

Presenting two color words on a single Stroop trial: Evidence for joint influence, not capture

COLIN M. MACLEOD and DOUGLAS A. BORS
University of Toronto, Scarborough, Ontario, Canada

MacLeod and Hodder (1998) demonstrated that presenting two different incongruent color words in the same color on a single Stroop trial resulted in no more interference than did presenting the same incongruent color word twice, and concluded that the first word captured attention, blocking out the second. They also showed that, within a trial, neither stimulus onset asynchrony between the two items nor the presence/absence of a visible gap between the two items had any effect. We replicated all of their empirical findings. Then, by extending their design and factorially combining three types of items—incongruent words, congruent words, and control nonwords—within a trial, we demonstrated that *both* items within a trial do influence processing, with the contribution of the second greater than that of the first. These results are incompatible with a capture account and suggest instead that the word dimension continues to be monitored during the attempt to identify and produce the name of the color.

Use of the Stroop task is by now so prevalent throughout psychology that it would be difficult to find anyone in the field not familiar with it. The simple fact that trying to name the print color of a mismatched color word is slowed by the presence of that word, despite instructions to ignore the word, has served as the basis for many hundreds of studies since Stroop (1935) introduced his task 67 years ago (see MacLeod, 1991, for a review). Variants of the task are now in widespread use in clinical psychology to investigate anxiety-related disorders (see Williams, Mathews, & MacLeod, 1996, for a review) and in cognitive neuroscience brain imaging studies to localize processes of attention (see MacLeod & MacDonald, 2000, for a review). Far from declining in popularity, Stroop's task seems to be in ascendancy.

In Dyer's (1973) review of the Stroop task literature, he argued strongly for a response competition account of the interference caused by the to-be-ignored color word in naming its print color. That is, the effect was seen as being due to conflict between two responses—one to the color word and one to the color—for output. This account became entrenched when, in setting out their views on automatic-

ity, Posner and Snyder (1975) said that "the usual Stroop effect arises because of response competition between vocal responses to the printed word and the ink color . . . color naming and reading go on in parallel and without interference until close to the output" (p. 57).

It would seem to be reasonable and consistent that such a response competition view would predict that the availability of a greater number of potential responses ought to cause more competition and, thereby, result in more interference. That is, the more incongruent words presented on a given trial, the slower the participant's response should be if each additional word provides an additional competing response. We know of only four studies in which such a situation has been investigated, all with the outcome that the presence of additional competing words did not, in fact, increase interference. The first three of these used a separated version of the classic color-word task, in which the word appeared adjacent to a color bar (e.g., the word BLUE below a red color bar, respond "red") and all the stimulus elements within a trial had coincident onsets. Kahneman and Chajczyk (1983, Experiment 4) showed that adding a second incongruent color word adjacent to the color bar (e.g., the word GREEN above the red bar) did not alter interference: Presenting two incongruent color words on a single trial produced no more interference than did presenting one, with both situations displaying around 65 msec of interference. Kahneman and Chajczyk also introduced the concept of *Stroop dilution*, showing that the presence of an extra noncolor word (in addition to the incongruent color word) reduced interference (and facilitation, in the congruent condition) by about half.

On the basis of this pattern, Kahneman and Chajczyk (1983) argued for a simple capture model wherein "one of the words presented on each trial is processed, as if the

This research was supported by Natural Sciences and Engineering Research Council of Canada Grant A7459. We are grateful to Shelley Hodder for assistance in programming and carrying out the experiments, to Ramona Bobocel for helpful comments on the initial draft, and to Daniel Algom and Avishai Henik for valuable reviews. The first author thanks Vrije Universiteit Amsterdam and Duke University for support during his research leave, when this paper was prepared. Correspondence may be sent to C. M. MacLeod or D. A. Bors, Division of Life Sciences, University of Toronto at Scarborough, Scarborough, ON, M1C 1A4 Canada (e-mail: macleod@utsc.utoronto.ca or bors@utsc.utoronto.ca).

—Accepted by previous editorial team

other were not present” (p. 499) and claimed strongly that “When two words are presented together, each has a 50% chance of being processed as if it had been shown alone; the other word is not processed at all” (p. 504). Under this serial-processing account, dilution was the result of participants processing the noncolor word on half of the trials and the incongruent color word on the other half of the trials. In further exploring the dilution effect, Yee and Hunt (1991) replicated their pattern, observing about 40 msec of interference regardless of the number of incompatible words on a trial.¹ Their conclusion was very similar to that of Kahneman and Chajczyk: “Only the word concept that first enters working memory is available to interfere with the color-naming response” (Yee & Hunt, 1991, p. 724).

Like Kahneman and Chajczyk (1983) and Yee and Hunt (1991), Brown, Roos-Gilbert, and Carr (1995) also used the separated version of the task, with simultaneous presentation of the color bar and one or two irrelevant words. They manipulated a number of physical features of the flanker words in the display, observing that dilution was quite unaffected by the nature of the extra flanker (unless it was a letter string like XXXX, which was especially fast). They concluded that Stroop dilution effects “originate in early visual encoding processes” (p. 1408). Consistent with the earlier studies, they found no evidence that increased interference resulted from multiple simultaneous incongruent color words, although their range of interference from a single item (45 msec) to two items (69 msec) was larger than that in prior studies.

