the stimulus in memory is abstract, then making the stimulus visually unfamiliar need not affect the strength of such associations.

Predictions From Four Variations of Context-Dependent Automaticity

MacLeod (2000): Highly specific schemas. As we noted earlier, MacLeod (2000) asserted that “automaticity is a schema that develops with practice and is *highly specific*” (italics ours). With reference to the Besner et al. (1997) results (i.e., that there is a reduction in the size of the Stroop effect when only a single letter is colored) he further stated that “trying to find one colored letter may change the nature of word processing.” By this logic, it would seem that virtually any departure from a learned schema should result in a decrease in the size of the Stroop effect because the participant will be less practiced at using this modified schema. Thus, MacLeod’s account appears to predict that making the color carrier word visually unfamiliar by case mixing and presenting each letter in a different font should cause a *decrease* in the size of the Stroop effect in comparison with that for a lowercase and font-consistent color carrier word, because the unfamiliar color carrier would be less likely to trigger a schema for processing the word that is automatic. The implication here is that veridical information, either for the word or the schema used to process the word (or both) exists; otherwise, this account is unable to accommodate the reduction in the size of the Stroop effect when a single letter is colored.

Cohen, Dunbar, and McClelland (1990): Strength of association. Cohen, Dunbar, and McClelland (1990) implemented MacLeod and Dunbar’s (1988) strength of processing idea in a connectionist model. The model contains a layer of input units for colors and a layer of input units for words, a layer of response units, and a layer of hidden units that are modulated by a set of task units. The task units allow the model to respond to either dimension (word or color). In order to simulate word reading being

more practiced than color naming, the word pathway received more training cycles than did the color pathway. Consequently, the word pathway had stronger connections between the input and response units than did the color pathway. The trained model was able to successfully reproduce the Stroop effect: Incongruent words interfered with color naming.

To simulate the MacLeod and Dunbar (1988) experiment described earlier, a shape pathway was added to the Cohen et al. model. Once this pathway received the same amount of training as did the participants in MacLeod and Dunbar's Experiment 3, shapes interfered with color naming. Thus, Cohen et al. (1990) demonstrated that the strength of association between input units and response units in a connectionist model (amount of practice on particular S-R pairings) is consistent with the continuum of automaticity framework. More generally, differences in the magnitude of the connection weights between the word and the color are responsible for the size of the Stroop effect.

In Cohen et al.'s (1990) implementation, the representation of the word is abstract. Thus, making the word visually unfamiliar would have no effect on the strength of the connection weights between the stimulus and its representation in memory. No change in the size of the Stroop effect would therefore be expected when the color carrier word is visually unfamiliar. However, there is nothing in Cohen et al.'s model that precludes it from being implemented so as to have veridical representations for the word (e.g., by storing information about the word's size, case, font, etc.) and color (e.g., hue, saturation, brightness, etc.). Under these circumstances, forms of the word most commonly encountered would have the greatest strength of association (e.g., yellow > yELLOW). Thus, a *smaller* Stroop effect is expected when the color carrier word is visually unfamiliar (yELLOW) than when it is visually familiar (yellow) because the strength of association is weaker for the former than for the latter. To see that this is the case, one can simply work backward from the fully trained state of the model to when there is less training on the word. This indeed yields a smaller Stroop effect (see Cohen et al., 1990, pp. 339–342).

Logan (1988): Instance theory of automaticity. Logan stated that

each encounter with a stimulus is encoded, stored, and retrieved separately. Each encounter with a stimulus is assumed to be represented as a processing episode, which consists of the goal the participant was trying to attain, the stimulus encountered in pursuit of the goal, the interpretation given to the stimulus with respect to that goal, and the response made to the stimulus. *When the stimulus is encountered again in the context of the same goal, some proportion of the processing episodes it participated in are retrieved.* (p. 495, italics ours)

Simply put, everything that is attended to is stored as part of the instance. Thus, if the visual format of the word is attended to, this is akin to a veridical representation; failing to attend to the form of the word is akin to an abstract rep-

resentation. If the format of the word is attended to, then making the stimulus visually unfamiliar would serve to reduce the number of instances in memory that are similar to the presented stimulus. The result would be a *reduction* in the size of the Stroop effect, because the probability of finding a match between the displayed stimulus and the instances stored in memory decreases as the number of matching instances decreases. Conversely, if the form of the word is not attended to, the number of matching instances will be the same for both visually familiar and unfamiliar words, and hence we would expect no difference in the size of the Stroop effect as a function of visual familiarity.

