Journal of Experimental Psychology:
Human Perception and Performance
1994, Vol. 20, No. 2, 219–234

0096-1523/94/$3.00

Stroop Process Dissociations: The Relationship Between Facilitation and Interference

D. Stephen Lindsay and Larry L. Jacoby

L. L. Jacoby's (1991) "process dissociation procedure" was used to quantitatively estimate the contributions of color-naming and word-reading processes to responding on the Stroop task. The results show that color naming and word reading can operate independently to determine responses. Degrading stimulus colors eliminated the typical asymmetry between Stroop facilitation and interference, as predicted by the equations (Experiments 1 and 2). Degrading stimulus colors reduced the estimated contribution of color naming to responding but had no effect on the estimated contribution of word reading (Experiment 2). In contrast, increasing the proportion of incongruent items reduced the estimated contribution of word reading but had no effect on the estimated contribution of color naming (Experiments 3 and 4). The results indicate that the facilitating and interfering effects of automatic processes cannot be accurately measured in terms of differences from baseline.

In this article we have three purposes. One is to introduce a procedure for quantitatively estimating the contributions of color-naming and word-reading processes to performance on the Stroop (1935) task. The second is to report research in which we used that procedure to explore the relationship between Stroop interference and facilitation. Our results provide powerful evidence that both interference and facilitation reflect the independent contributions of word-reading and color-naming processes. Our third aim is to argue that these findings have important implications for research on facilitation and interference effects in a wide variety of cognitive domains. We argue that the facilitating and interfering effects of automatic processes cannot be accurately measured in terms of differences from baseline (cf. Jonides & Mack, 1984).

D. Stephen Lindsay, Department of Psychology, University of Victoria, Victoria, British Columbia, Canada; Larry L. Jacoby, Department of Psychology, McMaster University, Hamilton, Ontario, Canada.

This research was supported by Natural Science and Engineering Research Council of Canada operating grants to Larry L. Jacoby and D. Stephen Lindsay. Some of the studies reported here were also presented at the 1992 annual meeting of the Midwestern Psychological Association, the 1993 annual meeting of the American Psychological Society, and the 1993 meeting of the Psychonomic Society.

We thank Colleen M. Kelley and Pamela Mink for testing subjects in Experiment 1 at Macalester College. We gratefully acknowledge Lorraine Allen, Vincenza Gruppuso, Colleen M. Kelley, Colin M. MacLeod, Michael E. J. Masson, Steve Tipper, Andy Yonelinas, and two anonymous reviewers for comments on earlier drafts of this article. Thanks are also due to Mike Hunter for advice on statistical analyses.

Correspondence concerning this article should be addressed to D. Stephen Lindsay, Department of Psychology, University of Victoria, P.O. Box 3050, Victoria, British Columbia, Canada V8W 3P5. Electronic mail may be sent to lindsay@uvvm.uvic.ca.

Measuring Facilitation and Interference in the Stroop Task

In the Stroop (1935) task, subjects are required to name the color in which words are printed. Stroop interference refers to the finding that color naming is slowed when the word is the name of another color (e.g., RED in green letters). Stroop interference is the classic example of a task in which a relatively automatic, unintended cognitive process (word reading) opposes a relatively controlled, intended cognitive process (color naming). As such, this phenomenon has been the subject of numerous research projects over the past half century (see MacLeod, 1991, for a review).

One recurrent debate in the Stroop literature concerns the measurement of Stroop interference and facilitation. Customarily, interference is indexed as the difference between color-naming latency on incongruent items (e.g., RED in blue ink) and on "neutral" control items (e.g., color patches, strings of symbols, pseudowords, or non-color-name words), and facilitation is indexed as the difference between congruent items and control items. There has been considerable debate over the years as to what constitutes the appropriate control item—that is, what kind of item provides a factor-pure measure of color-naming time (see MacLeod, 1991). We believe that there is no type of control item that can reliably provide a factor-pure baseline measure of color naming. Lacking a factor-pure control, one can demonstrate the existence of interference and facilitation but cannot accurately measure them. For example, if the control items themselves cause some degree of interference with color-naming processes, then the interfering effect of word-reading processes on incongruent items will be systematically underestimated. The lack of an accurate quantitative measure of the effects of automatic word-reading processes in Stroop interference and facilitation is a major stumbling block for evaluating theories of the cognitive processes that underlie these effects. This problem is not peculiar to the Stroop effect; rather, it is ubiquitous in cognitive psychology (cf. Jonides & Mack, 1984).

MacLeod (1991) suggested that Stroop interference and facilitation may arise from different mechanisms. One rationale for this hypothesis is that Stroop interference is generally much larger than Stroop facilitation. The reasoning is that if word reading and color naming were parallel processes one would expect word reading to help performance on congruent items to the same extent that it hurts performance on incongruent items. We argue, in contrast, that both facilitation and interference reflect the parallel operation of word-reading and color-naming processes. In the introduction to Experiment 1 we explain that, far from being evidence against parallel process models, the asymmetry that is typically observed between facilitation and interference is a natural consequence of the independent contributions of word-reading and color-naming processes. That experiment also provides a clear example of why the effects of word-reading processes on Stroop performance cannot be indexed in terms of differences from baseline. Our process dissociation procedure, described in the next section, provides a better way to measure the effects of word-reading and color-naming processes.

The Process Dissociation Procedure

Jacoby (1991) recently introduced a method for obtaining separate quantitative estimates of the concurrent contributions of controlled and automatic processes to task performance. The method, termed the *process dissociation procedure*, involves comparing performance when intended and unintended processes have the same effect (a facilitation condition) with performance when intended and unintended processes have opposing effects (an interference condition). For each condition, an equation can be written that represents how the two processes act together to determine performance in that condition. These equations express the hypothesis that the controlled and automatic processes act independently of one another to determine responding. Using the two simultaneous equations, the observed performance in the two conditions, and simple algebra, one can derive estimates of the contributions of the two processes to performance.

Consistent with a number of theorists (e.g., Cohen, Dunbar, & McClelland, 1990; Logan, 1980; Posner & Snyder, 1975a), we hypothesized that word reading and color naming make independent contributions to responding on the Stroop task. Under this hypothesis, the probability that a subject will name the color of a congruent item within a response deadline can be described as follows:

$$p(\text{correct} \mid \text{congruent}) = \text{Word} + \text{Color} - (\text{Word} \times \text{Color}),$$

which is to say that there are two ways a correct response can be produced on a congruent item: either because word-reading processes control the response or because color-naming processes control the response. Note that one implication of the equations is that either process (word reading or color naming) is sufficient to control responding on congruent items. By rearranging terms, this equation can be rewritten as follows:

$$p(\text{correct} \mid \text{congruent}) = \text{Word} + \text{Color}\,(1 - \text{Word}).$$

In contrast, the probability that a subject will name the color of an incongruent item within a deadline can be described as follows:

$$p(\text{correct} \mid \text{incongruent}) = \text{Color}\,(1 - \text{Word}),$$

which is to say that a subject can name the color of an incongruent item within the deadline only if the color-naming process and not the word-reading process controls the response.

Subtracting the second of these equations from the first yields an estimate of the contribution of word-reading processes to responding; using that estimate and simple algebra, an estimate of the contribution of color-naming processes to responding can be derived by using either equation. Thus one can obtain separate quantitative estimates of the contributions of word-reading and color-naming processes to performance.

These equations represent the hypothesis that word reading and color naming make independent contributions to color-naming performance on the Stroop task. Note that the terms in these equations do not refer to the occurrence of word reading or color naming but rather to the influence these processes have on overt performance. That is, it is possible that word reading or color naming could occur without contributing to performance (cf. Allport, 1989; Driver & Tipper, 1989). Another implication of the equations is that either process (word reading or color naming) is sufficient to control responding on congruent items. Finally, note that there is an asymmetry in the equations in that word-reading processes dominate over color-naming processes: On incongruent items, correct responses are said to occur only if color-naming and not word-naming processes control the response. In this sense, the equations describe color naming as less automatic than word reading, consistent with the observation that the major problem experienced by subjects in performing the Stroop task is avoiding the highly practiced and hence largely automatic influence of word-reading processes.

A process dissociation is obtained when an experimental manipulation changes the estimated contribution of one underlying process without affecting the other (e.g., a manipulation reduces the estimated contribution of color-naming processes to performance without affecting the estimated contribution of word-reading processes). On theoretical and empirical grounds we predict that a particular manipulation will affect only one process and not the other. We then use the observed measures of performance and our equations to estimate the contributions of word reading and color naming to responding in various conditions and determine whether the predicted pattern of results is obtained (i.e., an effect of the manipulation on one parameter but not on the other). Such findings indicate that the two kinds of process are indeed independent, as the equations imply. That is, if the equations do not accurately represent the way the two processes act together to determine responding, then it is unlikely that a manipulation of one process would affect only the corresponding estimate and not the other estimate (see the General Discussion section for a fuller discussion of this point). We also hypothesized that our estimate of the contribution of color-naming processes (derived using our equations from responses on congruent and incongruent words) will predict color-naming performance on nonletter control items. Finding such correlations increases our confidence in the accuracy of the estimates yielded by the equations.

Measuring Stroop Effects in Terms of Accuracy Rather Than Latency

MacLeod (1991) noted that the change to measuring Stroop effects in terms of reaction time (RT) on individual items, rather than total time on sets of items, was a major methodological advancement. Our deadline technique and the process dissociation procedure may constitute a new and more fundamental advancement: It allows us to measure Stroop effects in terms of accuracy rather than RT and to calculate separate quantitative estimates of the contributions of word-reading and color-naming processes to overt performance on the Stroop task.