The fourth study (MacLeod & Hodder, 1998) did not target the Stroop dilution phenomenon. Also in contrast to the three prior studies, MacLeod and Hodder used the traditional integrated Stroop task, in which the word itself is printed in color. Like the earlier studies, MacLeod and Hodder examined the simultaneous case (the word RED above the word GREEN, both printed in blue, respond “blue”; Experiment 2). However, they emphasized the sequential situation by presenting two incongruent color words successively within a single trial (e.g., RED in green followed by YELLOW in green, respond “green”; Experiment 1). They accomplished this by using relatively short stimulus onset asynchronies (SOAs) between the two words, from 50 to 150 to 250 msec, so that the response could not precede the item switch. As well, both items within a trial were always in the same color, so that the response-relevant dimension was consistent.

Once again, having two incongruent color words within the trial led to no more interference than did having only one incongruent word (presented twice) in the trial. Both the simultaneous and the sequential situations showed about 120 msec of interference; the greater interference relative to the three studies just discussed was due, no doubt, to the integrated stimuli in the MacLeod and Hodder (1998) study, as opposed to the separated stimuli in the prior studies (see MacLeod, 1998). Although there was a cost of about 50 msec for changing the irrelevant item within the trial, the switching costs were equivalent

in the control condition (WWW in blue to SSSSS in blue, respond “blue”) and in the incongruent condition (RED in blue to GREEN in blue, say “blue”). Moreover, whether the word changed or remained the same had no influence.

Like the earlier investigators, MacLeod and Hodder (1998) concluded that “the first word captures attention and ‘locks out’ others, preventing additional interference” (p. 212). This idea is consistent with the notion of stimulus-driven attentional capture (see, e.g., Kahneman & Chajczyk, 1983; Yantis, 1993), assuming that, early in processing, the *word channel* is captured by the first word encoded, so that presenting further words within a trial can have no further impact. Also consistent with the capture view, the abrupt visual onset of a second item within a trial does have a cost, although that cost is no greater for control items than for incongruent words. Put simply, the switching cost is independent of the Stroop interference caused by the incompatible word. MacLeod and Hodder’s results are inconsistent with the response competition view that processing occurs for more than one word on a given trial, at least to an extent sufficient to contribute to the disruption of color naming. Increasing potential competition did not increase interference either in sequential or in simultaneous displays.

In the present experiments, we put this capture account to a more stringent test by constructing more varieties of two-item Stroop trials. The logic is straightforward. If the first item does indeed capture attention on the word dimension, then the nature of that first item should be the only influence on the pattern of responding observed. Specifically, if that first item is a control item (e.g., WWW in blue), then regardless of whether the second item is an incongruent or a congruent word, there should be no interference or facilitation. If the first item is an incongruent word (e.g., RED in blue), then there should be stable interference regardless of the second item; switching to a control or congruent item should not reduce that interference. And if the first item is a congruent word (e.g., BLUE in blue), then there should be facilitation on that trial regardless of the second item—even if the second item is an incongruent word. It is also worth noting that the inclusion of these new conditions allows these same predictions to be assessed for the possibility that it is the second item on a trial, not the first, that captures attention, a possibility equally compatible with MacLeod and Hodder’s (1998) data.

In the two experiments we will report, we manipulated the irrelevant items within trials, producing the 18 combinations of conditions displayed in Table 1. It is perhaps easiest to conceptualize our design as a 4×4 factorial with two extra conditions owing to the possibility of keeping the two items the same or changing them when two incongruent items are involved. This 4×4 design is how we present the data in Tables 2–4. We first created a control item to correspond to each of the color words (i.e., WWW for RED, XXXX for BLUE, SSSSS for GREEN, and MMMMMM for YELLOW). By using incongruent words (e.g., RED or

Table 1
The 18 Conditions of Experiments 1A and 1B

Condition	Item for Each Position in Trial		Abbreviation
	First	Second	
Congruent word–congruent word	green	green	CW–CW
Congruent word–congruent control	green	sssss	CW–CC
Congruent word–incongruent word	green	red	CW–IW
Congruent word–incongruent control	green	www	CW–IC
Congruent control–congruent word	sssss	green	CC–CW
Congruent control–congruent control	sssss	sssss	CC–CC
Congruent control–incongruent word	sssss	red	CC–IW
Congruent control–incongruent control	sssss	www	CC–IC
Incongruent control–congruent word	www	green	IC–CW
Incongruent control–congruent control	www	sssss	IC–CC
Two incongruent controls (same)	www	www	IC–IC(S)*
Two incongruent controls (different)	www	xxxx	IC–IC(D)*
Incongruent control–incongruent word	www	red	IC–IW†
Incongruent word–congruent word	red	green	IW–CW
Incongruent word–congruent control	red	sssss	IW–CC
Two incongruent words (same)	red	red	IW–IW(S)*
Two incongruent words (different)	red	blue	IW–IW(D)*
Incongruent word–incongruent control	red	www	IW–IC†

Note—The conditions are listed in an order corresponding to the entries in Tables 2–4, moving down each column and from left to right. All of the illustrations are for the presentation color green. The condition labels are as follows: CW, congruent word (i.e., green); CC, congruent control (i.e., sssss); IW, incongruent word (e.g., red); IC, incongruent control (e.g., www). The four conditions followed by asterisks are those that were explored in MacLeod and Hodder's (1998) study. The two conditions followed by a dagger are shown in the same condition but actually contained a mixture of *same* and *different* trials (e.g., for IC–IW, the two items in a trial could be a *same* pair, such as www and red in green, as is shown, or a *different* pair, such as xxxx and red in green).