Coltheart, Woollams, Kinoshita, and Perry (1999): Cascaded processing with a deactivation process. Coltheart et al. (1999; see their note 2) explained the Stroop effect in terms of the assumption that the presentation of a word causes obligatory activation in the visual word-recognition system. Activation in their dual route model is cascaded, meaning that as soon as there is even the slightest activation at one level, this activation is passed to the next level. It is assumed that activation for a word begins earlier than for a color. Thus, activation for the phonemes associated with the word accumulates at the phoneme level and prevents the activation for phonemes associated with color identification from growing, because there is inhibition between the phonemes associated with the word and those associated with the color. Critically, activation for the word needs to reach a threshold before the word starts to deactivate in the orthographic lexicon. As a result of the cascaded processing integral to this dual route model, any deactivation occurring within the orthographic lexicon would flow on to the semantic system, the phonological output lexicon, and the phoneme system. Regardless of the level at which this is taking place, the deactivation mechanism allows for the activation of the phonemes associated with the color name to eventually dominate (which ultimately allows the model to correctly perform the task).

Coltheart et al.'s account of the Stroop effect thus predicts that anything that slows lexical activation should cause an increase in the size of the Stroop effect. This is because the obligatory activation of the phonemes associated with the color carrier word must reach some criterion before they can begin to be deactivated, and anything that slows that activation delays the time taken for activation to reach threshold. As noted earlier, rendering words visually unfamiliar by case mixing slows reading aloud and lexical decision time. According to the Coltheart et al. account then, making the words visually unfamiliar via case and font mixing should cause an *increase* in the size of the Stroop effect.

Summary

Four theoretical accounts have been discussed. The class of theories in which strength of association—and by extension, continuum of automaticity—plays a major role in producing the Stroop effect (e.g., Cohen et al., 1990;

Logan, 1988; MacLeod, 2000) predicts an outcome that is dependent on how the word is represented in memory. If the word is represented abstractly (e.g., in a lexical system that takes as its input the output of abstract letter identities) such that it is independent of visual features such as case or font, then these three accounts appear to predict no change in the size of the Stroop effect as a function of visual familiarity. On the other hand, if some representation of the color carrier word in memory contains information about case and font, then all of these accounts predict a *decrease* in the size of the Stroop effect, because there are fewer instances in memory that are similar to the presented stimulus. The Logan account stands alone in that it qualifies this so as to depend upon whether such information is attended to or not.

Coltheart et al.'s (1999) account predicts an *increase* in the size of the Stroop effect whenever word processing is slowed. Thus, visually unfamiliar words should produce a *larger* Stroop effect than their visually familiar counterparts.

EXPERIMENT 1

As previously noted, one manipulation known to reliably affect naming and lexical decision is case mixing. Naming and lexical decisions are slower for case-mixed words than for words displayed in lowercase. Such stimuli certainly look visually unfamiliar; case mixing was therefore chosen as a manipulation of visual familiarity. In addition, when the words were case mixed, each letter of the word was also presented in a different font, in order to make the word more visually unfamiliar. Thus, in Experiment 1, half of the words were displayed in lowercase letters and half were displayed as described above.¹ At issue here is whether the Stroop effect will be the same size, larger, or smaller for words rendered visually unfamiliar, in comparison with their visually familiar counterparts.

Method

Participants

Forty-seven undergraduates from the University of Waterloo were tested individually and paid C\$4.00 each for their participation.

Stimuli

The stimuli consisted of four color words displayed either in all lowercase letters or in ALTERNATING CASE, where the first letter of the word always appeared in lowercase. On half of the trials, the stimuli were congruent (i.e., the color and the word referred to the same color), and on the remaining trials they were incongruent (i.e., the color and the word referred to different colors; all possible incongruent pairings were used). For the congruent condition, each word appeared in the color to which it referred 12 times for each of the four color words for each of the lowercase (48 trials) and mixed-case (48 trials) conditions. For the incongruent condition, each word appeared in one of the three colors to which it did not refer, 4 times for each of the four color words for each of the lowercase (48 trials) and mixed-case (48 trials) conditions. The experiment was programmed in MEL (Schneider, 1988) running on a Pentium computer with a color monitor. The four colors used (and their corresponding MEL codes) were red (RGB 42, 0, 0), blue (RGB 0, 0, 42), green (RGB 0, 42, 0), and yellow (RGB 63, 63, 21). In the MIXED-CASE condition,

an additional manipulation of font was used. Each letter of the word in the experimental condition appeared in one of six unfamiliar fonts (MEL: md30.fnt, candy28.fnt, chicam30.fnt, berkly25.fnt, neon28.fnt, and alexis28.fnt) so that no font was presented more than once within a trial. In the lowercase control condition, all letters appeared in one font (MEL font: newyrk22.fnt).