Researchers in other domains have demonstrated that accuracy measures are sometimes qualitatively different from RT measures. For example, in research on letter detection, Santee and Egeth (1982) found qualitatively different effects of manipulations when performance was measured in terms of accuracy than when performance was measured in terms of RT (cf. Pashler, 1989). Such findings raise the question of whether measuring Stroop performance in terms of accuracy, as we have done, taps mechanisms that are qualitatively different from those that underlie Stroop effects measured in terms of response latency.

We think this is unlikely, in part because the mechanisms theorized to give rise to dissociations between accuracy and latency in other domains do not appear to map onto our procedure. Both Pashler (1989) and Santee and Egeth (1982) proposed two-component models: The first component is involved in basic perceptual processing of the stimulus, and the second is involved in response selection. Indeed, although Santee and Egeth (1982) presented their findings as showing that RT and accuracy are differently affected by certain manipulations, it might be closer to the mark to say that their findings show that early perceptual processing and response selection are differently affected by those manipulations. As these researchers pointed out, studies that measure letter identification in terms of accuracy typically use very brief exposure durations so as to avoid ceiling effects (e.g., in their studies, they used exposure durations between 8 ms and 20 ms). Thus, accuracy measures are primarily sensitive to variations in very early perceptual processing. In contrast, studies that measure letter identification in terms of RT typically use much longer exposure durations (e.g., 100 ms). Hence, they argued, RT measures are primarily sensitive to variations in response selection. Our exposure duration (800 ms) was relatively long—long enough for the basic perceptual processes involved in perceiving a word or a color. Indeed, to be counted as correct, a response had to be made before the stimulus was removed from the screen. Thus, our deadline procedure is primarily sensitive to variations in the response selection component, just as latency measures are. Our data support this argument.

Overview

In Experiment 1 we measured Stroop task performance in the standard way (i.e., without a deadline and with a response-latency measure) to show that degrading stimulus colors eliminates the asymmetry that is typically observed between Stroop interference and facilitation, just as our equations predict (see introduction to Experiment 1). In Experiment 2 we replicated Experiment 1, using our deadline procedure and measuring performance in terms of accuracy, which allowed us to use our equations to estimate the contributions of word-reading and color-naming processes to performance in the bright- and dull-colors conditions. We obtained a clear process dissociation in Experiment 2, in that degrading stimulus colors lowered the estimated contribution of color-naming processes to performance but had no effect on the estimated contribution of word-naming processes. This pair of experiments serves as a tutorial on the advantages of our approach because in both experiments there were large differences between the bright- and dull-colors conditions in the amounts of facilitation and interference relative to baseline, even though Experiment 2 revealed that the estimated contributions of word-reading processes to performance was the same in both conditions. Experiments 3 and 4 demonstrate the opposite process dissociation: Manipulating the proportion of incongruent items had a large effect on the estimated contribution of word-reading processes but no effect on the contribution of color-naming processes. The experiments also provide insight into correlations between facilitation and interference. Together, they provide strong evidence for the hypothesis that word reading and color naming can act independently to determine responding on the Stroop task.

At a more general level, the experiments reported here illustrate the advantages of looking for dissociations between two processes rather than seeking a type of item or task that provides a factor-pure baseline measure of one of those processes. This approach can also be applied in other contexts in which controlled and automatic processes are hypothesized to make independent contributions to performance. Indeed, we argue that without such an approach, measures of interference and facilitation are likely to be inaccurate.

Manipulating the Influence of Color Naming

Experiment 1

One of the most robust findings in the Stroop literature is that interference (poorer performance on incongruent items than on control items) is generally greater than facilitation (better performance on congruent items than on control items). (See MacLeod, 1991, for a review.) Indeed, facilitation effects sometimes fall short of statistical significance.

As noted earlier, at first glance this asymmetry between interference and facilitation might be taken as evidence against the hypothesis that word reading and color naming contribute to responding as independent processes. The argument would be that if word reading and color naming were parallel processes, one would expect word reading to help performance on congruent items to the same extent that it hurts performance on incongruent items. As explained below, however, the finding of greater interference than facilitation is not at all incompatible with parallel processing models.

Figure 1 depicts performance predicted by our equations on congruent items and incongruent items as a function of the contributions of color-naming (C) and word-reading (W) processes to responding. These equations, which describe word reading and color naming as parallel processes, predict that interference will be greater than facilitation whenever the contribution of color naming to performance is greater than .50. For example, when $C = .80$, interference is four times greater than facilitation, that is, if $C = .80$ and $W = .60$, the proportion correct on congruent items would be $W + C(1 - W) = .60 + (.80 \times .40) = .92$, which is only .12 facilitation relative to C alone; and the proportion correct on incongruent items would be $C(1 - W) = .80 \times .40 = .32$, which is .48 interference relative to C alone. Because color naming is what subjects are supposed to do in the Stroop task, one would generally expect the contribution of color-naming processes to performance to be relatively large (i.e., that color naming would control responding on more than 50% of the trials) and hence for interference to be greater than facilitation. On the other hand, if conditions are such that the contribution of color-naming processes is below .50, our equations predict greater facilitation than interference. For example, if $C = .30$ and $W = .60$, the proportion correct on congruent items would be .72 (.42 facilitation relative to C) and the proportion correct on incongruent items would be .12 (only .18 interference relative to C). Thus, we predicted that lowering the contribution of color naming to responding by degrading the stimulus colors would reduce the typical asymmetry between interference and facilitation.

To manipulate the contribution of color-naming processes to performance, we used the computer videographics array (VGA) palette to degrade the stimulus colors such that they were harder to identify in one condition than in the other. We predicted that degrading color information would eliminate the typical asymmetry observed between Stroop interference and facilitation.

Method

Subjects. The subjects were 40 undergraduate students at the University of Victoria and 20 undergraduate students at Macalester College who participated for extra credit in their introductory psychology courses. Half of the subjects from each institution were assigned to each condition.

Materials and procedure. The experiment was run on a PC clone with a VGA color monitor, using Schneider's (1988) Micro-Experimental Laboratory (MEL) software system. The stimuli were the words *black, blue, green, red,* and *brown,* and strings of percentage signs (%%%% or %%%%%). On each trial, one of these stimuli was presented in one of the five colors in the

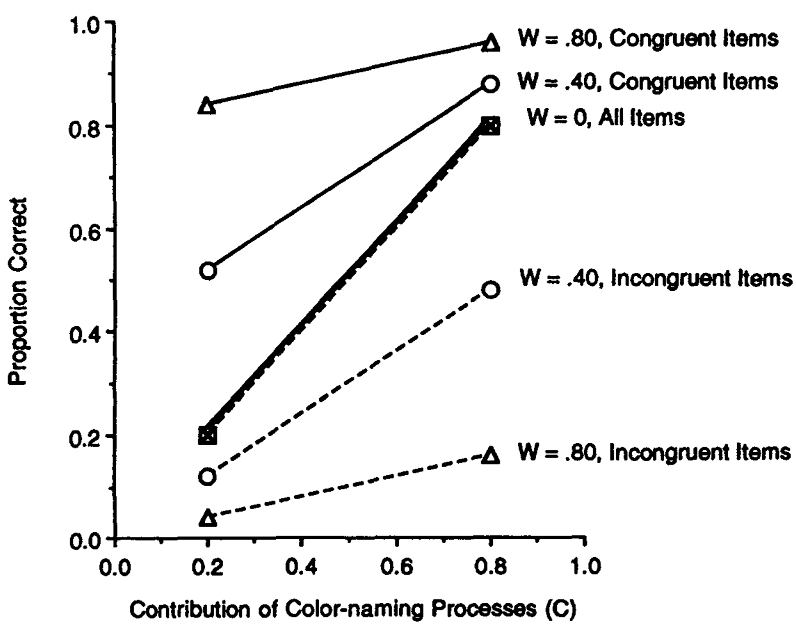

Figure 1. Idealized functions relating varying contributions of color-naming and word-reading processes to color-naming performance on congruent and incongruent items (using the equations in the text). W is the contribution of word-reading processes to performance on a color-naming task. When $W = 0$, there is no facilitation and no interference, and performance on congruent and incongruent items is identical because it is fully determined by color-naming processes (C). When $W > 0$ and $C > .50$, the amount of interference is greater than the amount of facilitation.

center of a light grey computer screen, with each combination of item and color used equally often. There were 100 test trials (20 nonletter control items, 40 incongruent items, and 40 congruent items), with each of the possible combinations of word or control item and color occurring equally often. The test trials were preceded by 40 practice trials, which were drawn from the same pool as the test items and so had approximately the same proportions of trials of each type as the test trials.

The VGA palette control was used to adjust the colors for the two conditions (see the Appendix for the VGA palette settings used in the two conditions). In the bright-colors condition the colors were bright and clear, and each appeared to D. S. Lindsay to be a prototypical exemplar of that color (e.g., a "good" red). In the dull-colors condition, the colors were "muddier" and less prototypical, and the distinctions between the different colors were less clear.[1]

Subjects were tested individually in a quiet room and were seated next to the experimenter. Subjects were instructed to say the color of each item aloud as quickly and accurately as they could. They spoke into a microphone connected to a voice key. The screen went blank when the voice key was tripped. The experimenter then pressed a key to code the subject's response (one of the five colors or *none* if the subject did not say a color name on that trial). When coding responses, the experimenter recorded the subject's initial response (e.g., if the subject said "black-blue" the experimenter keyed in *black*). The next item was presented 2 s after the experimenter keyed in the subject's response.