BLUE in green), incongruent controls (e.g., WWW or XXXX in green), congruent words (e.g., GREEN in green), and congruent controls (e.g., SSSSS in green), we could create the condition combinations required to provide the test just described. Thus, for example, we could have a congruent–incongruent trial (GREEN followed by RED, both in green), an incongruent–control trial (RED followed by XXXX, both in green), a control–control trial (XXXX followed by WWW, both in green).

MacLeod and Hodder (1998) also manipulated the SOA between the two items of a trial, ranging from 50 through 150 to 250 msec before the first item gave way to the second. They found that SOA had no effect and saw this as consistent with the capture account they offered. We retained the SOA manipulation for comparability with that prior study and also for generalizability with respect to when the switch between items occurred. We used only two SOAs (100 and 200 msec) but stayed within the same range. When the item changes within a trial (e.g., RED followed by BLUE, both in green), there should be a switching cost, relative to when the item does not change (RED followed by RED, both in green), just as MacLeod and Hodder observed. We refer to these as *different* and *same* trials, respectively, following their earlier nomenclature. However, as they also showed, this switching cost should not interact with the relative influence of the two items on the trial.

We also incorporated another element of MacLeod and Hodder's (1998) design: There were two versions of the

experiment, the only difference being that in Experiment 1A, the second item in a trial immediately replaced the first, whereas in Experiment 1B, there was a visible interruption, or *gap*, of 70 msec between the two items within a trial. This change was made in Experiment 1B to ensure that *same* trials would not appear to be continuous, introducing a confound between *same* and *different* trials, there necessarily being a discontinuity on *different* trials. However, as we will show, the presence or absence of this gap had no effect on the data pattern, just as was the case in MacLeod and Hodder.

We will present Experiments 1A and 1B together, treating 1B as a replication of 1A. An additional value of including the two experiments was that we could test the account developed for Experiment 1A with the independent data set of Experiment 1B. The critical question was whether the first (or the second) item would uniquely influence responding or whether, in fact, both items would contribute. In this more stringent test, would we still obtain evidence compatible with capture, or would we now see that both items within a trial exert their separate influences?

METHOD

Participants

All of the participants were University of Toronto at Scarborough introductory psychology students who received bonus points toward their final grade for their participation. In the final analyses, there were 28 participants in Experiment 1A and 24 in Experiment 1B.

The data of 2 further original participants in Experiment 1A and 6 in Experiment 1B were discarded owing to high error rates (greater than 33% errors in any one condition).

Materials

Four color words (RED, BLUE, GREEN, and YELLOW) and their corresponding control nonwords (WWW, XXXX, SSSS, MMMM) formed the trial items. The nonword controls were selected so as to have the same lengths as their color word counterparts and so as to use repeated letters that do not begin common color words. These verbal items were presented in standard 80-character lowercase font on the computer monitor.

Apparatus

The experiment was controlled by an IBM-AT compatible microcomputer with a 14-in. color VGA monitor. The controlling program, written in QuickBASIC 4.5, used the routines given by Graves and Bradley (1991) to achieve millisecond timing accuracy. Item displays were in one of four colors (red, blue, green, or yellow) against a black background. Oral responses were collected via a microphone positioned directly below the screen in front of the participant. Response times were recorded as the time between the onset of the first stimulus on the screen and the oral response into the microphone, which tripped a voice key that signaled the computer.

Design

The 18 conditions in the experiment are described in Table 1, which also includes an example and an abbreviated label for each condition. Trials always consisted of two items, either the same or different, and both items within a trial were always presented in the same color. The trial types included the 4 conditions of MacLeod and Hodder (1998), marked with asterisks in Table 1, plus 14 additional conditions. The MacLeod and Hodder conditions consisted of combinations of only two types: same or different incongruent words and same or different control nonwords. The new conditions added all remaining possible combinations by using congruent items and completely crossing the congruent, incongruent, and control item types. Table 1 also shows which cells in Tables 2–4 correspond to each of the 18 conditions.

There were 12 trials for each of the 18 conditions for each of the two SOAs, resulting in a total of 432 trials. Where conditions included 24 possible different stimuli, 12 were selected at random for the SOA of 100 msec, and the remaining 12 were used for the SOA of 200 msec. The entire trial sequence was randomized uniquely for each participant.