Procedure

On each trial, a single word was presented in the middle of the screen and was terminated as soon as a participant made a response. The participant was told to ignore the word and to identify quickly and accurately in which of the four colors (red, blue, yellow, green) it appeared. The participant responded by pressing the appropriate color on the keyboard; the "z," "x," ">," and "?" keys were covered, respectively, by red, blue, yellow, and green stickers. The intertrial interval was 400 msec. There was no fixation point (i.e., the words always appeared centered at the same location). Each participant saw a different random order of 192 test trials. The experimental trials were preceded by 48 practice trials whose composition mirrored the test trials. Given the low error rates in other Stroop experiments performed in this lab, we did not provide feedback concerning the correctness of the response.

Results

The correct RT data for each participant in each condition were subjected to an outlier analysis, where scores falling 3.5 or more standard deviations above or below the mean for that condition were eliminated from further analyses (Van Selst & Jolicoeur, 1994). This resulted in the elimination of 2.9% of the RT data. The remaining RT data and associated percentage errors can be seen in Table 1.

RT Analysis

A 2 (congruency: congruent vs. incongruent) \times 2 (format: lowercase vs. unfamiliar) ANOVA was carried out on the RT data. Congruent trials were 103 msec faster than incongruent trials [$F(1,46) = 59.7, MS_e = 8,633, p < .001$]. The main effect of format was not significant ($F < 1$). Critically, the interaction between format and congruency was significant [$F(1,46) = 5.7, MS_e = 5,042, p < .05$]. The Stroop effect was 49 msec *larger* for the unfamiliar format as compared to the familiar format.

Error Analysis

A two-way ANOVA of the same design was carried out on the error data. Participants made 3.1% fewer errors on congruent than on incongruent trials [$F(1,46) = 48.3, MS_e = 9.0, p < .001$]. Neither the main effect of format nor the interaction between format and congruency were significant ($F < 1$).

Table 1
Mean Reaction Time (RT, in Milliseconds) and Percentage Error According to Conditions in Experiment 1

Condition	RT				% Error			
	Familiar		Unfamiliar		Familiar		Unfamiliar	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Incongruent	729	17	764	23	5.4	0.6	5.6	0.5
Congruent	650	17	635	14	2.3	0.4	2.6	0.4
Difference	79	17	129	17	3.1	0.6	3.0	0.6

Discussion

The new finding reported here is that rendering the format of the color carrier word visually unfamiliar increases the size of the Stroop effect by more than 50% in comparison to a visually familiar color carrier word (i.e., one printed in lowercase that is also font consistent). With regard to the four theoretical accounts discussed in the introduction, the fact that the magnitude of the Stroop effect is *increased* by making the color carrier word visually unfamiliar appears inconsistent with the strength of association theories discussed in the introduction (i.e., Cohen et al., 1990; Logan, 1988; MacLeod, 2000). When the color carrier word is made visually unfamiliar, the strength of association theories predict either a *decrease* or *no change* in the size of the Stroop effect, depending, in turn, on whether the representation of the word in memory is veridical or abstract, respectively. These theories assume that the processes underlying word recognition become automatic through repeated S–R pairings. If the representation of these color carrier words in memory contains information about visual features such as case and font, then a word that has been rendered visually unfamiliar should produce a smaller Stroop effect, because the strength of association between the stimulus and the response would be reduced. However, if the representations of these color carrier words in memory do not contain information about visual features such as case or font, then the Stroop effect should be the same size for both visually familiar and unfamiliar words. At this juncture it is unclear how these strength of association accounts should be modified so as to produce an *increase* in the size of the Stroop effect for a visually unfamiliar word. This issue can therefore be seen as one ripe for theoretical development.

However, an *increase* in the size of the Stroop effect for visually unfamiliar words is consistent with accounts of the Stroop effect in which word recognition is assumed to be obligatory and the speed of lexical processing is an important factor (Coltheart et al., 1999). These accounts predict that anything that slows lexical processing will cause an increase in the size of the Stroop effect. Case alternation slows naming and lexical decision. Thus, it follows that slowing processing time even more (by also presenting each letter in a different font) should lead to an increase in the size of the Stroop effect, and this is exactly the pattern observed in Experiment 1.