Results and Discussion

The alpha level for all statistical analyses in this article was .05.

The median RTs for correct responses were analyzed in a 2 × 3 mixed-model analysis of variance (ANOVA), with condition as the between-subjects variable and item type as the repeated measure (see Table 1). There were reliable main effects of condition, with responding reliably faster in the bright-colors condition than in the dull-colors condition, $F(1, 58) = 19.69$, $MS_e = 21{,}042.43$, and item type, $F(1, 58) = 228.30$, $MS_e = 1{,}818.14$. More important, there was a reliable interaction between these variables, $F(2, 116) = 3.92$, $MS_e = 1{,}818.14$. We used Bonferroni-adjusted probabilities to test the differences between facilitation and interference in each condition. In the bright-colors condition, interference (129.4 ms; $t[29] = 10.62$) was significantly greater than facilitation (33.8 ms; $t[29] = 4.87$), $F(1, 29) = 37.80$, $MS_e = 3{,}628.0$. In the dull-colors condition, in contrast, facilitation (71.3 ms; $t[29] = 6.37$) was statistically equivalent to interference (91.5 ms; $t[29] = 8.28$), $F(1, 29) = 1.12$, $MS_e = 5{,}434.3$.

Table 1
Means and Standard Deviations for Response Latency (in Milliseconds) on Correct Responses Within Deadline: Experiment 1

	Item type					
	Congruent		Incongruent		Nonletter	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*
Bright colors	628	18	594	14	757	18
Dull colors	749	18	677	14	840	18

Note that facilitation was greater in the dull-colors condition (71.3 ms) than in the bright-colors condition (33.8 ms), $F(1, 29) = 8.13$, $MS_e = 2{,}602.0$. If the influence of word-reading processes were measured in terms of facilitation, one would conclude that word reading played a larger role in the dull-colors condition than in the bright-colors condition. Conversely, interference was greater in the bright-colors condition (129.4 ms) than in the dull-colors condition (91.5 ms), $F(1, 29) = 5.32$, $MS_e = 4{,}055.0$. If the influence of word-reading processes were measured in terms of interference, one would conclude that word reading played a larger role in the bright-colors condition than in the dull-colors condition, a conclusion opposite to that suggested by facilitation. Obviously, at least one of those conclusions must be wrong. The problem is, of course, that existing techniques cannot tell which conclusion is wrong or that both are wrong.

A critic might point out that it is immediately obvious that the effects of word reading cannot be identified with facilitation or interference in this experiment because performance on the baseline items was radically different in the bright- and dull-colors conditions. The obviousness of the problem in this case should be scant solace to those who would index facilitation and interference relative to baseline in other cases: Variations in baseline, both within and across experiments, are the norm in many domains. Indeed, it is not unusual for researchers to report only measures of facilitation or interference, without reporting the baselines at all. Even when differences in baseline within or across experiments are acknowledged, the significance of the problem is not appreciated; even if it were, existing techniques could not solve the problem (e.g., one cannot simply subtract baseline, or add baseline as a covariate to obtain an accurate index of facilitation or interference). Performance on facilitated items and on inhibited items always reflects the simultaneous contributions of independent processes, so one cannot measure priming or inhibition without having a means of separately estimating both contributing processes. As described later, our process dissociation procedure provides such a means.

Experiment 2

If word-reading processes and color-naming processes operate independently of one another, manipulations that affect one kind of process need not affect the other. Furthermore, such manipulations should affect the corresponding parameter estimate derived with our equations but should not affect the other parameter estimate. One purpose of Experiment 2 was to demonstrate that degrading stimulus colors reduces the estimated contribution of color-naming processes to responding on the Stroop color-naming task but has no effect on the estimated contribution of word-reading processes. We also hoped to replicate, with the deadline procedure and accuracy measure, the finding of Experiment 1 that degrading

[1] The exact psychophysical properties of the stimuli were not of interest here. Our aim was merely to contrast a condition in which colors were clear and distinct and therefore easy to identify with a condition in which they were less clear and distinct and therefore harder to identify.

stimulus colors eliminates the typical asymmetry between Stroop interference and facilitation.

Method

Subjects. The subjects were 60 University of Victoria undergraduate students who participated for extra credit in an introductory psychology course.

Materials and procedure. The materials and procedure were identical to Experiment 1 except that subjects were instructed to name the color of each stimulus before an 800-ms deadline. An error tone sounded when subjects did not respond within the deadline.

Results and Discussion

Accuracy. Mean proportions correct on congruent, incongruent, and nonletter control items were calculated at 50-ms intervals from 600 to 800 ms. (Below 600 ms too many subjects had zero correct on either incongruent or nonletter items.) Two subjects (one in each condition) were excluded from the analyses because even at 600 ms they had zero correct on either incongruent or nonletter items. The mean proportions correct on each type of item at each post hoc deadline are presented in Figure 2. These data were analyzed in a 2 × 2 × 2 mixed-model ANOVA, with condition (bright vs. dull colors) as the between-subjects variable and item type (congruent, incongruent, or nonletter control items) and the two extreme post hoc deadlines (600 vs. 800 ms) as repeated measures.[2]

There were reliable main effects of condition, $F(1, 56) = 13.26$, $MS_e = .085$; item type, $F(2, 112) = 187.17$, $MS_e = .009$; and deadline, $F(1, 56) = 411.80$, $MS_e = .007$. More important, there was a reliable interaction between condition and item type, $F(2, 112) = 6.66$, $MS_e = .009$. In the bright-colors condition, as predicted, there was substantial interference on incongruent items relative to nonletter control items but relatively little facilitation on congruent items. At the long deadline, this is doubtless partly due to ceiling effects, but the pattern holds even at the shortest deadline, when performance on the nonletter controls was below ceiling for every subject. An analysis of the proportion correct at the 600-ms deadline revealed that in the bright-colors condition, interference ($M = .41$) was reliably greater than facilitation ($M = .09$), $F(1, 28) = 43.79$, $MS_e = .032$, although both were significantly greater than zero, $F(1, 28) = 80.68$, $MS_e = .025$, and $F(1, 28) = 11.90$, $MS_e = .020$, respectively. In the dull-colors condition, in contrast, interference (.23) was not greater than facilitation (.24; $F < 1$); both were significantly greater than zero, $F(1, 28) = 81.84$, $MS_e = .020$, and $F(1, 28) = 65.86$, $MS_e = .023$, respectively. Thus in the bright-colors condition, facilitation was much smaller than interference (as is typically found); however, in the dull-colors condition, facilitation was substantial and equivalent to interference.

As in Experiment 1, if the influence of word-reading processes were measured in terms of interference, one would conclude that word reading had a greater influence in the bright-colors condition than in the dull-colors condition; however, if the influence of word reading were measured in terms of facilitation, one would conclude that word reading had a greater influence in the dull-colors condition than in the bright-colors condition. The estimate of the contribution of word-reading processes to responding derived by using our equations (reported in the next section) shows that both of these conclusions would be wrong.

Estimates. Estimated contributions of word-reading and color-naming processes to performance were calculated at each post hoc deadline, using the equations described in the

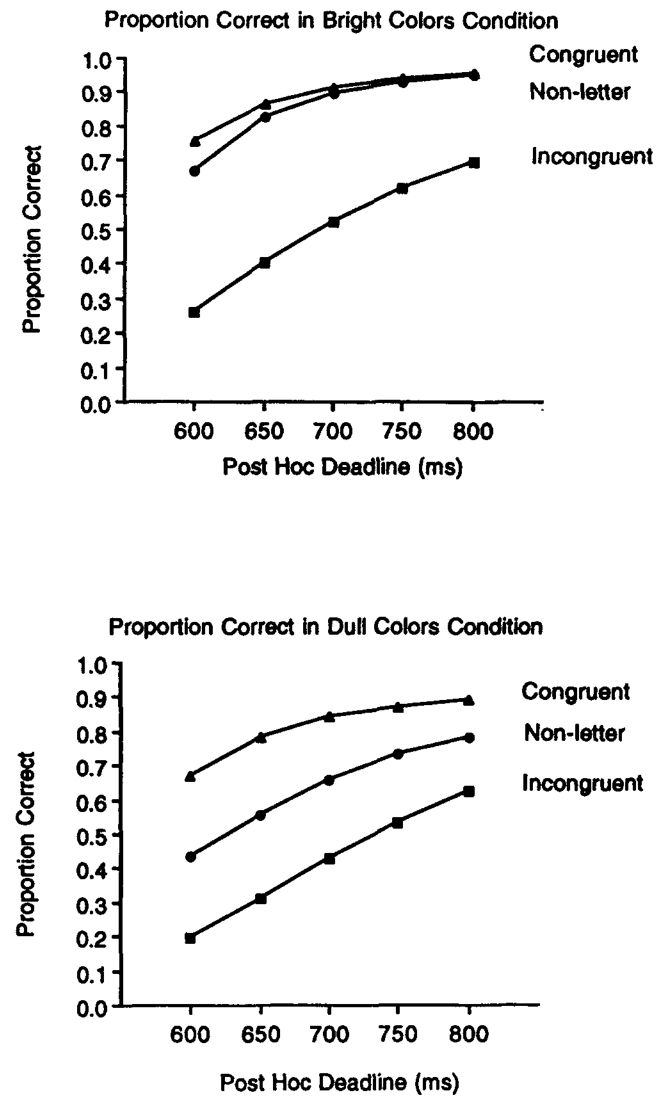

Figure 2. Experiment 2: Mean proportions correct on congruent, incongruent, and nonletter items at each post hoc deadline. Top: bright-colors condition. Bottom: dull-colors condition.