Procedure

The participants were told that their task was to name the color of print in which a word or string of letters was printed, ignoring the word or the string of letters. They were informed that there would be only four colors, that on some trials they would see two immediately successive words or rows of letters, but that when there were two items on a trial, both would always be in the same color. They were instructed to name the color aloud into the microphone as quickly as they could, avoiding errors. They were also given instructions on how to respond aloud into the microphone, with an emphasis on speaking in a loud voice and avoiding false starts.

Because of the large number of trials required to complete the design and the relative infrequency of errors in such experiments, the participants were instructed to score their own accuracy on each trial. They were to press the “/” key if that trial was correct or the “z” key if that trial was incorrect. MacLeod and Dunbar (1988) had tested this procedure and had found that it led to scoring very similar to that by an experimenter.

The participants had an initial 36 practice trials to become familiar with the task. If necessary, they were given feedback after this

practice session (e.g., they were asked to speak louder or to score their accuracy differently). The 432 experimental trials followed immediately after the practice trials, with short breaks after each quarter of the experimental trials (i.e., after Trials 108, 216, and 324).

The pacing of each trial was as follows. A 250-msec initial blank screen was immediately replaced by a string of eight asterisks at the center of the screen for 250 msec. Following this warning, there was another 250-msec blank screen. Then the first stimulus item appeared for 98 msec (SOA 100; 7 refresh cycles of 14 msec each) or 196 msec (SOA 200; 14 refresh cycles) at the center of the screen. In Experiment 1A, the second stimulus item then immediately replaced the first one at the same location; in Experiment 1B, there was a 70-msec (5 refresh cycles) blank screen, or gap, before the onset of the second stimulus item.² The second item remained on the screen until the participant responded. The word “Accuracy?” then appeared at the center of the screen to prompt the participant to enter his or her accuracy for the just-completed trial. The next trial followed immediately after the participant input the response accuracy.

RESULTS

The mean correct latencies for all of the conditions are shown in Table 2; their respective standard deviations are shown in Table 3. Table 4 displays the mean proportions of errors for all of the conditions. In all cases, the data are collapsed over SOA, in keeping with the initial set of analyses reported in the next section. Experiment 1A is in the top panel, and Experiment 1B is in the bottom panel, in all three tables.

Correspondence Between Experiments 1A and 1B

We first note the evident similarity between the data sets of Experiments 1A and 1B. The participants in Experiment 1B, with the 70-msec gap between items in a trial, were on average 30 msec faster than those in Experiment 1A, without the within-trial gap. However, this between-subjects effect of experiment was not significant in a 2×18 analysis of variance (ANOVA; $F < 1$). There was an overall significant effect of condition [$F(17,850) = 73.82, MS_e =$

Table 2
Mean Correct Response Latencies (in Milliseconds)
as a Function of Item Type, Item–Color Congruency,
and Item Order Within a Trial

Second Item	First Item			
	Congruent		Incongruent	
	Word	Control	Word	Control
Experiment 1A				
Congruent word	688	707	747	709
Congruent control	720	708	796	738
Incongruent word	815	843	824/863	847
Incongruent control	731	738	778	712/734
Experiment 1B				
Congruent word	652	691	753	695
Congruent control	672	684	783	712
Incongruent word	762	786	797/845	789
Incongruent control	687	693	771	691/702

Note—Cells with two entries show the *same* on the left and the *different* on the right.

Table 3
Standard Deviations for Mean Correct Response Latencies
(in Milliseconds) as a Function of Item Type, Item-Color
Congruency, and Item Order Within a Trial

Second Item	First Item			
	Congruent		Incongruent	
	Word	Control	Word	Control
Experiment 1A				
Congruent word	116.14	106.48	96.75	105.40
Congruent control	111.93	94.58	114.50	106.28
Incongruent word	138.29	141.88	120.93/121.57	130.68
Incongruent control	108.23	95.94	87.27	97.30/97.25
Experiment 1B				
Congruent word	131.60	107.22	118.85	114.45
Congruent control	119.36	102.06	117.55	112.54
Incongruent word	143.52	149.42	141.06/140.73	147.22
Incongruent control	114.93	102.80	125.10	113.19/115.30

Note—Cells with two entries show the *same* on the left and the *different* on the right.

2,088.50, $p < .001$] and a reliable interaction [$F(17,850) = 2.07$, $p < .01$]. By inspection, this interaction seemed to arise from a smaller between-experiment difference in those conditions in which an incongruent word came first. Because we have no interpretation for this small difference, we will not discuss it further. Comparison of the data in the top and the bottom of Table 2 clearly shows the strong similarity between the two experiments; the correlation between the latency means in the two experiments was $r = .94$, and the correlation between the accuracy means was $r = .70$.

Speed–Accuracy Relation

Rather than reporting corresponding analyses of the accuracy data for each analysis of response time data throughout all of the following sections, which would add a great many relatively uninformative analyses, we chose the following expedient to demonstrate that there was no speed–accuracy tradeoff overall. For each of the two experiments and for each of the two SOAs, we correlated the means in the 18 conditions for proportion of errors (accuracy) with those for correct response time (latency). These four correlations were all uniformly high (mean $r = .82$, median $r = .85$). As is characteristic of Stroop task data in general and as is clear from Table 4, most of the errors were in the incongruent conditions, which were also the slowest conditions. Thus, on the basis of this absence of a tradeoff between accuracy and latency, we will not report analyses of error patterns in the remainder of the article.