EXPERIMENTS 2A AND 2B

Stroop experiments are typically designed so that half of the trials are congruent and the other half are incongruent (i.e., a 50:50 congruency ratio) and they tend to use between four and six colors. Here, words carry information about color by virtue of a correlation between the word and color dimensions. For example, in Experiment 1, four different colors were used, so that the word *red* was three times more likely to appear in the color red (and similarly for all other colors; see Table 2A) than in any of the other three colors in order to preserve a 50:50

Table 2
Number of Presentations of Each Stimulus per Participant in Each of the Visually Familiar and Unfamiliar Conditions in Experiments 1, 2A, and 2B

Color Dimension	Word Dimension			
	Red	Green	Blue	Yellow
Experiment 1				
Red	12	4	4	4
Green	4	12	4	4
Blue	4	4	12	4
Yellow	4	4	4	12
Experiment 2A				
Red	4	4	4	4
Green	4	4	4	4
Blue	4	4	4	4
Yellow	4	4	4	4
Experiment 2B				
Red	6	6	–	–
Green	6	6	–	–
Blue	–	–	6	6
Yellow	–	–	6	6

Note—In Experiment 1, the congruency ratio was 50:50, and the word dimension was predictive of the color dimension. In Experiment 2A, the congruency ratio was 25:75, and the word dimension was not predictive of the color dimension. In Experiment 2B, the congruency ratio was 50:50, and the word dimension for an individual participant was not predictive of the color dimension (this represents one of the nine counterbalances).

congruency ratio. In short, word processing performance is better on congruent trials than on incongruent trials.

One way of removing this correlation is to pair each color with each word equally often, as in Table 2B. The consequence is that the ratio of congruent to incongruent trials is now 25:75. This is a potential problem, because it confounds congruency ratio with the correlation between word and color dimensions. One way of working around this problem is to control which incongruent items a given participant in the experiment sees, and to counterbalance these across participants. Table 2C provides an example in which there is no correlation between the word and the color, and the congruency ratio is 50:50.

It may also be that the correlation between word and color is more salient when the color carrier word appears in a novel format. If so, then one question is whether the effect observed in Experiment 1 is dependent on this correlation. Rendering a color carrier word visually unfamiliar may serve to increase the recruitment of processing in the service of looking for useful information (i.e., help the participant to “predict” the color [which need not be a conscious process]). A consequence of this may be that it is hard for the participant to disengage from such processing. The question addressed by Experiment 2 thus concerns whether the increase in the size of the Stroop effect observed in Experiment 1 will persist when the correlation between the color and the word is absent.

As discussed prior to Experiment 1, the strength of association theories (Cohen et al., 1990; Logan, 1988; MacLeod, 2000) are unable to accommodate an increase in the size of the Stroop effect in their current form. None-

theless, if the interaction between format and congruency observed in Experiment 1 is eliminated, then these theories can accommodate this result if they assume that the representation of the color carrier word in memory is functionally abstract because participants did not attend to some of the visual details of the color carrier word. On the other hand, Coltheart et al. (1999) assumed obligatory processing of the color carrier word. This account seems to predict that the effect observed in Experiment 1 should persist in Experiment 2, because there is currently no way to avoid the increased delay before deactivation starts.

The issue in Experiments 2A and 2B is simply whether the format of the color carrier word continues to modulate the size of the Stroop effect when the correlation between color and word is absent.

Method

Participants

Sixty-nine and 63 undergraduates from the University of Waterloo served as participants in Experiments 2A and 2B, respectively. Participants were tested individually and paid C\$4.00 each for their participation.

Stimuli and Procedure

The stimuli and procedure were identical to those used in Experiment 1. The only change was the way in which the correlation between the color and word dimensions was removed. In Experiment 2A, this was accomplished by changing the ratio of congruent to incongruent trials from 50:50 to 25:75. For example, the word *red* was equally likely to be presented in red, blue, yellow, or green.

In Experiment 2B, the correlation was removed by counterbalancing which color–word pairings a particular participant would see. The restrictions were that participants (1) saw all of the congruent color–word pairings, and (2) saw four incongruent pairings (so that the congruency ratio was 50:50) such that each word was presented once and each color was represented once. There are nine possible sets of four incongruent pairings that satisfy these restrictions. Hence, a counterbalance consisted of nine participants.

Results

An outlier analysis of the same form as in Experiment 1 was applied to the RT data; this resulted in the elimination of 2.4% and 3.1% of the RT data in Experiments 2A and 2B, respectively. The remaining RT data and associated percentage errors for each of these experiments are displayed in the top and bottom portions of Table 3.