[2] Only the extreme deadlines (one set by the 800-ms deadline in the procedure itself, and the other set by the fact that below 600 ms subjects often had zero correct on incongruent or nonletter items) were used to minimize the extent to which the independence assumption of analysis of variance was violated.

introduction. These data are shown in Figure 3. Separate 2 × 2 mixed-models ANOVAs were performed for each of these measures, with condition as the between-subjects variable and the two extreme deadlines as the repeated measure. As predicted, the estimated value of *C* was reliably greater in the bright-colors condition than in the dull-colors condition, $F(1, 56) = 5.27$, $MS_e = .054$. Thus our manipulation of "muddying" the colors reduced the probability that color-naming processes would contribute to responding within the deadline. It is important to note that the color manipulation had no effect on estimated values of word-reading processes ($F < 1$). Neither measure interacted with deadline ($Fs < 1$).

Inspection of Figure 3 reveals that word-reading processes often contributed to responding at the earliest deadlines. Over time, the estimated contribution of *W* declined, $F(1, 56) = 132.78$, $MS_e = .010$, and the estimated contribution of *C* increased, $F(1, 56) = 232.64$, $MS_e = .026$. There was no tendency toward an interaction between condition and deadline for either *C* or *W* (both $Fs < 1$).

As indicated earlier, one cannot assume that control items provide a process-pure measure of color-naming processes. However, *C* (an estimate of the contribution of color-naming processes to performance on the words) should nonetheless predict performance on the nonletter control items, albeit imperfectly, because a substantial proportion of the variability in performance on control items should be due to individual differences in the efficacy of color-naming processes. As predicted, the estimated contribution of color-naming processes to performance on the congruent and incongruent words predicted performance on the nonletter items: $r(56) = .74$ and $r(56) = .69$ at the 600-ms and 800-ms post hoc deadlines, respectively. As we would expect, there was no

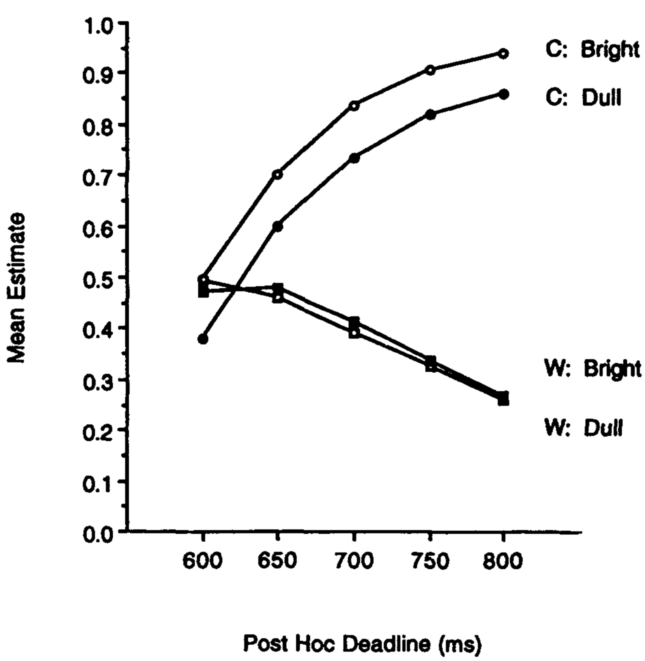

Figure 3. Experiment 2: Mean estimated contributions of word reading (*W*) and color naming (*C*) to performance at each post hoc deadline in the bright- and dull-colors conditions.

correlation between our estimate of the contribution of word-reading processes and performance on the nonletter items: $r(56) = -.01$ and $r(56) = -.21$ at the 600-ms and 800-ms post hoc deadlines, respectively, both *ns*.

Summary. Degrading the color of the stimuli in a Stroop task reduced the estimated contribution of color naming to responding but had no effect on the estimated contribution of word reading. Furthermore, reducing the contribution of color naming to performance eliminated the typical asymmetry between facilitation and interference, regardless of whether performance was measured in terms of latency (Experiment 1) or in terms of accuracy (Experiment 2). Finally, our estimate of the contribution of color-naming processes to performance on the congruent and incongruent words predicted performance on the nonletter control items. These findings support our procedure and the hypothesis that word reading and color naming operate as independent processes to determine accuracy on the Stroop task.

Manipulating the Influence of Word Reading

Experiment 3

Experiment 2 supported our parallel processing equations by showing that the stimulus-color manipulation affected the estimated contribution of color-naming processes but had no effect on the estimated contribution of word-reading processes. Experiments 3 and 4 were designed to demonstrate the opposite process dissociation: a manipulation that alters the estimated contribution of word-reading processes but has no effect on the estimated contribution of color-naming processes.

Several researchers have demonstrated that increasing the proportion of incongruent trials reduces Stroop effects measured in terms of response latency (e.g., Logan & Zbrodoff, 1979; Lowe & Mitterer, 1982). This finding suggests that when most items are incongruent, subjects somehow inhibit the influence of word-reading processes, relative to when most items are congruent. Therefore, in Experiments 3 and 4 we predicted that manipulating the proportion of congruent items would affect estimates of the contribution of word-reading processes but would not affect estimates of the contribution of color-naming processes. Such results would provide converging support for the hypothesis that word reading and color naming can operate as independent processes to determine responding and further evidence that our technique can document that independence.

Method

Subjects. The subjects were 28 undergraduate students at McMaster University who volunteered to participate for extra credit in an introductory psychology course.

Materials and procedure. The experiment was run on an IBM PC with a VGA color monitor, using Schneider's (1988) Micro-Experimental Laboratory (MEL) software system. The stimuli were the words *black, blue, green, red,* and *yellow*, and strings of percentage signs (%%%% and %%%%%). On each trial, one of these stimuli was presented in one of the five colors in the center of a light-grey computer screen. There were 40 practice trials followed by 140 test trials. For half of the subjects, 100 of the test trials were congruent and 20 were incongruent, whereas for the remaining subjects the proportions were reversed. There were 20 nonletter control trials for all subjects, 4 in each of the five colors (2 consisting of a string of four percentage signs and 2 consisting of a string of five percentage signs). The items were presented in a new random order for each subject. Each color and each word occurred equally often, and within conditions each color–word combination occurred equally often. The practice trials were randomly selected from the appropriate pool for each subject and so had approximately the same proportions of trials of each type. To equate the number of items in each condition, all statistical analyses used data from 60 preselected items that were representative of the entire set of items: 20 congruent color names (four occurrences in each of the five colors), 20 incongruent color names (the four possible incongruent versions of each of the five colors), and the 20 nonletter control items. These 60 critical items were randomly embedded within the 140 test items.

Subjects were instructed to say the color of each item aloud as quickly as they could and before an error tone sounded. The experimenter did not look at the screen during the test and so did not know what type of item was presented on any given trial. On each trial, the experimenter pressed the space bar on the computer keyboard when the subject said a color name, and then pressed another key to code the subject's response.[3] An error tone sounded if the space bar was not pressed within 800 ms. This was the same deadline as used in Experiment 2, but it was functionally much shorter because the experimenter had to detect the subject's response and press the space bar within the deadline in order for the response to be scored as within deadline. The screen went blank when the space bar was pressed or the deadline elapsed. The next item was presented 2 s after the experimenter keyed in the subject's response.

Results and Discussion

Accuracy. Mean proportions correct within the deadline were analyzed in a mixed-model ANOVA, with condition (most congruent vs. most incongruent) as the between-subjects variable and item type (congruent vs. incongruent vs. nonletter control items) as the within-subjects variable (see Table 2). Consistent with our predictions, there was a significant interaction between these variables, $F(2, 52) = 8.61$, $MS_e = .015$. Tests using Bonferroni-corrected probabilities showed that on incongruent items, subjects in the most incongruent condition did much better than subjects in the most congruent condition, $F(1, 26) = 12.63$, $MS_e = .034$, whereas on congruent items subjects in the most incongruent condition did slightly (although not significantly) more poorly than subjects in the most congruent condition ($F < 1$). The near-ceiling level of performance on congruent items may have obscured the tendency for subjects in the most congruent condition to benefit more from congruent items than subjects in the most incongruent condition. Thus, as others have found, the Stroop effect was reduced when most of the items were incongruent, indicating that subjects had some degree of strategic control over the influence of word-reading processes. It is important to note that the most congruent and most incongruent conditions did not reliably differ in the proportion correct on the nonletter control items ($F < 1$). This finding is important because it is consistent with our hypothesis that the proportion congruent manipulation would affect the contribution of word reading but not that of color naming.

Table 2
Raw Proportions Correct Within Deadline: Experiments 3 and 4

Measure	Congruent *M*	Congruent *SD*	Incongruent *M*	Incongruent *SD*	Nonletter *M*	Nonletter *SD*
Experiment 3						
Most congruent	.93	0.10	.37	0.20	.79	0.19
Most incongruent	.91	0.13	.61	0.17	.85	0.14
Experiment 4 (count dots)						
Most congruent	.80	0.15	.27	0.11	.62	0.21
Most incongruent	.68	0.23	.41	0.19	.67	0.23

Table 3
Mean Estimates (and Standard Deviations) of the Contributions of Color-Naming and Word-Reading Processes: Experiments 3 and 4

Measure	Color naming *M*	Color naming *SD*	Word reading *M*	Word reading *SD*
Experiment 3				
Most congruent	.84	0.22	.56	0.22
Most incongruent	.88	0.17	.30	0.16
Experiment 4 (count dots)				
Most congruent	.59	0.20	.53	0.18
Most incongruent	.58	0.25	.27	0.15

Estimates. Separate one-way ANOVAs were performed on the estimated contributions of word-reading and color-naming processes (see Table 3). As predicted, the estimated value of *C* did not differ in the two conditions ($F < 1$), but the estimate of *W* was higher in the most congruent condition than in the most incongruent condition, $F(1, 26) = 12.61$, $MS_e = .037$. Subjects in the most incongruent condition inhibited the effects of word-reading processes, and they did so without altering the contribution of color-naming processes.