Stimulus Onset Asynchrony

Initial analyses—18 (condition) \times 2 (SOA) within-subjects ANOVAs—were carried out to examine the influence of SOA on response time in both experiments. The SOA variable had been included to parallel the design in MacLeod and Hodder (1998), but the expectation based on that study was that SOA would have no effect. For Experiment 1A, neither the main effect of SOA nor the SOA

\times condition interaction was significant ($F < 1$). In Experiment 1B, there was again no reliable effect of SOA ($F < 1$), but here, for the first time, the interaction was reliable [$F(17,782) = 4.58$, $MS_e = 3,142.52$, $p < .001$]. However, because this interaction showed no interpretable pattern and it was the first time any effect involving SOA in either MacLeod and Hodder or the present study was statistically reliable, we will assume it to be spurious.

On the basis of these analyses, we are comfortable collapsing over SOA for all of the remaining analyses. It appears quite safe to conclude that the difference in onset time of the two items within a single trial is not important, at least within the range we have examined (50, 150, and 250 msec in MacLeod and Hodder's, 1998, study; 100 and 200 msec in the present study).

Comparison With MacLeod and Hodder (1998)

We begin the main analyses by examining the response latencies for only those conditions that had also appeared

Table 4
Mean Proportions of Errors as a Function of Item Type,
Item-Color Congruency, and Item Order Within a Trial

Second Item	First Item			
	Congruent		Incongruent	
	Word	Control	Word	Control
Experiment 1A				
Congruent word	.020	.006	.031	.020
Congruent control	.023	.011	.031	.019
Incongruent word	.058	.074	.062/.070	.099
Incongruent control	.019	.014	.028	.024/.019
Experiment 1B				
Congruent word	.010	.003	.020	.009
Congruent control	.008	.010	.020	.010
Incongruent word	.029	.034	.075/.055	.024
Incongruent control	.008	.005	.020	.009/.010

Note—Cells with two entries show the *same* on the left and the *different* on the right.

in MacLeod and Hodder's (1998) study: incongruent word pairs and incongruent control pairs. Thus, these are $2 \times 2 \times 2$ ANOVAs (SOA \times item type \times same/different). For Experiment 1A, just as in MacLeod and Hodder (1998, Experiment 1A), the main effect of SOA was nonsignificant ($F < 1$), and none of the interactions was significant (all $F_s < 1.68$). Both of the key main effects—item type [$F(1,54) = 135.62, MS_e = 6,001.79, p < .001$] and same/different [$F(1,54) = 12.79, MS_e = 4,129.40, p < .001$]—were significant. The replication was almost perfect, with color words (843 msec) producing 120 msec of Stroop interference relative to control items (723 msec) and with two different items (798 msec) producing an item-switching cost of 30 msec, as compared with one item (768 msec). These correspond to an interference effect of 104 msec and a switch cost of 37 msec in Experiment 1A of MacLeod and Hodder.

Experiment 1B also closely replicated MacLeod and Hodder (1998, Experiment 1B), with the main effect of SOA being nonsignificant ($F < 1$) and all but one of the interactions being nonsignificant (all $F_s < 3.29$). Both of the key main effects—item type [$F(1,54) = 176.35, MS_e = 4,213.98, p < .001$] and same/different [$F(1,54) = 19.54, MS_e = 2,094.00, p < .001$]—were significant. This time, however, these two variables interacted significantly [$F(1,54) = 5.52, MS_e = 3,020.68, p < .05$]. Otherwise, the replication pattern was very similar to MacLeod and Hodder, with color words (821 msec) producing 125 msec of Stroop interference relative to control items (696 msec) and a cost of 30 msec for two different items (774 msec) as opposed to one item (744 msec). These correspond to interference of 133 msec and a switch cost of 34 msec in Experiment 1B of MacLeod and Hodder.

Clearly, then, introduction of the new conditions here (congruent items and mixed word-control pairs) did not disrupt the patterns observed in the present study relative to the previous one; indeed, the means for the corresponding conditions were virtually identical. Once again, Stroop interference did not increase with the move from one to two incongruent words, regardless of the SOA between the two words. That the presence of a noticeable gap between the two items in a trial had no effect on the data pattern implies that the switching cost was caused by the presence of two different items, and not by an interruption per se.

The fact that two incongruent words produced no more interference than one incongruent word led MacLeod and Hodder (1998) to conclude, in keeping with Kahneman and Chajczyk (1983), Yee and Hunt (1991), and Brown et al. (1995), that the first word on a trial captured attention and that the second word contributed nothing further. Despite the appearance of the same pattern in the present experiments for incongruent items, this conclusion no longer is tenable on the basis of the evidence from the new conditions included here, to which we now turn.