Experiment 2A

RT analysis. A two-way ANOVA was carried out on the RT data. Congruent trials were 66 msec faster than incongruent trials [$F(1,68) = 49.2$, $MS_e = 6,060$, $p < .001$]. The 10-msec main effect of format was not significant ($F < 1.8$). Critically, the interaction between format and congruency was not significant ($F < 1$) even though this experiment had .98 power to detect the 49-msec difference observed in Experiment 1.²

Error analysis. A two-way ANOVA was carried out on the error data. There were 2.5% fewer errors on congruent trials than on incongruent trials [$F(1,68) = 14.2$, $MS_e = 30.8$, $p < .001$]. Neither the main effect of format nor the interaction between format and congruency was significant ($F_s < 1$).

Experiment 2B

RT analysis. A two-way ANOVA was carried out on the RT data. Congruent trials were 81 msec faster than incongruent trials [$F(1,62) = 71.0$, $MS_e = 5,840$, $p < .001$]. The main effect of format was significant; visually unfamiliar stimuli were processed 13 msec more slowly than visually familiar stimuli [$F(1,62) = 4.6$, $MS_e = 5,840$, $p < .05$]. More critically, the interaction between format and congruency was not significant ($F < 1$) even though this experiment had .97 power to detect the 49 msec difference observed in Experiment 1.²

Error analysis. A two-way ANOVA was carried out on the error data. There were 1.8% fewer errors on congruent trials than on incongruent trials [$F(1,62) = 13.4$, $MS_e = 12.1$, $p < .001$]. Neither the main effect of format nor the interaction between format and congruency was significant ($F_s < 1$).

Discussion

Experiments 2A and 2B were identical to Experiment 1 except that the correlation between the word and color was removed, either by changing the congruency ratio to 25:75 congruent:incongruent or by counterbalancing incongruent trials between participants (and maintaining the 50:50 congruency ratio for each participant). The new result is that an unfamiliar format had no impact on the size of the Stroop effect in comparison to the size of the Stroop effect in a visually familiar format.³

Table 3
Mean Reaction Time (RT, in Milliseconds) and Percentage Errors
According to Conditions in Experiments 2A and 2B

Condition	RT				% Error				
	Familiar		Unfamiliar		Familiar		Unfamiliar		
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	
Experiment 2A	Incongruent	756	16	770	17	5.3	0.6	5.7	0.7
	Congruent	694	17	701	17	2.9	0.5	3.0	0.6
	Difference	62	11	69	11	2.4	0.7	2.7	0.9
Experiment 2B	Incongruent	785	20	798	21	3.8	0.5	4.1	0.5
	Congruent	704	16	717	17	2.4	0.4	2.2	0.3
	Difference	81	10	81	13	1.4	0.5	1.9	0.5

GENERAL DISCUSSION

The present experiments investigated the impact on the magnitude of the Stroop effect of rendering the color carrier word visually unfamiliar. Experiment 1 showed that when there was a correlation between the color and word dimensions, the size of the Stroop effect was 49 msec larger for the visually unfamiliar color carrier words than for the visually familiar ones. In Experiments 2A and 2B, the correlation between the color and word dimensions was removed, either by changing the congruency ratio to 25:75 congruent:incongruent or by displaying only a subset of all possible incongruent color–word pairings. Even though these experiments had high power to detect (1) the increase in the size of the Stroop effect for visually unfamiliar words observed in Experiment 1 and (2) comparable size Stroop effects observed for lowercase words (79 msec in Experiment 1; 62 msec in Experiment 2A; and 81 msec in Experiment 2B), Experiments 2A and 2B yielded no difference in the size of the Stroop effect as a function of the color carrier word's visual familiarity.

The results of Experiment 1 appear problematic for the strength of association and instance accounts discussed earlier (Cohen et al., 1990; Logan, 1988; MacLeod, 2000). The reason for this is that these theories predict either a decrease in the size of the Stroop effect (if the representation of the color carrier word in memory contains information about case and font) or no change in the size of the Stroop effect (if the representation of the color carrier word in memory is abstract and as such contains no information about case and font). These theories could explain the results from Experiments 2A and 2B, if they assumed that the representation of the color carrier word in memory is functionally abstract, in that font and case were not attended to and therefore played no role in the retrieval process from long-term memory.