The Pearson product–moment correlation coefficient for the relationship between *C* and the proportion correct on the nonletter control items, across subjects, was $r(26) = .77$. Thus, our estimate of *C* accounted for 59% of the variation in performance on the nonletter items. The relationship was equally strong in the most congruent ($r[12] = .76$) and most

[3] As others have noted (e.g., Whittlesea, 1993), this procedure for measuring response latencies has advantages over the use of a voice key in that voice keys are sometimes tripped by nonresponse noises and sometimes fail to detect responses. Given that the experimenter was blind to item type and the hypotheses under test were relatively complex, we think it very unlikely that experimenter bias affected recording of naming latencies. Furthermore, as noted in the text, the obtained pattern of results mirrored that reported by other investigators who conducted a similar proportion congruent manipulation and measured response latency with a voice key.

incongruent ($r[12] = .77$) conditions. Furthermore, the absolute values of C and of the proportion correct on nonletter control strings were very close to one another. As we would expect, there was no correlation between W and the proportion correct on the nonletter items ($r[26] = -.17$, *ns*).

Summary. The process dissociation obtained in Experiment 3—a large effect of the proportion congruent manipulation on estimated W and no effect on estimated C—supports our procedure and the theory implied by our equations (i.e., that word reading and color naming make independent contributions to performance). The correlation between estimated C and performance on the nonletter control items provides further support for our approach: The estimate of the contribution of color-naming processes to performance on the congruent and incongruent words predicted performance on the nonletter items.

It is clear that the performance of subjects in the most incongruent condition was less affected by word-reading processes than was that of subjects in the most congruent condition. What did these subjects do to inhibit the effects of word-reading processes? In debriefing, subjects in the most incongruent condition often indicated that they had degraded the visual clarity of the words (e.g., by squinting their eyes or looking at the bottom of the screen). Such strategies could effectively reduce the input to word-reading processes (e.g., blur letter features) while leaving the input to color-naming processes intact, hence accounting for the process dissociations we obtained. We designed Experiment 4 to prevent use of such crude input-degradation tactics so as to determine whether subjects are capable of a more central inhibition process and, if so, whether more central inhibition of the effects of word-reading processes can be independent of the contribution of color-naming processes.

Experiment 4

Method

Subjects. The subjects were 30 undergraduate students at McMaster University who volunteered to participate for extra credit in an introductory psychology course.

Materials and procedure. The materials and procedure were identical to Experiment 3, except that each test item was accompanied by zero to four small dots (periods) placed immediately to the right, left, above, or below the test item, and subjects were required to report the number of dots after naming the color of the item. Our rationale was that, because the dots were very small and very close to the test item, the task would force subjects to focus on the test item (i.e., they could not look at the bottom of the screen or grossly blur their vision and still count the number of dots). The dots were the same color as the item. Each number of dots appeared equally often with each color–word combination. The screen was blanked as soon as the color-naming response was recorded by the experimenter pressing the space bar or the 800-ms response deadline elapsed.

Results and Discussion

Dot-counting performance. Performance on the dot-counting task was very accurate, with an overall mean proportion correct of .97. There was a small but reliable main effect of item type, $F(2, 56) = 5.74$, $MS_e = .001$, with performance being best on the congruent items ($M = .99$) and worst on the incongruent items ($M = .96$). There was also a reliable effect of condition, $F(1, 28) = 8.51$, $MS_e = .002$, with performance being better in the most congruent condition ($M = .98$) than in the most incongruent condition ($M = .96$). Finally, there was a nonsignificant tendency for these two variables to interact, $F(2, 56) = 2.47$, $MS_e = .001$, $p < .10$, such that the difference between dot-counting performance on the congruent and incongruent items tended to be slightly smaller in the most congruent condition than in the most incongruent condition (most likely because of a ceiling effect in the most congruent condition). It is important to note that none of these effects compromises our interpretation of the primary analyses reported later. It makes sense that the dot-counting task would be harder to perform for subjects in the most incongruent condition than for those in the most congruent condition because the former subjects were more heavily taxed by the color-naming task. The important point is that subjects did, in fact, focus on test items with sufficient visual clarity to note the number of dots.

Accuracy. The pattern of results for the color-naming task was the same as that obtained in Experiment 3, except that performance was somewhat depressed (just as one would expect because of the additional task demand of reporting the number of dots). Analysis of the raw proportion correct revealed a significant interaction between condition and item type, $F(2, 56) = 8.82$, $MS_e = .015$ (see Table 2). Tests with Bonferroni-corrected probabilities showed that on the incongruent items, performance was better in the most incongruent condition than in the most congruent condition, $F(1, 28) = 5.84$, $MS_e = .025$, whereas on the congruent items there was a nonsignificant tendency for performance to be better in the most congruent condition than in the most incongruent condition, $F(1, 28) = 2.89$, $MS_e = .037$, $p < .11$. Thus, as in Experiment 3, subjects in the most incongruent condition suffered less cost when word and color were incongruent, and showed a nonsignificant tendency to gain less benefit when they were congruent, than subjects in the most congruent condition. Also consistent with Experiment 3, the two groups did not differ on performance on the nonletter items ($F < 1$), suggesting that the effect of the proportion congruent manipulation was on the contribution of word reading, not color naming.

Estimates. The pattern of results for the algebraically derived estimates of the contributions of word-reading and color-naming processes to performance also mirrored those of Experiment 3, except that estimates of color naming were lower because of interference from the secondary task (see Table 3). The proportion congruent had no effect on estimates of C ($F < 1$), but estimates of W were much higher in the most congruent condition than in the most incongruent condition, $F(1, 28) = 18.61$, $MS_e = .027$. Thus, subjects in the most incongruent condition were able to suppress the effects of word-reading processes without interfering with color-naming processes.

As in Experiments 2 and 3, we examined the correlation between the estimated contribution of color-naming processes to performance on the words and proportion correct on the nonletter controls. Replicating the earlier findings, estimates of C predicted performance on the nonletter controls ($r[28] = .70$; in the most congruent condition, $r[13] = .67$, and in the most incongruent condition, $r[13] = .75$). As we would expect, and consistent with our earlier findings, there was no reliable relationship between our estimate of W and performance on the nonletter items ($r[28] = .24$, *ns*).

Summary. The results of Experiments 3 and 4 support our procedure and the hypothesis that word reading and color naming can act as parallel processes to determine Stroop performance. Consistent with earlier findings using latency measures, increasing the proportion of incongruent items reduced Stroop effects with our accuracy measure. Moreover, this manipulation produced a clean process dissociation: The estimated contribution of word-reading processes to responding was substantially reduced in the most incongruent condition relative to the most congruent condition, but the estimated contribution of color-naming processes to responding was unaffected by the manipulation. Converging support for our approach is provided by the finding that our estimate of C, algebraically derived from performance on the congruent and incongruent words, predicted performance on the nonletter items.

General Discussion

The research reported here demonstrates how the process dissociation procedure (Jacoby, 1991) can be used to obtain separate quantitative estimates of the contributions of word-reading and color-naming processes to accuracy in Stroop tasks. Our results provide compelling evidence for the hypothesis that word reading and color naming can act in parallel to determine responses. The process dissociation procedure has implications for the definition of automaticity. Our approach also provides important insight into the relationship between Stroop interference and facilitation. Perhaps most important, our results indicate that the influence of automatic processes cannot accurately be measured in terms of the difference between performance on primed or inhibited items and performance on baseline items. Each of these issues is discussed in the following subsections.

Word Reading and Color Naming as Independent Processes

Process Dissociations as Evidence of Process Independence

An anonymous reviewer of an earlier version of this article commented that our rationale seemed circular: The equations, which assume that word reading and color naming make independent contributions to performance, were used to garner evidence of that independence. In fact, however, our argument is not circular. The rationale for our approach is to use a manipulation that, on a priori and empirical grounds, can be hypothesized to affect one kind of process but not the other: In this context, the fact that the manipulation dramatically affects one of our estimates but has no effect on the other constitutes evidence that the two processes are indeed independent and that the equations provide meaningful estimates of their contributions to performance.

For example, in line with others (e.g., Logan & Zbrodoff, 1979; Lowe & Mitterer, 1982; Tzelgov, Henik, & Leiser, 1990), we proposed that subjects inhibit the influence of word-reading processes when most of the stimuli in a Stroop color-naming task are incongruent items. Consistent with this, subjects in the most incongruent condition demonstrated less interference and less facilitation than those in the most congruent condition. It is important to note that color-naming performance on the nonletter control items was equivalent in the two conditions, supporting the claim that the manipulation affected the influence of word reading but not that of color naming. In that context, the fact that estimates of the influence of word reading were dramatically lower in the most incongruent than in the most congruent condition, whereas estimates of the influence of color naming were equivalent in the two conditions, provides strong support for our method and for the estimates derived with our equations.