Beyond the Incongruent Condition(s)

From here on, we will report the data involving all 18 of the conditions in a less standard format, but one that we

hope makes presentation of our new explanation more straightforward. As a preface, we note that the data support three main conclusions: (1) Any switch between items within a trial (the item itself must change; a visible interruption of the same item is not adequate) has a constant cost relative to no switch, consistent with MacLeod and Hodder's (1998) claim; (2) neither item within a trial captures word processing, contrary to previous claims (Brown et al., 1995; Kahneman & Chajczyk, 1983; MacLeod & Hodder, 1998; Yee & Hunt, 1991); and (3) the two items within a trial exert a combined influence, with the second item contributing more than the first item to that influence. In the following, we will also make more use of the abbreviated labels for the conditions, shown in Table 1.

Switch cost. In comparing the present results with those of MacLeod and Hodder (1998, Experiment 1), we have already shown that there was a switch cost of about 30 msec in the incongruent conditions (IW-IW and IC-IC) when the two items within a trial were different, as opposed to the same. It is less straightforward, but one can also examine *different* versus *same* trials for the congruent stimuli by making one assumption: that a suitable control for CW-CW is CW-CC and that a suitable control for CC-CC is CC-CW, thereby emphasizing the correspondence in the first item on the trial. A $2 \times 2 \times 2$ ANOVA (experiment \times item type \times same/different) revealed no reliable main effect of experiment ($F < 1.10$). Experiment also did not interact with item type in the two-way or the three-way interaction (both $F_s < 1$). The reliable main effect of item type [$F(1,50) = 7.71, MS_e = 1,346.20, p < .01$] showed that words (683 msec) were 15 msec faster to process than controls (698 msec), although the experiment \times item type interaction demonstrates that this was slightly more evident in Experiment 1B than in Experiment 1A [$F(1,50) = 4.79, p < .05$]. That this is evidence of facilitation for the CW-CW case alone is underlined by the significant item type \times same/different interaction [$F(1,50) = 4.32, MS_e = 1,509.18, p < .05$], in which only CW-CW (670 msec) was faster than the other three combinations (CW-CC, 696 msec; CC-CW, 696 msec; CC-CC, 699 msec), showing about 27 msec of facilitation. Finally, and most important in this section, the reliable main effect of same/different showed a 15-msec disadvantage for *different* items (698 msec), relative to *same* items [683 msec; $F(1,50) = 9.00, MS_e = 1,232.56, p < .01$], demonstrating a clear cost for switching, although this cost appears lower for all-congruent items (15 msec) than for all-incongruent items (30 msec), possibly owing to proximity to the performance floor in the congruent case.

Combined influence, not capture. If only the first item in a trial exerted any influence on color-naming time, all instances of incongruent first should display the same interference, all instances of congruent first should display the same facilitation, and all instances of control first should be identical. Table 2 shows that this is certainly not the case. Condition \times experiment ANOVAs for each of these cases confirm this.

For congruent words first (column 1 of the tables), there were four conditions: CW–CW, CW–CC, CW–IW, and CW–IC. The effect of experiment was nonsignificant ($F < 1.93$), as was its interaction with condition ($F < 1$). However, the effect of condition was highly reliable [$F(3,150) = 69.74$, $MS_e = 1,939.81$, $p < .001$]. For control congruent items first (column 2 of the tables), including CC–CW, CC–CC, CC–IW, and CC–IC, both the main effect of experiment ($F < 1.44$) and the interaction [$F(3,150) = 2.03$, $MS_e = 2,274.13$, $p > .10$] were nonsignificant. The main effect of condition was again highly significant [$F(4,200) = 72.23$, $p < .001$]. For incongruent words first, there were five conditions (column 3 of the tables): IW–CW, IW–CC, IW–IW(S), IW–IW(D), and IW–IC. The effect of experiment was nonsignificant ($F < 1$), as was its interaction with condition ($F = 1.04$). Again, however, the effect of condition was highly reliable [$F(4,200) = 42.69$, $MS_e = 1,880.08$, $p < .001$]. For control incongruent items first (shown in column 4 of the tables), including IC–CW, IC–CC, IC–IW, IC–IC(S), and IC–IC(D), both the main effect of experiment and the interaction were nonsignificant ($F_s < 1$). The effect of condition was again significant [$F(4,200) = 5.50$, $MS_e = 1,246.56$, $p < .001$].

In all cases, then, where the nature of the first item on a trial was held constant, there were reliable differences across condition as a function of the different second items on each trial. Note that reanalyzing the data by holding the second item on a trial constant and examining the effect of variation in the first item would lead to identical conclusions. Thus, both items within a trial exerted influences on responding. We now consider the nature of this joint influence.

A simple account of joint influence. We developed this account for Experiment 1A and then tested it with Experiment 1B. Because the two data sets were so similar and the interexperiment correspondence was so close, we will present the model for the data collapsed over the two experiments. First, we had to determine a baseline control mean. There were two obvious control conditions—CC–CC and IC–IC(S)—both of which involved repeated control items and both of which, therefore, should have had no switch cost or congruency effect(s). Indeed, these two conditions produced very similar means of 696 msec for CC–CC and 702 msec for IC–IC(S), for a mean of 699 msec. We took this value as our control mean. These data also showed that the distinction between a *congruent* control item (which contained the same number of characters as its corresponding congruent word; e.g., ssss for GREEN in green) and an *incongruent* control item (which contained a different number of letters from its corresponding incongruent word; e.g., xxx for GREEN in red) did not influence participant performance, as we had anticipated.