Coltheart et al.'s (1999) deactivation account is consistent with the results from Experiment 1 because it predicts that anything that slows lexical processing should cause an increase in the size of the Stroop effect. Whether this account can accommodate the results from Experiments 2A and 2B depends on how the visual qualities of the color carrier word are conceptualized as being stored and retrieved from long-term memory under the conditions of these experiments.

In summary, none of the four extant accounts can explain all of the results from these three experiments. We favor the view that identifying words and colors utilizes common attentional resources, particularly when both dimensions are in the response set. In keeping with the orienting reflex described by Pavlov (1927), we also assume that a visually unfamiliar word attracts more of such resources (by virtue of its novelty) than does a familiar word. The saliency of the unfamiliar word makes it more likely that participants will pick up on the contingency between word and color when the congruency ratio is 50:50, thus yielding a larger Stroop effect for visually unfamiliar words. However, when the congruency ratio is 25:75, participants pick up, during the practice trials, on

the fact that the word and color are not correlated. The result is that they now process the word in the same way for both familiar and unfamiliar conditions, and therefore the interaction between format and congruency disappears. Nonetheless, the novel words still attract attention, and a main effect of format is still expected. This small effect (<10 msec) was marginal in Experiment 2A and significant in Experiment 2B.

Whatever theoretical view prevails, the present experiments provide the first demonstration that (1) visual format of the color carrier word can modulate the size of the Stroop effect, and (2) this effect appears bound up with the correlation between word and color dimensions. The present experiments and their results thus serve as a useful reminder that issues surrounding the interaction between selective attention, the representation of the stimulus, and its retrieval from long-term memory cannot be easily sidestepped.

REFERENCES

- BESNER, D. (1983). Basic decoding components in reading: Two dissociable feature extraction processes. *Canadian Journal of Experimental Psychology*, **37**, 429-438.
- BESNER, D., COLTHEART, M., & DAVELAAR, E. (1984). Basic processes in reading: Computation of abstract letter identities. *Canadian Journal of Psychology*, **38**, 126-134.
- BESNER, D., & McCANN, R. S. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 201-219). Hillsdale, NJ: Erlbaum.
- BESNER, D., STOLZ, J. A., & BOUTILIER, C. (1997). The Stroop effect and the myth of automaticity. *Psychonomic Bulletin & Review*, **4**, 221-225.
- 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.
- COHEN, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- 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.
- COLTHEART, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language*, **15**, 245-286.
- COLTHEART, M., WOOLLAMS, A., KINOSHITA, S., & PERRY, C. (1999). A position-sensitive Stroop effect: Further evidence for a left-to-right component in print-to-speech conversion. *Psychonomic Bulletin & Review*, **6**, 456-463.
- DISHON-BERKOVITS, M., & ALGOM, D. (2000). The Stroop effect: It is not the robust phenomenon that you have thought it to be. *Memory & Cognition*, **28**, 1437-1449.
- EVETT, L. J., & HUMPHREYS, G. W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology*, **33A**, 325-350.
- LOGAN, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492-527.
- MACLEOD, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, **109**, 163-203.
- MACLEOD, C. M. (2000, July). *Putting automaticity in context: Reducing the Stroop effect*. Paper presented at the joint meeting of the Experimental Psychological Society and the Canadian Society for Brain, Behaviour & Cognitive Science, Cambridge, U.K.
- MACLEOD, C. M., & DUNBAR, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 126-135.
- PAVLOV, I. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex* (G. V. Anrep, Trans.). London: Oxford University Press.

- POSNER, M. I., & SNYDER, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55-85). Hillsdale, NJ: Erlbaum.
- RAYNER, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*, 65-81.
- ROELOFS, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, *110*, 88-125.
- SCHNEIDER, W. (1988). Micro Experimental Laboratory: An integrated system for IBM-PC compatibles. *Behavior Research Methods, Instruments, & Computers*, *20*, 206-217.
- STROOP, J. R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology*, *18*, 643-662.
- VAN SELST, M., & JOLICŒUR, P. (1994). A solution to the effect of sample size on outlier elimination. *Quarterly Journal of Experimental Psychology*, *47A*, 631-650.

NOTES

1. An experiment similar to Experiment 1 compared the size of the Stroop effect for uppercase versus lowercase color carrier words. Since no difference in the effect was found between these conditions, lowercase was arbitrarily chosen as the baseline.
2. Power calculations were computed using J. Cohen's (1988) method for within-subjects designs.
3. It is important to note that removing the correlation between color and word does not eliminate the Stroop effect as implied by Dishon-Berkovits and Algom (2000).

(Manuscript received July 24, 2003;
revision accepted for publication November 24, 2004.)