It is important to understand that the process dissociations we obtained are not artifacts of the equations. That the equations do not compel a particular pattern of findings is demonstrated by the fact that the nature of the process dissociation was determined by the nature of the manipulation: Whereas the proportion congruent manipulation affected estimates of W but not estimates of C, the stimulus color manipulation affected estimates of C but not estimates of W.[4]

Converging evidence for our approach comes from the finding that our estimate of the influence of color-naming processes predicted performance on the nonletter control items quite well (e.g., $r = .77$ in Experiment 3). The number of subjects in each experiment was relatively small, and the estimate of C was derived from performance on a relatively small number of congruent and incongruent words and was used to predict performance on a relatively small number of nonletter items. In view of the likely measurement error in this situation, we think this is an impressive level of prediction.

Alternative Accounts

Critics could argue that there are, in fact, dependencies between word-reading and color-naming processes (e.g., that the effect of word reading is greater on incongruent items than on congruent items) and claim that our evidence of process dissociations is really just a quirk of chance. By such a view, the appropriate equation for performance on congruent items would be $W_c + C(1 - W_c)$ and that for performance on incongruent items would be $W_i + C(1 - W_i)$, that is, the contribution of W would differ for congruent and incongruent

[4] Of course, one would expect that some manipulations would affect estimates of the contributions of both word-reading and color-naming processes, not because the two are dependent on one another but simply because they are both affected by that manipulation. For example, in pilot research we found that dimming the illumination on the computer screen lowered our estimates of the influence of both color naming and word reading.

items. Of course, these equations cannot be solved by applying algebra to performance on congruent and incongruent items—there are three unknowns and only two equations. But a critic could argue that these three-parameter equations represent what is really going on in the Stroop task and point out that there are an infinite number of possible values of C, W_c, and W_i that could be fitted to any given pattern of performance on congruent and incongruent items.

The crux of such an argument would be that, by chance, we hit on conditions under which our equations produce clean process dissociations (i.e., findings that one estimate differed across conditions whereas the other remained invariant) even though the core assumptions of the equations are false. It is true that this is possible (e.g., one can produce post hoc estimates for three-parameter dependent-process equations that fit our findings). However, for this to happen the levels of the three parameters would have to be in a delicate balance such that the dependency perfectly offset a true difference (e.g., a difference in color naming produced by the proportion–congruent manipulation). It is impossible to calculate the chances of such a happy accident, but we believe that the odds against such serendipity occurring in three different experiments are very small. The plausibility of such dependent-process accounts is further weakened by the fact that our estimates of C predicted performance on nonletter items; if there were strong dependencies between word reading and color naming, our estimates of C would be grossly distorted and therefore unlikely to predict performance on nonletter items. Thus the current findings are strong evidence against dependent-process accounts of Stroop effects.

We do not claim that our estimates are perfect nor that the hypothesized independence between word reading and color naming is 100% complete. It is possible that there is some degree of dependence between W and C (e.g., that the contribution of word reading is slightly greater on incongruent than on congruent items). Our results suggest that if there are such dependencies, they are small, and if our estimates are biased, the bias is slight. Occam's razor would favor the simpler equations unless and until they are shown to be inadequate. Undoubtedly, it will eventually be necessary to complicate our simple model. However, as further discussed later, that simple model is a substantial advancement over measuring facilitation and interference in comparison to baseline, and it has allowed us to gain important new insights into Stroop phenomena.

Measuring Stroop Effects in Terms of Accuracy Rather Than Latency

As noted in the introduction, a potential concern about our method is that our accuracy-within-deadline measure might tap processes that are qualitatively different from those that underlie Stroop effects on latency measures (cf. Pashler, 1989; Santee & Egeth, 1982). We see little grounds for this concern. Because our response deadlines were relatively long and subjects had to respond before the stimulus was removed from the screen to be scored as correct, our accuracy measure is likely to be sensitive to response selection mechanisms, just as latency measures are, rather than to very early perceptual processing mechanisms. Moreover, we obtained exactly the same pattern of results in our color-degrading experiments regardless of whether a standard response-latency measure (Experiment 1) or an accuracy-within-deadline measure (Experiment 2) was used; in addition, our pattern of results with the proportion congruent manipulation (Experiments 3 and 4) mirror findings others have obtained with that manipulation and a standard response-latency measure. Ultimately, of course, the proof will be in the pudding. Perhaps we will find that, in some cases, Stroop effects measured in terms of accuracy are qualitatively different from Stroop effects measured in terms of response latency. Such findings would be highly informative.

In summary, our results are compatible with theories of Stroop interference that describe word reading and color naming as parallel processes (e.g., Cohen et al.'s [1990] parallel distributed processing [PDP] model, Logan's [1980] evidence accrual model, and Posner's [1978] automaticity model; see also Dempster, 1990; Melara, 1990; Tzelgov et al., 1990), and our results conflict with models that describe word reading and color naming as interactive processes. For example, our finding that the proportion congruent manipulation had pronounced effects on the contribution of word reading but no effect on the contribution of color naming challenges Hock and Egeth's (1970) hypothesis that word meanings affect perception of color.

Our studies were not directed toward discriminating between the different parallel processing models that have been proposed (e.g., our results do not indicate whether Logan's [1980] evidence accrual model is more or less useful than Cohen et al.'s [1990] PDP model). What our approach offers that previous research does not is a way to measure the independent processes that contribute to performance on the Stroop task. Future research using the process dissociation technique may enable researchers to select between and further refine particular parallel processing models.

Automaticity

The process dissociation procedure defines automaticity in contrast to control: A process is automatic to the extent that its influence is the same regardless of whether it helps performance (as in congruent Stroop items) or hinders performance (as in incongruent items). That is, a process is automatic in a particular situation to the extent that it is as likely to affect performance when the person wants it to as when the person wants it not to. Automaticity is not an inherent property of particular kinds of stimuli, tasks, or processes; for example, reading is not automatic in a generic sense. Rather, automaticity is an emergent property of particular constellations of stimuli, tasks, and processes (cf. Allport, 1989; Neumann, 1984). For example, in Experiments 3 and 4 the influence of automatic word reading played a larger role in the most congruent conditions than in the most incongruent conditions.

Another important point is this: The fact that two processes can be performed in parallel does not imply that either of them can be performed in parallel with some third process.

The estimated contribution of color-naming processes was smaller in Experiment 4 than in Experiment 3 because color-naming processes were impaired by the secondary task used in Experiment 4. On the other hand, the secondary task had no effect on estimated word-reading processes. This indicates that word reading and dot counting were performed as parallel processes but that color naming and dot counting were not. Because color naming is much less practiced than word reading, one would expect a variety of secondary tasks to interfere with color naming but not with word reading, just as our results indicate. It is important that word reading and color naming were performed in parallel (as indicated by the finding that the proportion congruent manipulation affected word reading but not color naming) even though the secondary task interfered with color naming (as indicated by poorer overall performance and lower estimates of C in Experiment 4 than in Experiment 3).

Our findings extend those of others (e.g., Logan & Zbrodoff, 1979; Lowe & Mitterer, 1982; Tzelgov et al., 1990) in demonstrating that subjects can exert a remarkable degree of control over the effects of automatic word-reading processes in the Stroop task (cf. Allport, 1989; Logan & Cowan, 1984; Neill, 1977; Neumann, 1984). In Experiments 3 and 4 the estimated contribution of word-reading processes to performance in the most congruent condition was approximately halved in the most incongruent condition. Our results indicate that this control is not limited to peripheral input-degradation tactics (e.g., grossly blurring one's vision or focusing away from the stimuli). Rather, these data suggest central mechanisms by which subjects can inhibit the influence of word-reading processes without affecting the influence of color-naming processes.

The Relationship Between Facilitation and Interference

MacLeod (1991) suggested that Stroop interference and facilitation may arise from different mechanisms. In the following sections we examine the relation between facilitation and interference and argue that, rather than the two arising from separate mechanisms, facilitation and interference reflect the independent contributions of separate bases for responding. This perspective has important implications for research on inhibition and priming in a wide variety of research domains.

Asymmetry Between Facilitation and Interference

Stroop interference is typically much greater than Stroop facilitation. Our approach provides an independent-processes account of this asymmetry. We argue that it is precisely the independence of word-reading and color-naming processes (as in our equations) that gives rise to asymmetrical effects of congruence and incongruence around control performance. As explained in the introduction to Experiment 1, given a relatively large contribution of color-naming processes (which one would generally expect on a color-naming task), the equations predict that the difference between incongruent and control items (interference) will be greater than that between congruent and control items (facilitation). This is illustrated in Figure 1, which plots the proportion correct on congruent and incongruent items as a function of various levels of contribution of color-naming and word-reading processes. So long as the contribution of color-naming processes exceeds .50, interference will be greater than facilitation.

This asymmetry between facilitation and interference is evident in the data from Experiments 3 and 4. In the most congruent condition of Experiment 3, for example, relative to nonletter control items (mean proportion correct = .85) there was substantial interference on incongruent items (M = .61) but small and nonsignificant facilitation on congruent items (M = .91). This is not because word-reading processes did not contribute to color-naming performance on the congruent items; on the contrary, the estimated contribution of word reading to color naming was .30. Rather, larger interference than facilitation is a natural consequence of the independent relation between word-reading and color-naming processes when color-naming processes make relatively large contributions to performance.

Experiments 1 and 2 offered direct evidence on this point. Experiment 2 demonstrated that degrading the color of the stimuli lowered the estimated contribution of color-naming processes to performance without affecting the estimated contribution of word-reading processes. Furthermore, as predicted by our equations, the typical finding of greater interference than facilitation (which was replicated in the bright-colors condition) was eliminated in the dull-colors condition in both experiments: As the contribution of color-naming processes to performance decreases, facilitation increases relative to interference. Thus asymmetries between interference and facilitation are easily accounted for by parallel process models.