Next, we examined the cost of switching by using combinations of control items only. There are three such conditions: IC–CC, CC–IC, and IC–IC(D). A comparison showed these three condition means to be within 9 msec

of each other, with a mean of 720 msec, again showing that the participants did not differentiate congruent control items from incongruent control items. Subtracting the control mean (without switching) from this mean (with switching, but without congruency effects) produced a difference of 21 msec, our estimate of overall switching cost. By removing the control mean (699 msec) and this switching cost (21 msec) from 11 of the 13 other conditions, we could estimate the cost or benefit for each of the various congruency combinations. For 2 of the remaining 13 conditions—CW–CW and IW–IW(S)—there was no switch so, for these 2, only the control mean (and not the switching cost estimate) was removed.

There were four conditions in which a congruent word occurred either first or second, paired with a control item (either congruent or incongruent). Our estimates of the facilitation resulting in these cases are as follows: CW–CC = –24 msec, CW–IC = –11 msec, CC–CW = –21 msec, and IC–CW = –18 msec. These in turn produced estimated facilitation of –18 msec for a congruent word presented first and –20 msec for a congruent item presented second. These are quite typical amounts of facilitation for Stroop experiments (see MacLeod, 1998).

There were also four conditions in which an incongruent word occurred either first or second, paired with a control item (either congruent or incongruent). Our estimates of the interference resulting in these cases are as follows: IW–CC = 70 msec, IW–IC = 55 msec, CC–IW = 94 msec, and IC–IW = 98 msec. These in turn produced estimated interference of 62 msec for an incongruent word presented first and 96 msec for an incongruent item presented second. These are in the modest-to-normal range for amounts of interference in Stroop experiments (see MacLeod, 1998).

The remaining five conditions all involved combinations of congruent and/or incongruent words without control items. For a repeated congruent item (CW–CW), there was –29 msec of facilitation. For a repeated incongruent word (IW–IW[S]), there was 111 msec of interference. For two different incongruent words (IW–IW[D]), there was 134 msec of interference. Finally, for trials combining a congruent and an incongruent word, there was 30 msec of interference for IW–CW and 68 msec of interference for CW–IW.

Can we predict the performance in these two-word conditions from the performance in the conditions that involved only one word (together with a nonword control)? We can. Our parameter values are –18 msec for congruent first, –20 msec for congruent second, 62 msec for incongruent first, and 96 msec for incongruent second. Right away, these parameters indicate that both words in a trial influenced overall response time. That the second word had a larger effect than the first word is clearly evident for incongruent words, but less so for congruent words, where proximity to a performance floor may prevent further differentiation. Certainly, it is not the case that either word captured all processing on the “word channel.”

By using these parameters, IW–CW can be assembled from the incongruent-first (62 msec) and the congruent-second (–20 msec) cases, resulting in a “predicted” value of 42 msec. The CW–IW case can be assembled from the congruent-first (–18 msec) and the incongruent-second (96 msec) cases, resulting in a predicted value of 78 msec. The IW–IW case can be assembled from the incongruent-first (62 msec) and the incongruent-second (96 msec) cases, resulting in a predicted value of 158 msec. Finally, the CW–CW case can be assembled from the congruent-first (–18 msec) and the congruent-second (–20 msec) cases, resulting in a predicted value of –38 msec.

The predicted pattern is quite similar to the observed pattern, but all of the predicted values somewhat exceed the observed values. We assume that combining two words in a trial is slightly underadditive and estimate the resulting overall reduction of effect at 15%. There may be a role for within-trial priming in this reduction (cf. Seifert & Johnson, 1994). Decrementing predicted values by 15% produces a quite good fit: For CW–CW, the predicted value is 32 msec, and the observed value is 29 msec; for IW–IW(S), the predicted value is 134 msec, and the observed value is 111 msec; for IW–IW(D), the predicted value is 134 msec, and the observed value is 134 msec; for IW–CW, the predicted value is 36 msec, and the observed value is 30 msec; and for CW–IW, the predicted value is 66 msec, and the observed value is 68 msec.

DISCUSSION

Along with Kahneman and Chajczyk (1983), Yee and Hunt (1991), and Brown et al. (1995), MacLeod and Hodder (1998) have demonstrated that when two incongruent words appear within a single Stroop trial, there is no more interference than when only one incongruent word appears. In keeping with the idea of attentional capture (see, e.g., Yantis, 1993), all of these investigators concluded that one of the words captured attention, essentially totally blocking out any influence of the other. In the earlier three studies, the two words were presented simultaneously; MacLeod and Hodder also presented the two words successively and further argued that it was the first word that captured attention in this case as well.