Correlation Between Facilitation and Interference

In his review of an earlier version of this article, C. M. MacLeod (personal communication, October 26, 1992) noted that a graduate student in his lab, Marina Vanayan, in her dissertation research, found no correlation between interference and facilitation. MacLeod argued that this suggests that different mechanisms underlie facilitation and interference. This prompted us to calculate the correlation between interference and facilitation in the experiments reported here. Those values are reported in Table 4.

There was a consistent tendency for facilitation and interference to be negatively correlated, and the relationship was statistically reliable whenever performance on the nonletter controls was low enough to allow for substantial facilitation. This suggests that one possible reason Vanayan (cited by C. M. MacLeod, personal communication, October 26, 1992) found no correlation between facilitation and interference is that ceiling effects may have obscured the relationship.

Why were facilitation and interference negatively, rather than positively, correlated in our experiments? Our equations suggest that this counterintuitive relationship will occur when the influence of color-naming processes varies more

Table 4
Correlations Between Interference and Facilitation Across Experiments

Measure	Correlation
Experiment 1 (latency)	
Bright colors	−.27*
Dull colors	−.47*
Experiment 2 (600-ms deadline)	
Bright colors	−.62*
Dull colors	−.50*
Experiment 3	
Most congruent	−.30
Most incongruent	−.30
Experiment 4	
Most congruent	−.60*
Most incongruent	−.61*

* Pearson correlation significant at the .05 level.

from subject to subject than the influence of word-naming processes does. According to our equations, if performance on the control items is $C^{\wedge}$ then interference $= C^{\wedge} - C(1 - W)$ and facilitation $= [W + C(1 - W)] - C^{\wedge}$. If $C^{\wedge}$ is assumed to be equivalent to C (i.e., if C is a perfect estimate and $C^{\wedge}$ a perfect control), then if W remains relatively constant across subjects, those with relatively high values of C will tend to have higher interference and lower facilitation than those with lower values of C. For example, if $W = .6$ and $C^{\wedge} = C$, then:

If $C = .8$, then facilitation $= [.6 + .8(.4)] - .8 = .12$,
and interference $= .8 - .8(.4) = .48$.
If $C = .7$, then facilitation $= [.6 + .7(.4)] - .7 = .18$,
and interference $= .7 - .7(.4) = .42$.

Thus, as C decreases when W is held constant, facilitation increases (from .12 to .18) and interference decreases (from .48 to .42). Therefore, variability from person to person in the contribution of color-naming processes to Stroop task performance will tend to produce negative correlations between facilitation and interference if there is less variability in the contribution of word-reading processes. On the other hand, facilitation and interference will tend to be positively correlated if the contribution of word reading varies considerably and that of color naming is relatively stable across individuals. For example, if $C = .6$, then:

If $W = .8$, then facilitation $= [.8 + .6(.2)] - .6 = .32$,
and interference $= .6 - .6(.2) = .48$.
If $W = .7$, then facilitation $= [.7 + .6(.3)] - .6 = .28$,
and interference $= .6 - .6(.3) = .42$.

Thus, as W decreases when C is held constant, facilitation decreases (from .32 to .28) and interference also decreases (from .48 to .42). Therefore, variability from person to person in the contribution of word-reading processes to Stroop task performance will tend to produce positive correlations between facilitation and interference if there is less variability in the contribution of color-naming processes. If both of these independent bases for responding vary to the same extent, then no correlation will emerge between facilitation and interference.

According to our equations, then, our finding of negative correlations between facilitation and interference should be due to greater variability from subject to subject in the contribution of color-naming processes than in the contribution of word-reading processes. Consistent with this, estimates of W in each condition were virtually identical in Experiments 3 and 4. Furthermore, as shown in Table 5, the variance of C was greater than the variance of W in Experiments 2 and 4, that is, in the deadline experiments in which interference and facilitation were reliably negatively correlated. Thus, just as the equations suggest, the negative correlations between facilitation and interference in these experiments reflects the fact that there was greater variability in C than in W. In other situations, W and C may be equally variable (and hence there will no correlation between facilitation and interference, as in Experiment 3). One might also find situations in which there is greater variability in W than in C (and hence a positive correlation between facilitation and interference).

Evidence Against Separate Mechanisms for Facilitation and Interference

Thus far, we have explained how our approach can easily account for findings that have been offered as challenges to parallel processing models of Stroop phenomena. We have shown that the typical asymmetry between facilitation and interference is a direct consequence of the independence of word reading and color naming as bases for responding, and we have offered an account of variations in the correlation between facilitation and interference. Having responded to these challenges, we now highlight aspects of our results that directly conflict with the hypothesis that Stroop interference reflects mechanisms that are different from those that underlie Stroop facilitation (e.g., an interactive mechanism that impairs color-naming processes on incongruent items).

Our findings provide clear evidence against such separate-mechanisms hypotheses, suggesting instead that both facilitation and interference reflect the independent contributions of word-reading and color-naming processes to performance. For example, subjects in the most incongruent conditions of

Table 5
Variances in Estimates of Color Naming and Word Reading Across Experiments and Conditions

Measure	Estimate	
	Color naming	Word reading
Experiment 2		
Bright colors	.077	.015
Dull colors	.051	.022
Experiment 3		
Most congruent	.046	.048
Most incongruent	.027	.026
Experiment 4		
Most congruent	.041	.031
Most incongruent	.063	.023

Experiments 3 and 4 reduced interference by inhibiting the effects of word reading, but by doing so they lowered facilitation as well as interference. If interference reflected a special mechanism that operates only when word and color are incongruent, the proportion congruent manipulation would be expected to have different effects on congruent and incongruent items and hence different effects on facilitation and interference. Furthermore, if a special mechanism affected color naming on incongruent items, our equations would not work: *C* would not be the same on congruent and incongruent items and thus our algebraically derived estimates would be distorted. Therefore, the fact that we obtained very clean process dissociations (e.g., that the stimulus-color manipulation affected estimates of *C* but not estimates of *W*) and that our estimates of *C* predicted performance on nonletter items is strong evidence against the notion that interference is due to special mechanisms that operate only on incongruent items. Instead, interference, like facilitation, reflects the independent contributions of word-reading and color-naming processes.

Measuring Priming and Inhibition

Our results indicate that neither priming nor inhibition can be accurately measured by comparing them to baseline. Consider, for example, the results of Experiments 1 and 2. If the facilitating effect of word reading on congruent items were indexed by subtracting performance on the nonletter items from performance on the congruent items, one would conclude that the facilitating effect of word reading was greater in the dull-colors condition than in the bright-colors condition. In contrast, if one indexed inhibition of word-reading processes in terms of the difference between incongruent and nonletter items (on the rationale that total inhibition would lead to equal performance on incongruent and nonletter items), one would conclude that inhibition of word reading was greater in the dull-colors condition than in the bright-colors condition because interference was less in the dull-colors condition. Of course, those effects do not reflect a difference in the contribution of word-reading processes to performance in the dull- and bright colors conditions—our estimate of the effects of word-reading processes were nearly identical for the two conditions; it was color-naming processes that differed in the two conditions and were responsible for the differences in facilitation and interference.

The important point here is that the effect one process has on performance depends on the level of the other process. That is why the amounts of facilitation and interference differed so markedly in the two conditions of Experiment 2, even though the contribution of word reading to responding was the same in both conditions. Likewise, that is why interference was much greater in Experiment 4 than in Experiment 3, even though the contribution of word reading to responding was the same in those two experiments. As illustrated in Figure 1, variations in the contribution of color-naming processes to performance mediate the effects of word-reading processes (and vice versa).

Thus, our results show that facilitation is nothing like an accurate index of the effects of word reading on congruent Stroop items and that interference is nothing like an accurate index of inhibition of the effects of word reading on incongruent Stroop items. The same problems may arise in studies of other kinds of facilitation and inhibition effects.

Cognitive psychology includes a massive literature on priming (e.g., Meyer & Schvaneveldt, 1971; Morton, 1969; Neely, 1976; Posner & Snyder, 1975b) and a largely separate literature on inhibition (e.g., Bjork, 1989; Tipper, 1992). The standard technique for measuring priming and inhibition effects in a wide variety of domains is to compare items in a facilitation or interference condition to items in a supposedly neutral baseline condition. The resultant measure is taken as an index of the effect of the priming or inhibiting process.

Facilitation and interference effects in many domains can be described as analogous to Stroop effects, that is, as situations in which two independent bases for responding contribute to performance. For example, regardless of whether one conceptualizes priming as reflecting the influence of episodic memories (e.g., Ratcliff & McKoon, 1988) or as reflecting residual subthreshold activation of a logogenlike representation (e.g., McNamara, 1992), performance on primed items in lexical decision tasks may be described as resulting from the simultaneous contributions of memory for the priming event and of current word-reading processes.

If this independent-processes view is correct, priming and inhibition cannot be accurately assessed by comparing performance on primed or inhibited items with performance on baseline items. Differences in performance on baseline items within or across experiments will lead to different measures of facilitation and inhibition even when the effects of the facilitating or inhibiting process do not actually differ (cf. Jonides & Mack, 1984). This problem cannot be solved by simply subtracting out baseline or using it as a covariate. Performance on primed items and on inhibited items reflects the joint contributions of independent processes, so one cannot measure priming or inhibition without having a means of separately estimating both contributing processes. The process dissociation procedure provides such a means.

The process dissociation procedure could also be used to investigate deficits in inhibition. For example, it has been proposed that the efficacy of inhibitory processes declines with age (e.g., Hasher, Stoltzfus, Zacks, & Rypma, 1991; Tipper, 1991). One finding cited as support for that claim is that the elderly suffer greater Stroop interference than do younger subjects (e.g., Panek, Rush, & Slade, 1984). As we argued earlier, the difference between performance on incongruent and control items always reflects the independent contributions of word-reading and color-naming processes: A given degree of inhibition of word-reading processes can cause either slight or great interference, depending on the level of the contribution of color-naming processes. Consequently, a finding of greater interference for the elderly could be because of a deficit in color-naming processes, in the ability to inhibit the effects of word reading, or some combination of the two. Therefore, inhibitory mechanisms may well decline with age, but that decline should be measured as a change in the effects of the to-be-inhibited process (e.g., our estimate of W) rather than as a change in the amount of interference observed.

In summary, the standard approach has been to measure facilitation or interference relative to a supposedly neutral control condition. Our results show that the standard approach will not work. First, there are serious problems surrounding attempts to find a neutral control condition. Second, and more important, facilitation and interference do not always reflect the operation of different mechanisms. Rather, as in our experiments, interference and facilitation can arise from the contributions of the same two independent bases for responding. To reveal that independence, it is necessary to compare interference and facilitation conditions, as we have done with the process dissociation procedure, so as to estimate the contributions of different processes to responding. Doing otherwise runs the risk of creating confusing and conflicting results because of the inability to take differences in baseline into account.

Summary and Conclusions

By using the process dissociation procedure, we obtained separate quantitative estimates of the contributions of word-reading and color-naming processes to accuracy of responding on the Stroop task. We then used the procedure to explore substantive issues in the Stroop literature. Our results indicate that word reading and color naming can act in parallel to determine responses: Degrading the color of Stroop stimuli decreased the influence of color-naming processes but did not affect the influence of word-reading processes, and increasing the proportion of incongruent items led subjects to substantially reduce the influence of word-reading processes on performance without affecting the influence of color-naming processes. The pattern of results across experiments amounts to a double dissociation between word-reading and color-naming processes. Furthermore, our algebraically derived estimates of the contribution of color-naming processes to performance on the words predicted performance on the nonletter control items. We have also shown that parallel process models can easily account for the typical asymmetry between Stroop interference and facilitation and can be used to understand correlations between these two measures. Our procedure could also be used to determine whether training programs or group differences associated with superior performance on the Stroop task are mediated by facilitation of the effects of color-naming processes, inhibition of the effects of word-reading processes, or both.

The Stroop effect is the paradigmatic example of situations in which two types of cognitive processes, one intended and the other automatic, simultaneously contribute to performance. The Stroop effect is paradigmatic because the effect is large and robust, the two classes of processes (word reading and color naming) are reasonably well defined, and the materials are easy to work with. But this situation, in which intended and automatic processes work together (in opposition or in concert) to determine performance is common to many psychological tasks (see Jacoby & Kelley, 1991; Jacoby, Lindsay, & Toth, 1992; Jacoby, Ste-Marie, & Toth, 1993). Thus the problem of isolating the contributions of controlled and automatic processes is ubiquitous in experimental psychology.

Our results illustrate the advantages of looking for process dissociations as evidence of independent processes, and they have important implications for the investigation of interference and facilitation effects. Many of the domains in which such effects have been investigated can be framed as analogous to the Stroop task, that is, as situations in which performance is determined by the contributions of independent bases for responding. One may then use the process dissociation procedure to seek evidence of that independence and, if such evidence is found, to estimate the contributions of the different processes to performance. This approach allows discoveries that would have otherwise remained elusive; for example, measuring interference relative to the nonletter items in Experiment 1 could not have revealed that the influence of word-reading processes remained invariant across changes in baseline. We think this change in perspective is important and generally applicable.

To call for a shift from the traditional latency procedures to our process dissociation procedure is to demand a major change in the way experiments are done and indicts a huge amount of past research (e.g., 20 years of research using the lexical decision task). Such a broad call for change will undoubtedly bring out critics who will question our assumptions, the most vulnerable of which is the assumed independence of processes (e.g., that W and C are independent bases for responding on the Stroop task). Careful scrutiny of our approach is necessary, and we have gone to considerable length to show that the assumption of independence is well justified in these studies. We suggest that the assumptions underlying the well-established response latency procedures (e.g., that control items provide a factor-pure baseline measure, or that it makes sense to study facilitation and interference separately) should be scrutinized just as closely (cf. Jonides & Mack, 1984).

References

Allport, A. (1989). Visual attention. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 631–682). Cambridge, MA: MIT Press.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger, III, & F. I. M. Craik (Eds.), *Varieties of memory & consciousness* (pp. 309–330). 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.

Dempster, F. N. (1990, November). *Resistance to interference: A neglected dimension of cognition.* Paper presented at the meeting of the Psychonomic Society, New Orleans, LA.

Driver, J., & Tipper, S. P. (1989). On the nonselectivity of "selective" seeing: Contrasts between interference and priming in selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 304–314.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 163–169.

Hock, H. S., & Egeth, H. (1970). Verbal interference with encoding

in a perceptual classification task. *Journal of Experimental Psychology, 83,* 299–303.
Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513–541.
Jacoby, L. L., & Kelley, C. M. (1991). Unconscious influences of memory: Dissociations and automaticity. In D. Milner & M. Rugg (Eds.), *The neuropsychology of consciousness* (pp. 201–223). San Diego, CA: Academic Press.
Jacoby, L. L., Lindsay, D. S., & Toth, J. P. (1992). Unconscious influences revealed: Attention, awareness, and control. *American Psychologist, 47,* 802–809.
Jacoby, L. L., Ste-Marie, D., & Toth, J. P. (1993). Redefining automaticity: Unconscious influences, awareness, and control. In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention, selection, awareness, and control: A tribute to Donald Broadbent* (pp. 261–282). New York: Oxford University Press.
Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin, 96,* 29–44.
Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology, 12,* 523–553.
Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91,* 295–327.
Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition, 7,* 166–174.
Lowe, D., & Mitterer, J. O. (1982). Selective and divided attention in a Stroop task. *Canadian Journal of Psychology, 36,* 684–700.
MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.
McNamara, P. T. (1992). Theories of priming: I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18,* 1173–1190.
Melara, R. D. (1990, November). *Mandatory processes: Consider your options.* Paper presented at the annual meeting of the Psychonomic Society, New Orleans, LA.
Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227–234.
Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165–178.
Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory & Cognition, 4,* 648–654.
Neill, W. T. (1977). Inhibitory and facilitatory processes in selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 444–450.
Neumann, O. (1984). Automatic processing: A review of recent findings and a plea for an old theory. In W. Prinz & A. F. Sanders (Eds.), *Cognition and motor processes* (pp. 225–293). Berlin: Springer-Verlag.
Panek, P. E., Rush, M. C., & Slade, L. A. (1984). Locus of the age–Stroop interference relationship. *Journal of Genetic Psychology, 145,* 209–216.
Pashler, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology, 21,* 469–514.
Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Erlbaum.
Posner, M. I., & Snyder, C. R. R. (1975a). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium.* Hillsdale, NJ: Erlbaum.
Posner, M. I., & Snyder, C. R. R. (1975b). Facilitation and inhibition in the processing of signals. In P. M. Rabbitt & S. Dornic (Eds.), *Attention and performance* (Vol. 5, pp. 669–682). San Diego, CA: Academic Press.
Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review, 95,* 385–408.
Santee, J. L., & Egeth, H. E. (1982). Do reaction time and accuracy measure the same aspects of letter recognition? *Journal of Experimental Psychology: Human Perception and Performance, 8,* 489–501.
Schneider, W. (1988). Micro-Experimental Laboratory: An integrated system for IBM PC compatibles. *Behavior Research Methods, Instruments, and Computers, 20,* 206–217.
Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.
Tipper, S. P. (1991). Less attentional selectivity as a result of declining inhibition in older adults. *Bulletin of the Psychonomic Society, 29,* 45–47.
Tipper, S. P. (1992). Selection for action: The role of inhibitory mechanisms. *Current Directions in Psychological Science, 1,* 105–109.
Tzelgov, J., Henik, A., & Leiser, D. (1990, November). *Controlling Stroop effect.* Paper presented at the annual meeting of the Psychonomic Society, New Orleans, LA.
Whittlesea, B. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1235–1253.

Appendix

MEL Code Specifications to Alter Stimulus Colors in Experiments 1 and 2

Bright-colors condition	Dull-colors condition
SET_PALETTE_VGA(1, 0, 0, 63) ! blue	SET_PALETTE_VGA(1, 0, 0, 24) ! blue
SET_PALETTE_VGA(2, 10, 42, 10) ! green	SET_PALETTE_VGA(2, 2, 12, 2) ! green
SET_PALETTE_VGA(4, 50, 0, 0) ! red	SET_PALETTE_VGA(4, 23, 2, 2) ! red
SET_PALETTE_VGA(6, 30, 15, 4) ! brown	SET_PALETTE_VGA(6, 20, 12, 2) ! brown

Note. The first number in the parentheses is the videographic array (VGA) label for a particular color. The second, third, and fourth numbers specify the amounts of red, green, and blue in that color. Due to an oversight the default color for black was used in both conditions.

Received August 5, 1992
Revision received May 18, 1993
Accepted May 18, 1993 ■