The present study argues strongly against these earlier claims. By factorially combining irrelevant congruent words, incongruent words, and control items within a trial, we have shown that both items within the trial influence response time (and accuracy). If capture were indeed occurring, only the first (or perhaps only the second) word would exert an influence on responding. We hasten to point out that our results were not simply a consequence of including so many trial types: The case of two incongruent words still behaved just as in the earlier studies. On the basis of our results, we do suggest that, whether congruent or incongruent, the impact of the first word is somewhat less than that of the second word—perhaps as much as 50% less. Clearly, we continue to encode information from the display even well after we have begun processing the initially perceived information; this is true for at least

200 msec after display onset, given the identical pattern in our data for SOAs of 200 and 100 msec.

One intriguing implication of our findings is that because participants continue to monitor the display, there should be effects of changes within a trial even in the more standard single-item-per-trial format. La Heij, van der Heijden, and Plooiij (2001) have recently reported just such a finding, one that initially seems counterintuitive but makes sense within our explanation. They showed that interference actually decreased if the color offset during the trial. This was true both in the integrated case (e.g., after 150 msec of RED in green, the word RED turned white) and in the separated case (e.g., after 150 msec of a green color bar presented above the word RED printed in white, the color bar disappeared). One might quite readily imagine that removing the word would reduce interference, but it seems surprising that they obtained their reduction in interference when they removed the color.

For our purposes, what is most interesting about La Heij et al.’s (2001) result is that a change in the display well after the processing of a trial is under way does influence responding, showing that continued monitoring of the display is indeed occurring. In fact, our conditions in which the word was removed and replaced by a nonword—whether congruent (CW–CC, CW–IC) or incongruent (IW–CC and IW–IC)—all showed reduced effects relative to the conditions in which the same word remained on the screen throughout the trial (CW–CW and IW–IW). Demarcating the two dimensions in any way seems to assist in keeping the irrelevant dimension from interfering with the relevant one.

Why do both words from the irrelevant verbal dimension affect color naming? We hypothesize that people continue to sample the stimulus display while they are preparing their color naming response, even though the color information in these experiments never changes within a trial, so in fact the first color sample is always all that is needed. Although the switch on the verbal dimension is not response relevant, it is highly noticeable and, no doubt, demands attention (for a related argument, see Seifert & Johnson, 1994). Allocating that attention to the irrelevant dimension thus permits the new irrelevant information to influence responding, apparently more than the original irrelevant information, at least within the range of SOAs we have explored. Thus, there is a role for capture in ensuring that both of the irrelevant verbal items are provided with the chance to affect processing, but not in ensuring absorption by one or the other. We suspect that the most recent of these items has the greater influence because it is the one that, with short SOAs such as the ones we used, remains on display for the longest time, until the color-naming response is output.

REFERENCES

- BROWN, T. L., ROOS-GILBERT, L., & CARR, T. H. (1995). Automaticity and word perception: Evidence from Stroop and Stroop dilution effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*, 1395–1411.
- DYER, F. N. (1973). The Stroop phenomenon and its use in the study of

- perceptual, cognitive, and response processes. *Memory & Cognition*, **1**, 106-120.
- GRAVES, R. E., & BRADLEY, R. (1991). Millisecond timing on the IBM PC/XT/AT and PS/2: A review of the options and corrections for the Graves and Bradley algorithm. *Behavior Research Methods, Instruments, & Computers*, **23**, 377-379.
- KAHNEMAN, D., & CHAJCZYK, D. (1983). Tests of the automaticity of reading: Dilution of Stroop effects by color-irrelevant stimuli. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 497-509.
- LA HEIJ, W., VAN DER HEIJDEN, A. H. C., & PLOOIJ, P. (2001). A paradoxical exposure-duration effect in the Stroop task: Temporal segregation between stimulus attributes facilitates selection. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 622-632.
- MACLEOD, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, **109**, 163-203.
- MACLEOD, C. M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. *Memory & Cognition*, **26**, 201-211.
- 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.
- MACLEOD, C. M., & HODDER, S. L. (1998). Presenting two incongruent color words on a single trial does not alter Stroop interference. *Memory & Cognition*, **26**, 212-219.
- MACLEOD, C. M., & MACDONALD, P. A. (2000). Inter-dimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences*, **4**, 383-391.
- 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.
- SEIFERT, L. S., & JOHNSON, N. F. (1994). On the naming of color words and color patches. *Memory & Cognition*, **22**, 169-180.
- STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.
- WILLIAMS, J. M. G., MATHEWS, A., & MACLEOD, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, **120**, 3-24.
- YANTIS, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science*, **2**, 156-161.
- YEE, P. L., & HUNT, E. (1991). Individual differences in Stroop dilution: Tests of the attention-capture hypothesis. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 715-725.

NOTES

1. Intriguingly, Yee and Hunt (1991) argued that "dilution" was really the result of averaging two kinds of participants: those who always processed the noncolor word and those who always processed the incongruent color word.

2. In Experiment 1B, then, the SOA was actually 170 or 270 msec, given inclusion of this 70-ms gap. However, for ease of description, we will refer to the SOAs in both experiments as 100 and 200 msec.

(Manuscript received May 23, 2001;
revision accepted for publication May 7, 2002.)